大学共同利用機関法人 情報・システム研究機構
統計数理研究所

第4回

統計数理研究所
NOE（Network Of Excellence）形成事業顧問会議

報告集

N O E

令和6年3月

統計数理研究所
NOE形成事業

第4回

統計数理研究所
NOE（Network Of Excellence）形成事業 顧問会議

報 告 集

■ まえがき …… 1

■ プログラム …… 2

■ 出席者名簿 …… 3

■ オンライン会議の様子 …… 4

■ 開会・出席者紹介 …… 5

■ 所長挨拶・統計数理研究所の概要 …… 7

■ 国際戦略アドバイザーからの助言への対応 …… 27

■ NOE形成事業の概要 …… 35

■ 統計的機械学習研究センター NOE 活動紹介・事業計画 … 59

■ ものづくりデータ科学研究センター NOE 活動紹介・事業計画 … 74

■ 統計数理研究所のスーパーコンピュータ …… 98

■ 統計エキスパート人材育成プロジェクト …… 111

■ 調査科学NOE 活動紹介・事業計画 …… 126

■ 総合討論および顧問お一人ずつからの総括コメント …… 150

■ 謝辞および閉会 …… 156

Network Of Excellence

まえがき

統計数理研究所では、2005（平成17）年度に北川源四郎所長（当時）のCOE（Center of Excellence）からNOE（Network of Excellence）活動へとの構想を基に、リスク解析戦略研究センターを設置し、国内外の研究者が様々な分野の研究者・実務家と共同して、統計的・数理的側面を有するリスク研究プロジェクト並びにリスクに関わる公共的活動を企画・実施することを目的として、「リスク研究ネットワーク」の運営が始まりました。その後、2010（平成22）年4月からの大学共同利用機関法人情報・システム研究機構の第2期中期目標・中期計画の中で正式に「統計数理NOEの形成」を掲げました。人材、設備環境をネットワークでつなぐことで、個別分野および横断分野のより高度な研究推進のみならず異分野融合研究や新しい研究分野の創成を展開していく「NOE（Network Of Excellence）形成事業」を始動し、以降の中期目標・中期計画の中でも、本事業を研究所がハブとなるプロジェクト型研究の中心として位置づけ、運営・推進に注力しております。

本研究所では、様々な分野を横断する基幹研究系を横軸、それをもとにプロジェクト型応用研究をする各NOE型研究センターを縦軸とした二軸構造を主軸に、研究教育活動を展開して参りました。

NOE型研究センターである「統計的機械学習研究センター」と「ものづくりデータ科学研究センター」は、これまで統計的機械学習の横断的理論研究と先端学術分野でのプロジェクト型実践研究で世界をリードする研究成果を生み出してきました。この2つのセンターは、特色的な大学共同利用・国際共同研究・産学連携ネットワークを築いており、基幹研究系と兼務する現センター教員の活動に対する社会的ニーズが急増しました。今後益々増加する共同研究（産学連携・国際共同研究、大型競争的研究資金に基づくプロジェクト等）への対応が研究所としても喫緊の課題となっていました。

そこで、本研究所は2022（令和4）年4月からの第4期中期目標・中期計画において2023（令和5）年度内の基幹研究系の改組を掲げました。所長直下の新基幹研究系として先端データサイエンス研究系の設置し、その下にこの2センターを横断型（統計的機械学習）と特定領域型（マテリアルズ・インフォマティクス研究推進）の高等研究センターとして移管設置する改組を企図したのです。

これを受けて2023年5月16日に第4回NOE形成事業顧問会議を開催いたしました。これまで、そして改組を含むこれからのNOE活動の計画や構想に係る研究所からの報告に対し、7名の顧問の先生方から非常に本質的かつ鋭い御指摘、今後の事業運営に向けての大変貴重な御助言、応援のお言葉を頂戴しました。そのおかげで本年度内に改組計画をさらに熟成し、2024（令和6）年3月1日付けで基幹研究系の改組、高等研究センターの設置を実施するに至りました。御多用の中、顧問会議に御参集くださり、さまざまな見地から御教示くださった顧問の先生方には、改めてここに深甚なる感謝の意を表します。

ここに今後の研究所の運営、研究企画の指針とすべく、顧問会議での詳細議事録を含めた『第4回統計数理研究所NOE（Network Of Excellence）形成事業顧問会議報告集』を編纂した次第です。

今後とも本研究所に対する皆様の御理解と御支援を何卒よろしくお願い申し上げます。

2024年3月

統計数理研究所長

椿　広計

第4回 統計数理研究所 NOE形成事業 顧問会議 プログラム

日　　時：令和５（2023）年５月16日（火）13時30分～17時00分
会議形式：Zoomミーティング

13：30 ～ 13：35	開会	司　会：山下　智志 （副所長／NOE 推進室長）
13：35 ～ 13：40	出席者紹介	
13：40 ～ 13：55	所長挨拶・ 統計数理研究所の概要について	所　長　椿　広計
13：55 ～ 14：05	国際戦略アドバイザーからの 助言への対応について	副所長　川崎　能典
14：05 ～ 14：25	NOE形成事業の概要について	NOE 推進室長 副所長　山下　智志
14：25 ～ 14：50	統計的機械学習研究センター NOE活動紹介・事業計画について	統計的機械学習研究センター長 教　授　福水　健次
14：50 ～ 15：00	休　憩	
15：00 ～ 15：25	ものづくりデータ科学研究センター NOE活動紹介・事業計画について	ものづくりデータ科学研究センター長 教　授　吉田　亮
15：25 ～ 15：45	統計数理研究所の スーパーコンピュータについて	統計科学技術センター長 教　授　上野　玄太
15：45 ～ 16：05	統計エキスパート 人材育成プロジェクトについて	大学統計教員育成センター長 特任教授　千野　雅人
16：05 ～ 16：25	調査科学NOE 活動紹介・事業計画について	データ科学研究系調査科学グループ 准教授　前田　忠彦
16：25 ～ 16：35	休　憩	
16：35 ～ 16：50	総合討論および 顧問お一人ずつからの総括コメント	
16：50 ～ 17：00	謝辞および閉会	所　長　椿　広計

出席者名簿

NOE形成事業 顧問

伊藤　聡	計算科学振興財団 チーフコーディネータ
上田　修功	理化学研究所 革新的知能統合研究センター 副センター長
蒲地　政文	株式会社オーシャンアイズ 取締役
小宮山　靖	日本製薬工業協会 医薬品評価委員会 データサイエンス部会 副部会長 ※ 日本製薬工業協会会長代理
佐藤　嘉倫	京都先端科学大学 人文学部 学部長／教授
副島　豊	日本銀行金融研究所 所長
山形 与志樹	慶應義塾大学大学院 システムデザイン・マネジメント研究科 教授

NOE形成事業 運営委員会委員

椿　広計	所長
山下　智志	副所長／NOE 推進室長／リスク解析戦略研究センター 副センター長
加藤　昇吾	リスク解析戦略研究センター長／准教授
上野　玄太	統計科学技術センター長／モデリング研究系 データ同化グループ 教授 （次世代シミュレーションNOE 代表）
中野　慎也	モデリング研究系 データ同化グループ 教授 （次世代シミュレーションNOE 副代表）
前田　忠彦	データ科学研究系 調査科学グループ 准教授 （調査科学 NOE 代表）
福水　健次	統計的機械学習研究センター長／教授
松井　知子	統計的機械学習研究センター 副センター長／教授
吉田　亮	ものづくりデータ科学研究センター長／教授
藤澤　洋徳	ものづくりデータ科学研究センター 副センター長／教授
松井　茂之	医療健康データ科学研究センター長／教授
野間　久史	医療健康データ科学研究センター 副センター長／教授

※ その他、統計数理研究所教員、立川共通事務部長他関係者が傍聴

オンライン会議の様子

NOE形成事業 顧問の方々

統計数理研究所の出席者

● 開会・出席者紹介

【司会：山下副所長／NOE 推進室長】私は、本日の司会進行を務めます統計数理研究所副所長、兼、NOE 推進室長の山下です。この顧問会議は、各専門分野の先生方から、いろいろな意見、ご助言を伺う機会です。今後の NOE 形成事業の運営の参考とさせていただくべく、今回、この会議の後には、詳細な議事録を含めた報告集を作成させていただきます。

本日は、7 名の顧問の先生方にご出席いただくことがかないました。

- 計算科学振興財団　チーフコーディネーター　伊藤聡先生
- 理化学研究所革新的知能統合研究センター　副センター長　上田修功先生
- 株式会社オーシャンアイズ取締役　蒲地政文先生
- 日本製薬工業協会医薬品評価委員会データサイエンス部会副部会長　小宮山靖先生

なお、小宮山先生は、日本製薬工業協会の岡田安史会長の代理でご出席いただいているということになっています。

- 京都先端科学大学人文学部学部長、教授　佐藤嘉倫先生

佐藤先生は、本務のご都合上、16 時からのご参加の予定となっています。

- 日本銀行金融研究所　所長　副島豊先生
- 慶應義塾大学大学院システムデザイン・マネジメント研究科教授　山形与志樹先生

山形先生は、初めと終わりの 1 時間以外は、移動等のご都合により、お聞きになるだけのご参加ということで伺っています。

続きまして、統計数理研究所の NOE 形成事業運営委員会委員の紹介をします。
所長の椿広計です。

【椿所長】よろしくお願いします。

【司会】リスク解析戦略研究センター長、准教授、加藤昇吾です。

【加藤准教授】よろしくお願いします。

【司会】リスク解析戦略研究センターの副センター長は、私、山下智志が務めさせていただいています。よろしくお願いします。モデリング研究系データ同化グループ教授、兼、統計科学技術センター長、上野玄太です。

【上野教授】よろしくお願いします。

【司会】モデリング研究系データ同化グループ教授、統計科学技術センター副センター長、中野慎也です。

【中野教授】よろしくお願いします。

【司会】データ科学研究系調査科学グループ准教授、前田忠彦です。

【前田准教授】よろしくお願いします。

【司会】統計的機械学習研究センター長、教授、福水健次です。

【福水教授】よろしくお願いします。

【司会】統計的機械学習研究センター副センター長、教授、松井知子です。

【松井（知）教授】よろしくお願いします。

【司会】ものづくりデータ科学研究センター長、教授、吉田亮です。後ほど吉田先生は、発表予定なので、その時に、ごあいさつをさせていただこうと思います。ものづくりデータ科学研究センター副センター長、教授、藤澤洋徳です。

【藤澤教授】よろしくお願いします。

【司会】医療健康データ科学研究センター長、教授、松井茂之です。

【松井（茂）教授】松井茂之です。よろしくお願いします。

【司会】医療健康データ科学研究センター副センター長、教授、野間久史です。

【野間教授】よろしくお願いします。

【司会】今回は、顧問の先生方のご意見を直接拝聴できる大変貴重な機会です。オンラインでの顧問会議開催ということもありまして、出席者名簿に名前のある者以外に、研究所の教員、もしくは、立川共通事務部長他、事務部の関係者等にも傍聴させていただけるように設定しました。その点をご了承いただければと思います。

The Institute of Statistical Mathematics

第4回 統計数理研究所 NOE形成事業 顧問会議

所長挨拶・統計数理研究所の概要

統計数理研究所長
椿　広計

資料 1

第2代および5代
末綱恕一所長による揮毫

所長挨拶・統計数理研究所の概要

第４回　統計数理研究所
NOE(Network Of Excellence)形成事業顧問会議

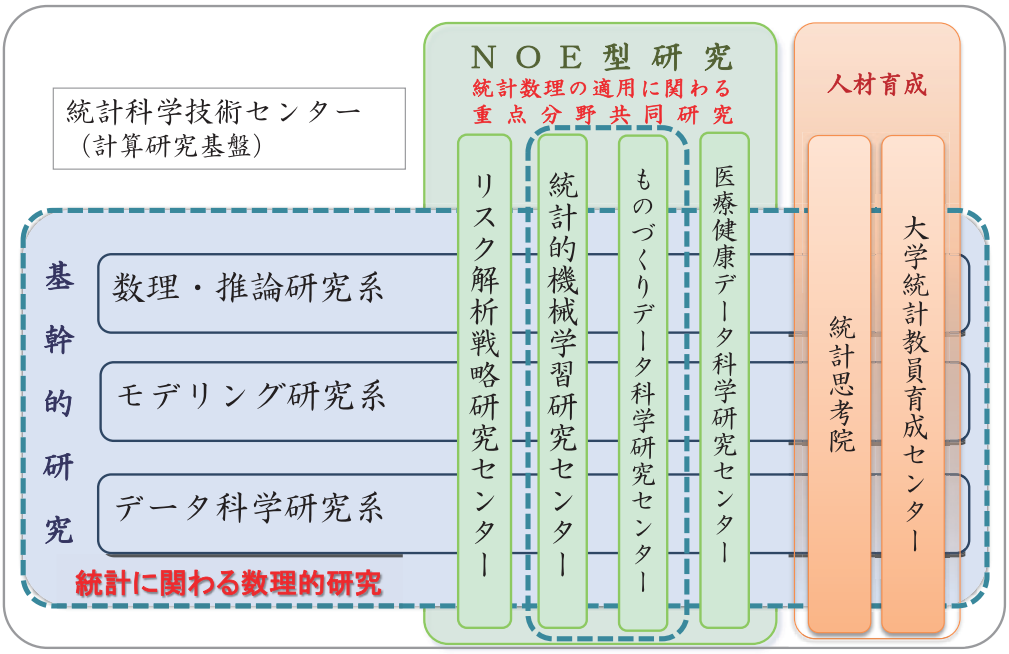
令和５（2023）年５月16日

所長　椿　広計

大学共同利用機関法人 情報・システム研究機構
統計数理研究所
As of, May 1, 2023

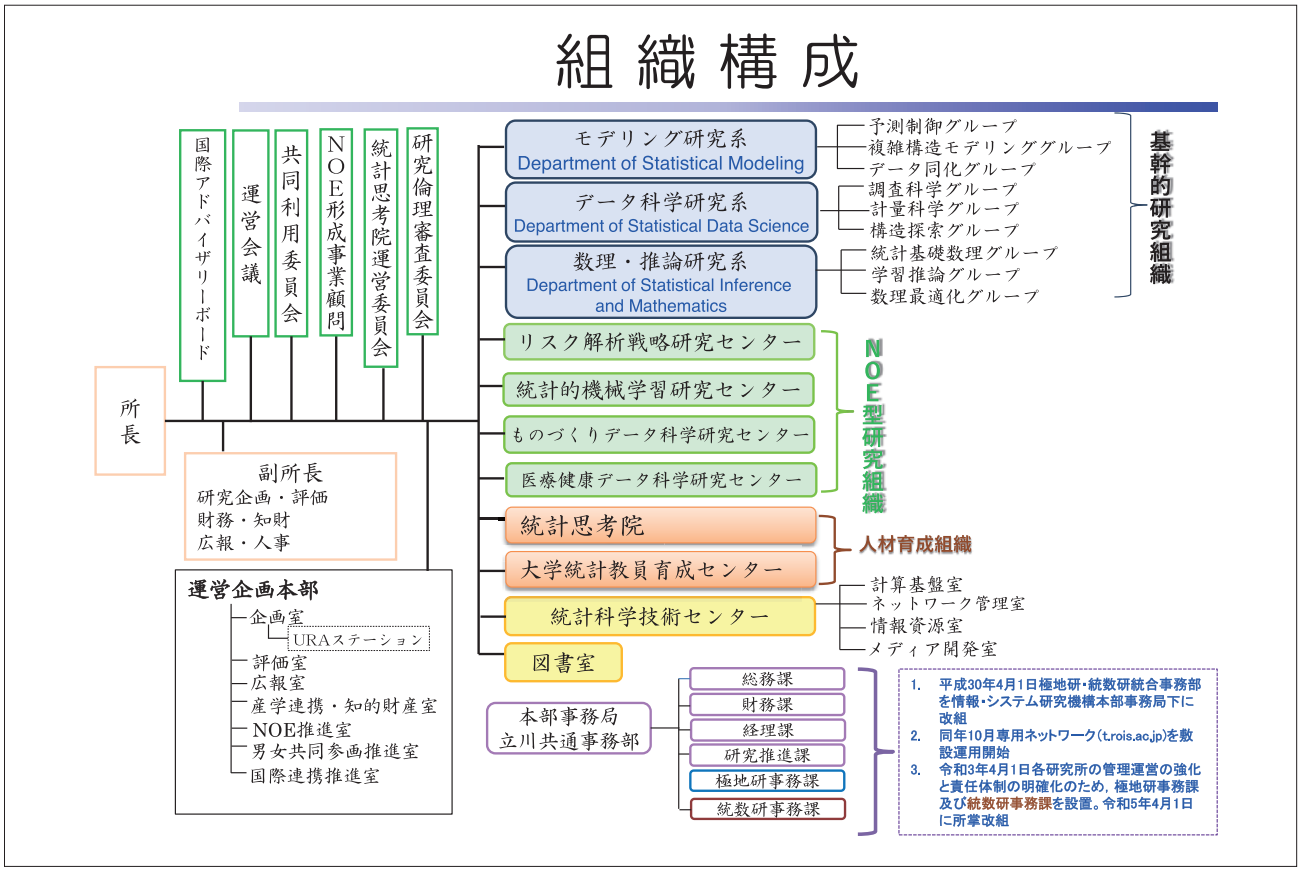
● スライド 1

主要事業

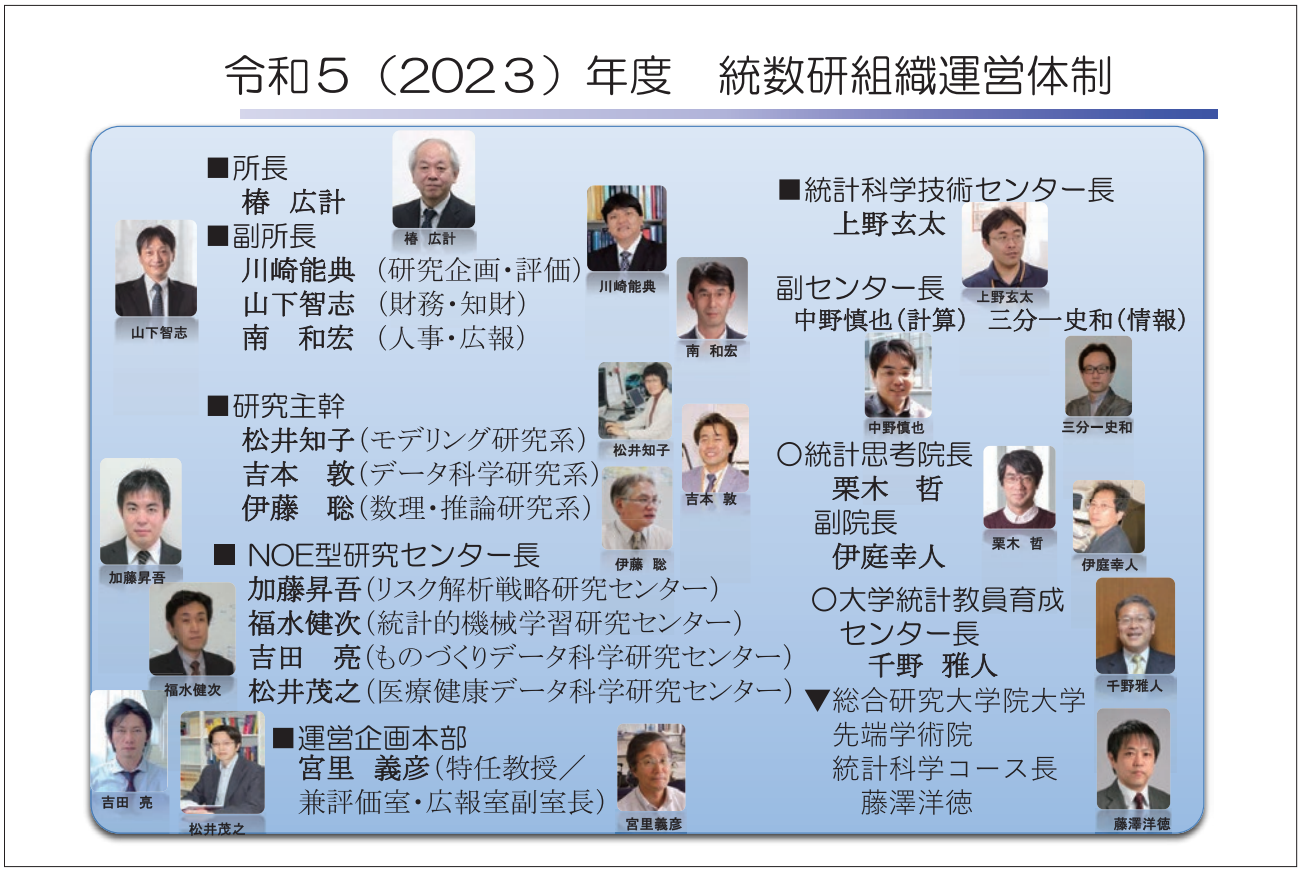


令和5年度内改組予定

● スライド 2


組織構成
国際アドバイザリーボード
運営会議
共同利用委員会
NOE形成事業顧問
統計思考院運営委員会
研究倫理審査委員会
所長
副所長
研究企画・評価
財務・知財
広報・人事
運営企画本部
企画室
URAステーション
評価室
広報室
産学連携・知的財産室
NOE推進室
男女共同参画推進室
国際連携推進室
モデリング研究系
Department of Statistical Modeling
データ科学研究系
Department of Statistical Data Science
数理・推論研究系
Department of Statistical Inference and Mathematics
予測制御グループ
複雑構造モデリンググループ
データ同化グループ
調査科学グループ
計量科学グループ
構造探索グループ
統計基礎数理グループ
学習推論グループ
数理最適化グループ
基幹的研究組織
リスク解析戦略研究センター
統計的機械学習研究センター
ものづくりデータ科学研究センター
医療健康データ科学研究センター
NOE型研究組織
統計思考院
大学統計教員育成センター
人材育成組織
統計科学技術センター
計算基盤室
ネットワーク管理室
情報資源室
メディア開発室
図書室
本部事務局
立川共通事務部
総務課
財務課
経理課
研究推進課
極地研事務課
統数研事務課
1. 平成30年4月1日極地研・統数研統合事務部を情報・システム研究機構本部事務局下に改組
2. 同年10月専用ネットワーク(t.rois.ac.jp)を敷設運用開始
3. 令和3年4月1日各研究所の管理運営の強化と責任体制の明確化のため，極地研事務課及び統数研事務課を設置。令和5年4月1日に所掌改組

● スライド 3

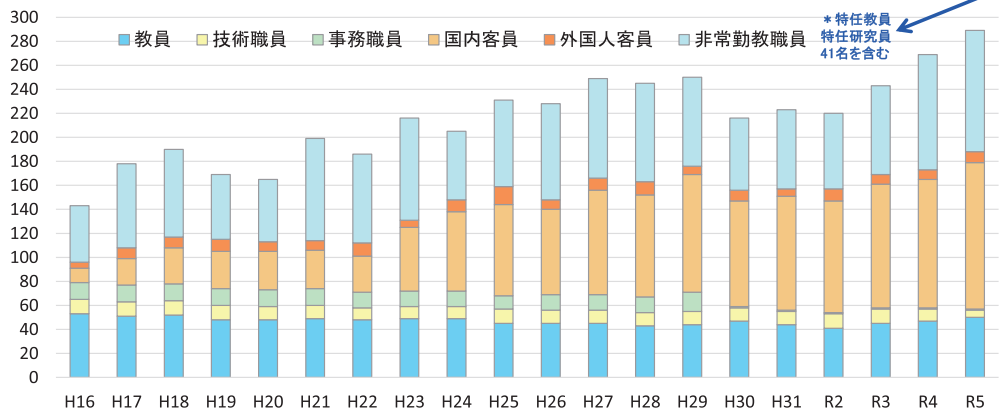

令和５（2023）年度　統数研組織運営体制
■所長
椿　広計
■副所長
川崎能典（研究企画・評価）
山下智志（財務・知財）
南　和宏（人事・広報）
■研究主幹
松井知子（モデリング研究系）
吉本　敦（データ科学研究系）
伊藤　聡（数理・推論研究系）
■ NOE型研究センター長
加藤昇吾（リスク解析戦略研究センター）
福水健次（統計的機械学習研究センター）
吉田　亮（ものづくりデータ科学研究センター）
松井茂之（医療健康データ科学研究センター）
■運営企画本部
宮里　義彦（特任教授／兼評価室・広報室副室長）
■統計科学技術センター長
上野玄太
副センター長
中野慎也（計算）　三分一史和（情報）
○統計思考院長
栗木　哲
副院長
伊庭幸人
○大学統計教員育成センター長
千野　雅人
▼総合研究大学院大学先端学術院統計科学コース長
藤澤洋徳

● スライド 4

職員数

※　各年４月１日付の数字（所長は含まない）
※　教授・准教授・助教はクロスアポイント，機構本部DS統括兼務者，出向者を含む
※　事務・技術職員は再雇用を含む
※　客員は各年度末実績数／令和３年度以降は期初見込み数
※　平成30年4月1日事務部が本部事務局立川共通刃部に改組。連携して業務を遂行。

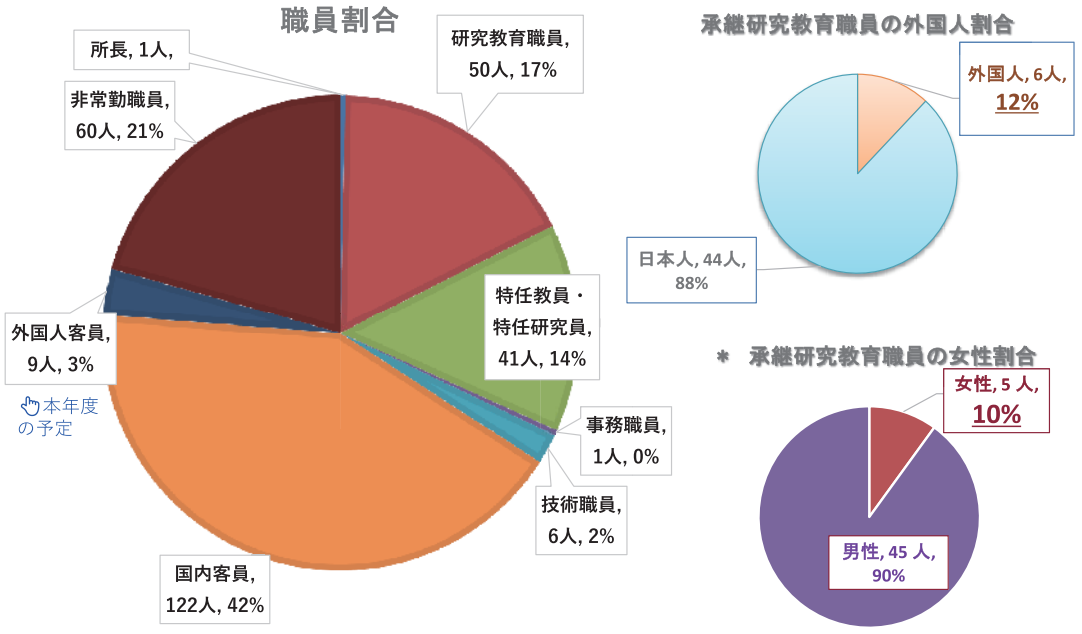
	H16	H17	H18	H19	H20	H21	H22	H23	H24	H25	H26	H27	H28	H29	H30	H31	R2	R3	R4	R5
教授	19	20	19	16	17	19	16	18	17	16	18	18	21	17	19	18	18	19	18	22
准教授	19	18	19	18	16	16	18	19	20	19	18	19	18	18	18	18	20	21	21	20
助教	14	12	13	16	15	15	14	11	12	10	9	8	6	9	10	8	3	5	8	8
技術職員	12	12	12	12	11	11	10	10	10	11	11	11	11	11	11	11	12	12	10	6
事務職員	14	14	14	14	14	14	13	13	13	12	13	13	1	1	1	1	1	1	1	1
国内客員	12	22	30	31	32	32	48	54	69	76	71	87	85	98	88	95	95	103	107	122
外国人客員	5	9	9	10	8	8	10	6	10	15	11	10	11	7	9	6	10	8	8	9
非常勤職員	47	70	73	54	52	85	74	85	57	72	80	83	82	74	60	66	63	74	96	101



● スライド５

統数研教員構成の多様性

統数研の教職員構成　【令和5年度期初】



＊客員を除く統数研所属職員は159人（所長を含め）

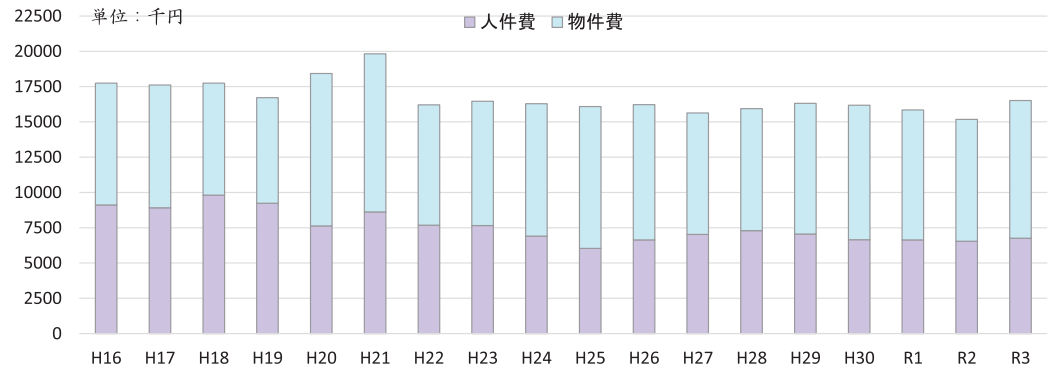
● スライド６

財

歳出決算額

＊運営費交付金 百万以下四捨五入　　単位：百万円

	H16	H17	H18	H19	H20	H21	H22	H23	H24	H25	H26	H27	H28	H29	H30	R1	R2	R3
人件費	911	891	981	897	763	832	769	765	691	604	663	702	729	705	666	664	654	675
物件費	863	870	793	737	1081	1121	853	881	938	1005	958	861	864	926	953	921	864	976
合計	1774	1761	1774	1634	1843	1983	1622	1647	1629	1609	1621	1563	1594	1631	1619	1585	1518	1651



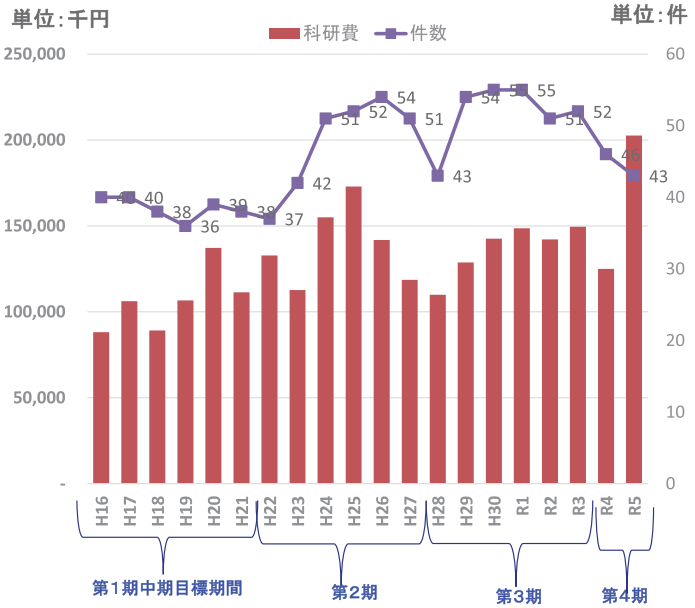
● スライド 7

外

【令和5年度　暫定速報値】科学研究費

研究種目	件数
新学術領域	1
学術変革領域 ＊A公募に新規採択	3
基盤研究（S）	–
基盤研究（A）	3
基盤研究（B）	8
基盤研究（C）	15
挑戦的研究	2
若手研究	9
研究活動スタート支援	2
特別研究員奨励費	–
研究成果公開促進費 ＜学術図書＞	–

新規採択＋継続件数 ➡ 計43件



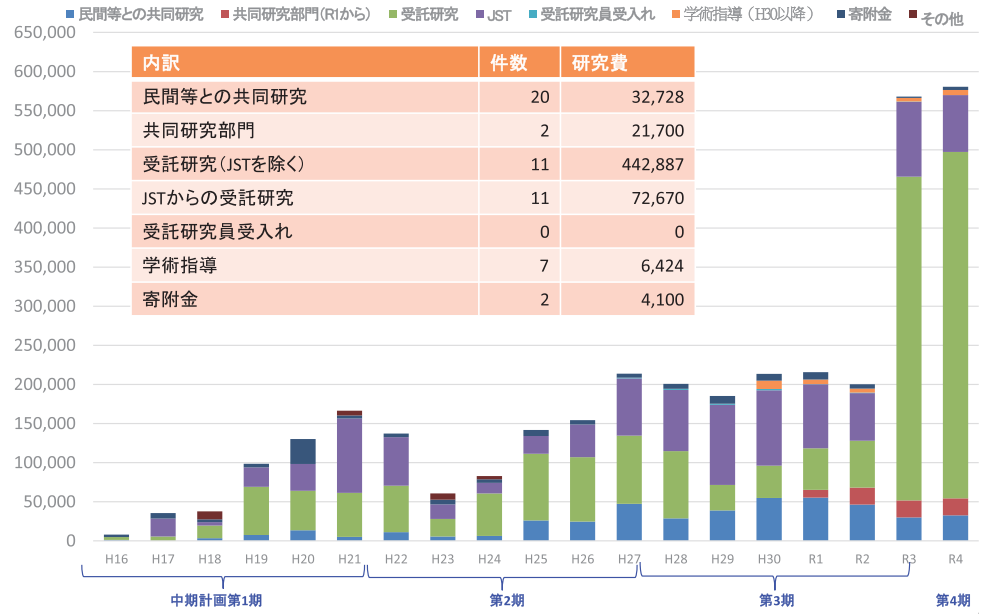
● スライド 8

外

外部資金獲得状況

●令和4年度

単位：千円

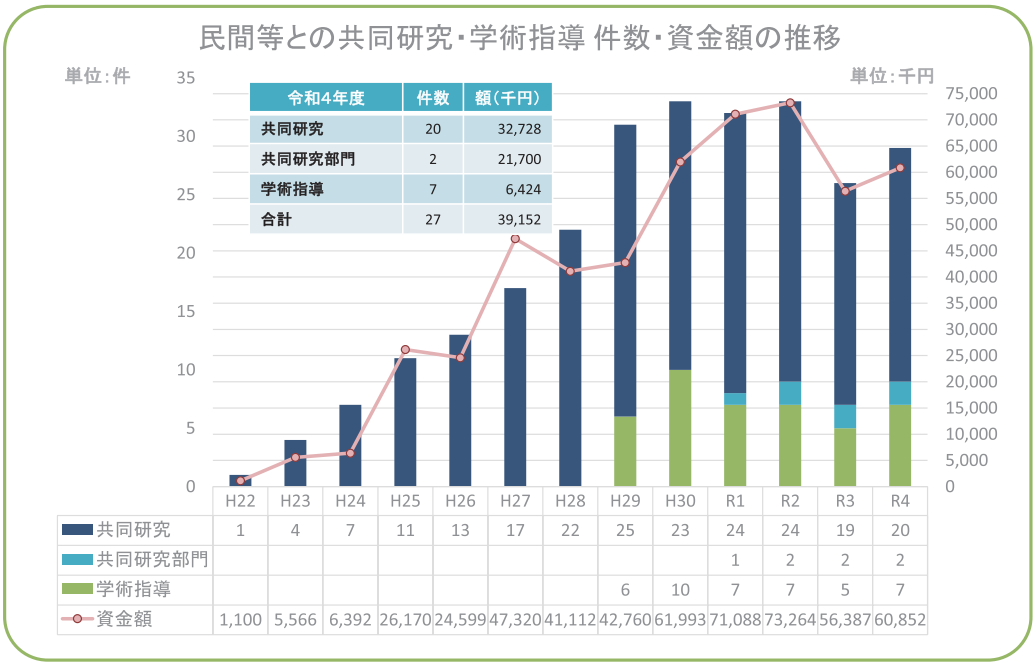


内訳	件数	研究費
民間等との共同研究	20	32,728
共同研究部門	2	21,700
受託研究（JSTを除く）	11	442,887
JSTからの受託研究	11	72,670
受託研究員受入れ	0	0
学術指導	7	6,424
寄附金	2	4,100

● スライド 9

外

民間等との共同研究・学術指導の実績



令和4年度	件数	額（千円）
共同研究	20	32,728
共同研究部門	2	21,700
学術指導	7	6,424
合計	27	39,152

	H22	H23	H24	H25	H26	H27	H28	H29	H30	R1	R2	R3	R4
共同研究	1	4	7	11	13	17	22	25	23	24	24	19	20
共同研究部門										1	2	2	2
学術指導								6	10	7	7	5	7
資金額	1,100	5,566	6,392	26,170	24,599	47,320	41,112	42,760	61,993	71,088	73,264	56,387	60,852

● スライド 10

令和４（2022）～令和９（2027）年度

第４期中期目標・中期計画

情報・システム研究機構の中目・中計の内，統数研に係る主なもの

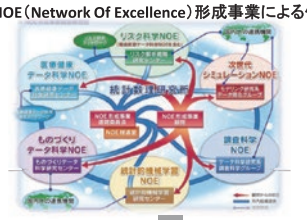
【研究】
- 先端的な基幹研究の推進

【共同利用・共同研究】
- 産官学連携／共同研究の強化
- NOE（Network Of Excellence）形成事業による異分野融合の推進・国内外の研究組織との連携機能の強化
- 様々な応用分野との共同利用・共同研究の支援体制を充実させるための統計科学スーパーコンピュータの提供
- 高分子物性データベース等，スパコンを活用した次世代シミュレーション・統計的機械学習・データ駆動型研究に資する多様なデータ基盤の整備異分野融合の進展

【教育・人材育成】
- 統計エキスパート
- データサイエンス教員
- 医療統計

共同利用・共同研究の加速
複数年度予算措置制度新設による国際共同研究の実施

分野融合の推進・国内外研究組織との連携強化
NOE（Network Of Excellence）形成事業による促進



人材育成
統計エキスパート人材，統計専門家人材，データサイエンス人材

● スライド 11

第4期中期計画 P3: 先端的な基幹研究の推進

大規模・複雑なデータに基づく統計数理科学の中核的研究機関として、国際的評価の高い研究の持続的創出を目指すため、運営企画機能及び研究支援機能の強化、基幹研究系の再編及び新領域創生部門の新設を行い、先端的な基幹研究を推進する。

評価指標	(1) **令和5年度に基幹研究系を改組し、先端データサイエンス研究系を設置する。新研究系には高等研究センターと革新統計数理研究センターを配置し、マテリアルズインフォマティクス共同研究拠点を含む、統計数理科学の新たな分野を先導・開拓する研究プロジェクトを累計3件立ち上げ、国際共同研究、産学共同研究を加速**する。
	(2) **大規模統計数理研究のコーディネートを念頭に、研究所の複数部署に散在する研究支援機能を抽出・一本化した新組織を設置することで効率化し、教員の研究専従率を36.5%に向上させ、先端的な基礎研究を推進**する。（第3期実績見込み：35.1%）
	(3) 先端的な研究成果に関連する厳選したプレスリリースを第４期末までに累計22件とする。（第３期実績見込み：19件）

● スライド 12

第4期中期計画 P23: スパコンを活用したデータ基盤の整備

環境負荷低減等の革新的特性を持つ新材料の探索のための高分子物性データベース等、スパコンを活用した次世代シミュレーション・統計的機械学習・データ駆動型研究に資する多様なデータ基盤の整備を行う。データベース構築のためのソフトウェア提供も行う。

評価指標	(1) 新規の取組として、高分子物性データベースへのポリマー（高分子化合物）登録数を第4期末までに12万件以上（注：世界最大）とする。

● スライド 13

（大学等番号： ｘｘ　大学等名：情報・システム研究機構　統計数理研究所）　令和６年度　教育研究組織改革分（組織整備）の概要

整理番号：　事業名　データ駆動型社会実現を目指す「先端データサイエンス研究系」新設による統計的機械学習理論と先端学術分野との共同研究・共同利用拠点機能の強化

令和６年度要求教員数

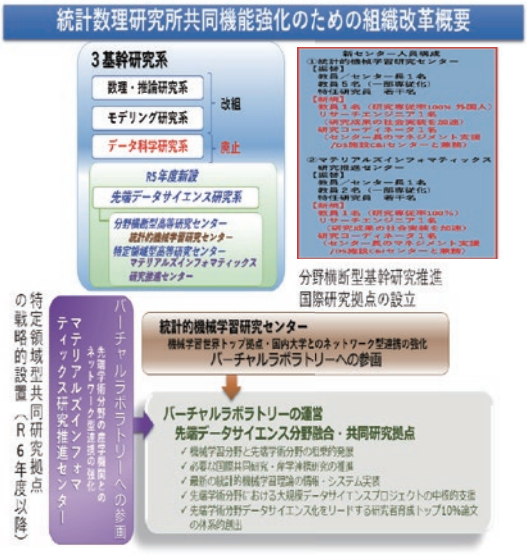
【要求教員数】　６人
（うち継続分０人、拡充分０人）

教育研究組織整備概要（新規）

教育研究組織の概要

統計数理研究所（以下、統数研）は、統計的機械学習の理論・方法・実践において世界をリードする成果を次々に生み出し、その結果、分野・組織・国境の垣根を越えた共同研究が急速に拡大している。本組織改革では、爆発的に増大する共同研究へのニーズに応えるべく、中長期的ビジョンに基づき機動的に戦略目標を設定し、統計的機械学習の重点分野における研究力ならびに共同研究機能の抜本的強化を図る。令和5年度に現行の基幹研究系を改廃し、所長直下に配置する基幹研究系「先端データサイエンス研究系（以下、新研究系）」を新設する。まずは新研究系の下に「統計的機械学習研究センター（以下、機械学習センター）」を移管し、国内外トップ研究者らとの連携の下で分野横断型基幹研究を推進する国際研究拠点を設立する。

さらに、統計的機械学習の重点分野の共同研究を推進する特定領域型共同研究拠点を順次設置していく。令和5年度には、マテリアルズインフォマティックス分野の共同研究拠点（以下、マテリアルズインフォマティックス研究推進センター）を設置する。令和6年度以降も関連分野の動向を注視しながら、研究所全体の戦略目標の下、特定領域型共同研究拠点を順次設置していく。限られた人的資源の下で拠点の活動を最大化するために、産学及び国内外から多様な専門性を有する人材が集う「バーチャルラボラトリー」体制を構築し、世界の巨大研究グループと伍する複合的・挑戦的・長期的研究を推進していく。上述の2研究センターには、令和5年度に研究専従を原則とする教員6名を含む9名を配置転換し、令和6年度に外国人研究者を含む優秀な若手人材等6名を新たに採用して組織の規模拡大を図る。令和6年度要求人員の内訳は、若手教員2名、研究成果のシステム実装を担うリサーチエンジニア2名、大型プロジェクトの管理・運営や研究成果の分野横断的技術移転等を担う研究コーディネータ2名である。新研究系では、センター長のイニシアティブの下、研究コーディネータが外部連携を支援・促進する体制を構築し、機動的かつ総合的に研究戦略を立案・遂行していく。



教育研究組織の教員の配置状況

教員の配置状況
※（ ）は外国人教員数、[ ]は若手教員数、<>は全学的な研究マネジメントを担う者の数

教員数	うち、組織整備における措置人数	うち、学内からの振替人数	うち、学外からの新規採用
15人(3人) [9人]<2人>	6人(1人) [4人]<2人>	9人(2人) [5人]<0人>	0人(0人) [0人]<0人>

措置教員等の状況（役割等）

【教育研究組織の教員体制についての考え方】
① 複合的・挑戦的・長期的ハイリスク・ハイリターン研究に研究資源を集中的に投下していくために、新研究系では研究専従率100%の教員の恒常的ポジションを確保する。
② 基礎研究の成果を分野横断的かつ即時的に技術移転していくために、ソフトウェアやデータベース開発、システム実装に関する高度専門性を有するリサーチエンジニア教員を配置する。
③ バーチャルラボにおける大型プロジェクトのマネジメント、技術移転可能な研究シーズの発掘、国内外及び産学連携ネットワークの形成を支援する研究コーディネータを配置する。
【新規要求教員等の必要性・役割】
① 若手教員（准教授・助教）2人（うち１人は外国人）：統計的機械学習を専門とする若手教員（研究専従率100%）を採用し、研究実務を担当する。
② リサーチエンジニア（若手）2人：ソフトウェア・データベース開発、計算高度化支援を推進する技術系教員を採用する。研究成果の社会実装の加速を図る。
③ 研究コーディネータ（教授相当）2人：バーチャルラボの大型プロジェクトのマネジメント、研究系ないしは情報・システム研究機構のアウトリーチ活動も担う。

● スライド 14

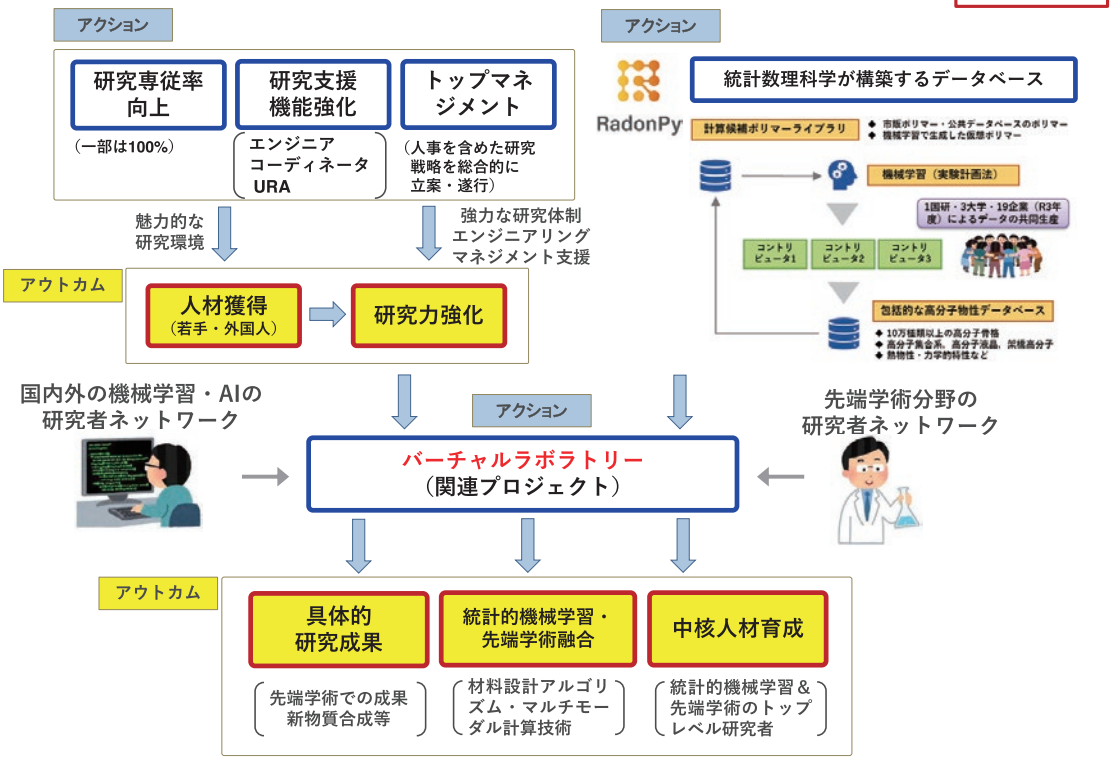
新研究系組織改革・拡充の工程表

	第4期中期計画期間					
	令和5年	令和6年	令和7年	令和8年	令和9年	令和10年
新研究系	2024/01 先端データサイエンス研究系 正式発足・教員配置転換	2024/10~ 特定領域型高等研究センター新設計画（理学系の学際分野を検討）	2025/06 新センター設置拡充要求（特定領域型高等研究センター 新規人員2名配置）	2026/10~ 特定領域型高等研究センター新設計画（情報学・数理科学分野を検討）	2027/06 新センター設置及び分野横断型高等研究センター拡充要求（新規人員4名配置）	2028/10~ 特定領域型高等研究センター新設計画
分野横断型高等研究センター	2024/01 統計的機械学習研究センター設置 教員公募開始	2024/07以降 採用者3名着任 国際研究ネットワーク活動	国際研究ネットワーク活動	国際研究ネットワーク活動	2028/01 教員公募開始 国際研究ネットワーク活動	2028/07以降 採用者2名着任 国際研究ネットワーク活動
特定領域型高等研究センター	2024/01 マテリアルズインフォマティックス研究推進センター設置 教員公募開始	2024/07以降 採用者3名着任 産学官研究ネットワーク活動	産学官研究ネットワーク活動	2026/01 新センター設置 教員公募開始 産学官研究ネットワーク活動	2027/07以降 採用者2名着任 産学官研究ネットワーク活動	2028/01 新センター設置 教員公募開始 産学官研究ネットワーク活動
バーチャルラボラトリー	2023/09~2024/01 設立準備 2024/01~03 共同研究者・滞在型研究者 公募・決定	2024/04~07 活動インフラ整備 2024/08 活動開始 2025/01~03 共同研究者・滞在型研究者 公募・決定	2025/09~ バーチャルラボ拡充準備作業開始 2025/04~06 共同研究者・滞在型研究者 公募・決定	2026/08 拡充されたラボの活動開始 2026/04~06 共同研究者・滞在型研究者 公募・決定	2027/09~ バーチャルラボ拡充準備作業開始 2027/04~06 共同研究者・滞在型研究者 公募・決定	2028/08 拡充されたラボの活動開始 2028/04~06 共同研究者・滞在型研究者 公募・決定

● スライド 15

組織整備の戦略と期待される効果

別添 図1



● スライド 16

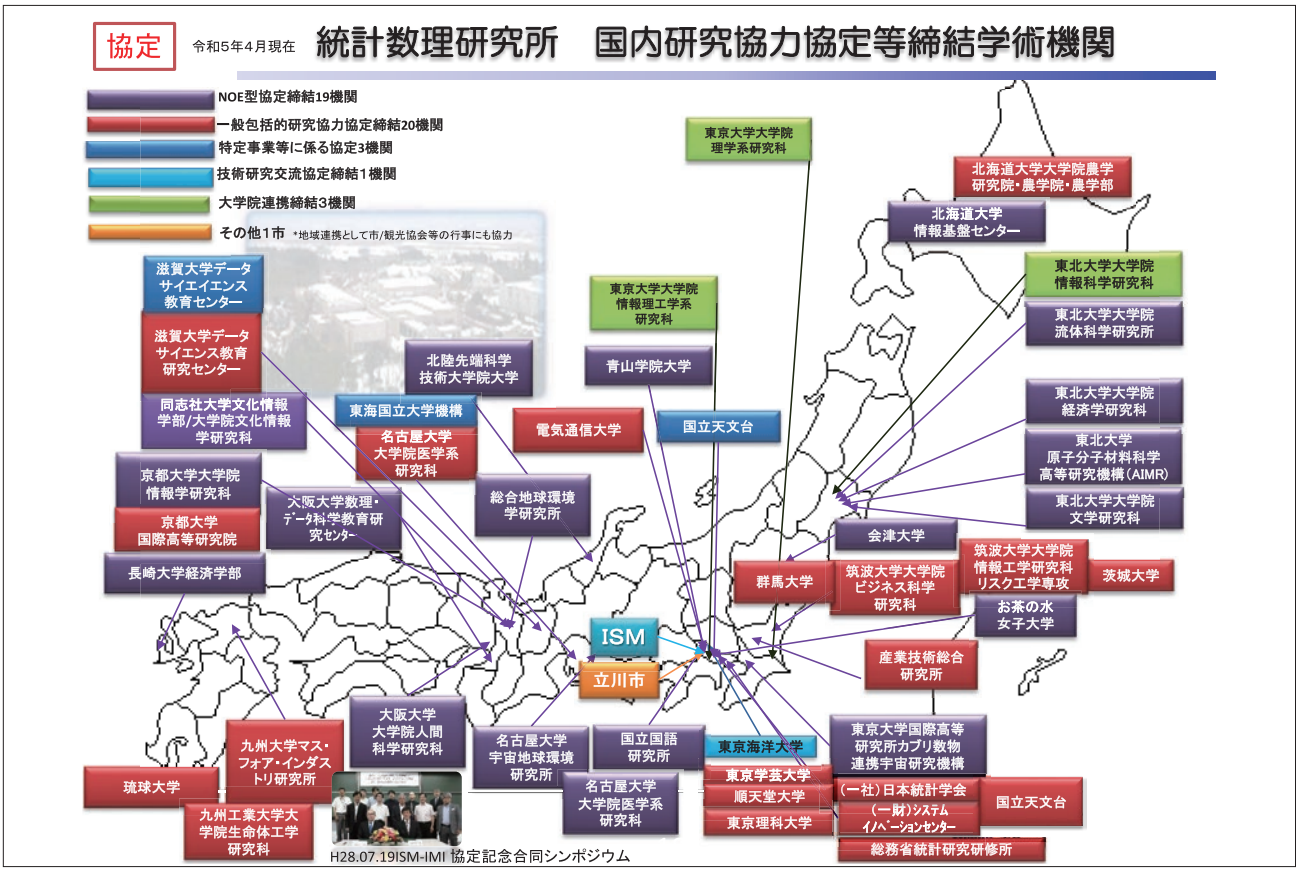
第4期中期計画 P10: 共同利用・共同研究

～産官学連携／共同研究の強化
～NOE形成事業による異分野融合の推進
～国内外研究組織との連携強化

研究者コミュニティからの要望や、SDGs達成、環境負荷低減等の社会の要請に応えるため、大規模データの利用技術の中核的役割を果たす統計数理科学が中心となる分野横断型のNOE（Network Of Excellence）形成事業による異分野融合を推進する。産業界、学界等、幅広く外部機関と連携して課題解決を行うNOE型研究組織の見直しにより産官学連携機能を強化し、新制度の導入により国内外の研究組織との連携機能を強化する。

評価指標	(1) 複数年の予算措置を行う国際共同研究制度を新たに設置し、第4期末までに2件の研究課題を実施する。
	(2) 先端的数理技術の進展と産官学連携の進展を促進すべく、**既存のNOE型研究組織を方法論型と分野・領域型に再編**することにより国内外の研究組織との連携機能を強化し、NOE型の連携協定を第4期末までに新規に7件締結し、累計24件とする。（第3期実績見込み：17件）

● スライド 17

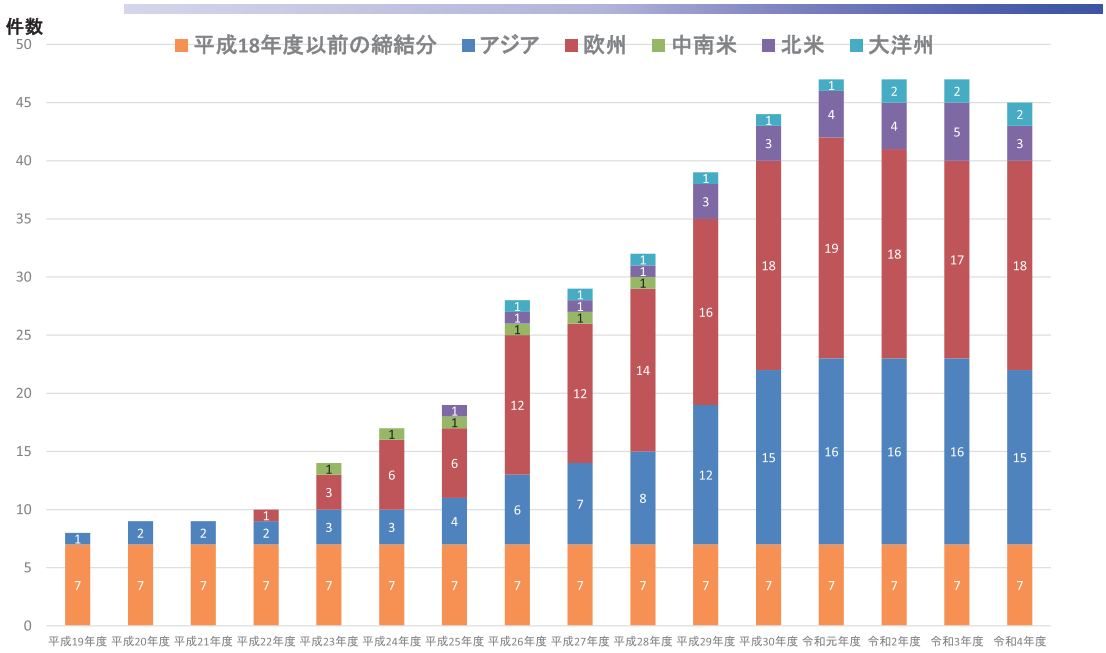


● スライド 18

協定

有効国際交流協定(MOU)締結数の推移

＊LOIも含む



＊平成19年度以降の有効MOUのみここに掲載

＊平成18年度以前のMOU締結実績は10件（**※内7件が現在も有効**）

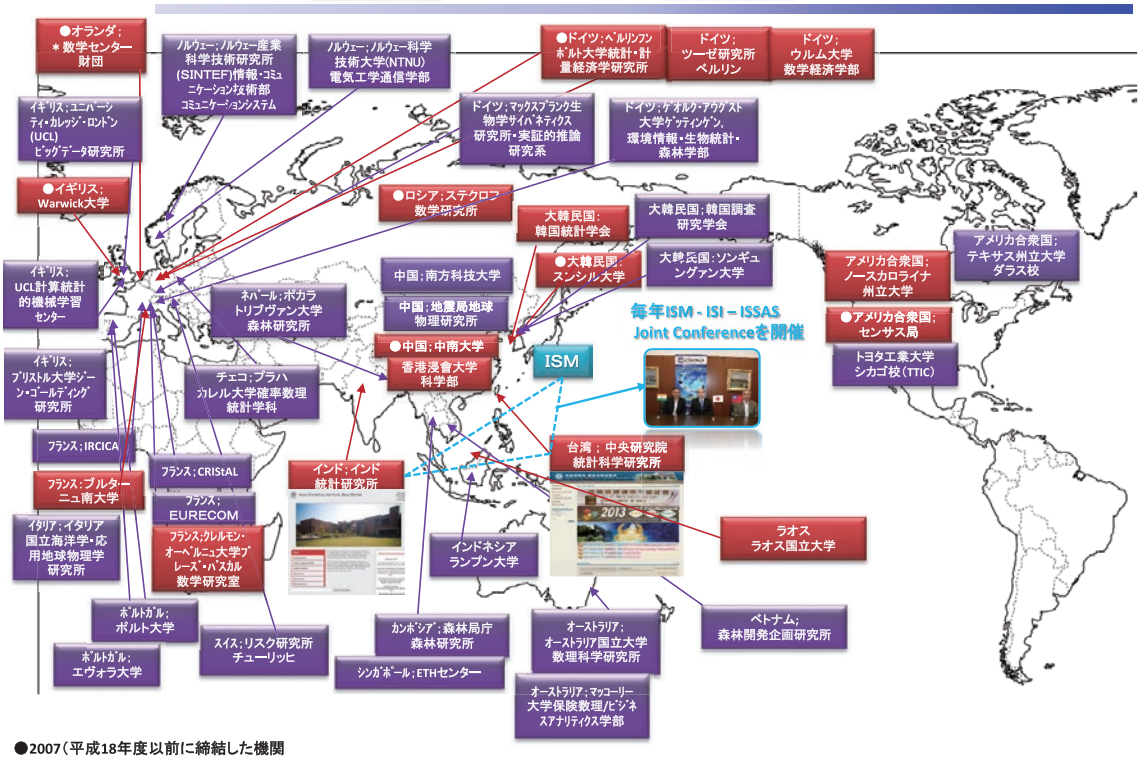
● スライド 19

協定 令和5年4月現在

統計数理研究所　海外学術交流協定締結機関

NOE型協定締結機関　28機関

一般協定　17機関(内*1機関はLetter of Intent)



●2007（平成18年度以前に締結した機関）

● スライド 20

第4期中期計画 P37: 人材育成（統計エキスパート）

ビッグデータやAI活用の社会的浸透に伴い、全国の大学で急拡大しているデータサイエンス教育の体制整備に資するべく、「数理及びデータサイエンスに係る教育強化」の拠点校、「数理・データサイエンス・AI教育の全国展開」の協力校、「データサイエンス系大学教育組織連絡会」とも連携し、**大学等における統計学の教育研究を担う若手中核人材**（統計エキスパート人材）の育成を行うとともに、大学のデータサイエンス研究力強化のために、**DS施設で第3期に立ち上げたデータサイエンス教員養成パイロット事業を、DS施設と統数研との連携の下で第4期では本格事業**とする。

評価指標	(1) 新規の取組として統計エキスパート人材育成プロジェクトを実施し、第4期末までに累計35名以上を育成する。
	(2) 統計エキスパート人材育成コンソーシアムへの参画機関数を、第4期末までに累計27以上とする。（第3期実績見込み：20機関）
	(3) 新規の取組としてデータサイエンス教員養成事業を実施し、第4期末までに累計8名以上を育成する。

● スライド 21

第4期中期計画 P38・P39: 人材育成（医療健康データサイエンス・産業界のデータサイエンス）

第4期中期計画 P38：人材育成(医療健康データサイエンス人材)

医療・健康科学分野における大学等とのネットワーク形成に基づき、統計科学の高度な実践力・研究力を備えた統計専門家人材（医療健康データサイエンス人材）の育成を行う。

評価指標	(1) 医療健康データサイエンス人材育成数を第4期末までに累計48名以上とする。（第3期末実績見込み：24名）

第4期中期計画 P39: 人材育成(産業界のDS人材)

社会人を含めた幅広い研究人材に対して、社会課題の解決等、現実社会での実践に向けた多様なスキルアップ機会を講座事業を通じて提供し、産業界のデータサイエンス高度人材育成を行う。

評価指標	(1) リーディングDAT関連講座受講者数を第4期末までに累計1,400名以上とする。（第3期末実績見込み：1,255名）

● スライド 22

● 所長挨拶・統計数理研究所の概要について

【司会】それでは、プログラムに移りたいと思います。まず、所長の椿広計から、本会議開催に当たってのごあいさつと、統計数理研究所の概要について、ご説明させていただきます。画面共有もしますが、お手元の資料で資料 1 をご覧いただければと思います。椿所長、よろしくお願いします。

【椿所長】本日は、ご多忙の中、私どもの NOE 形成事業顧問会議にご出席いただきましたことを、研究所を代表して、心からお礼申し上げます。本日のこの顧問会議は、統計数理研究所の今後の中期計画の実施、特に、組織改革、あるいは人材育成に関して、顧問の先生方に、いろいろな意見を頂戴したいと思います。個別の NOE 形成事業に関しても、ものづくり、あるいは統計的機械学習、それから今年は、国民性調査の実施年なのですけれども、そのようなことに関しても、いろいろとご意見を賜ればと思っています。

私のほうから、まず研究所の概要を、特に今日お諮りしたいことについて、少し話させていただければと思います。資料 1 は、PDF で取りあえず画面共有できているのではないかと思いますけれども、お手元の資料も見ていただければ、幸いです。

さて、統数研の主要事業というのは、私は所長で研究者ではありませんが、研究者は全員が、この基幹研究系に入っています。それが、横型で基礎研究を行います。それに対して、縦型で、統計数理の適用に関わる重点分野というものを設定して、ここに Network Of Excellence というものが配置され、センターがあります。それから、もう一つ、第 3 の事業として、人材育成というもので、統計思考院と、近年、大学統計教員育成センターを設置したのですけれども、ここにありますように、令和 5 年度、改組を予定しています。

なお、この主要事業を支えるための組織構成というものが、このような形になっています。われわれの運営に対して、外部の先生方にいろいろな意見を伺い、調整していただくために国際アドバイザリーボード、運営会議等々があるのです。NOE 形成事業顧問会議は、われわれの中では、このような全体のことについて、積極的にご意見を頂戴できる貴重な機会という形になっているところです。

この組織を支える構成として、現在、統数研のマネジメントは、このようなメンバーで行われています。今日の出席者も、この中からという形になっています。

統数研の概要ですけれども、職員数の推移を見ていただきますと、この水色のところが、承継職員といわれている、教授、准教授、助教で、現在 50 名ぐらいの組織です。ここは、いわゆる予算が毎年 1%削減といわれる運営費交付金の事情から、必ずしも伸びているという状況ではありません。むしろ非常勤の特任教員、特任研究員といわれている、事業型に 41 名ぐらいいて、伸びてきています。ここに NOE や、人材育成関係の事業があります。それから、特に国内客員が 122 名いらっしゃるわけですけれども、ここは、かなりの部分が、NOE 活動を支えてくださっているということがあります。

教職員構成に関しては、今、申し上げたところですが、このような形です。この中で、当

然、外国人の方、女性の方がどれくらいかということに関しては、われわれにとっても、マネジメントの非常に重要な関心事です。研究教職員に関しては、女性のほうで苦戦しているというのが、率直なところです。

これが、運営費交付金、国から来ている予算なのですけれども、これも、先ほど言いましたように、一定の削減率というものがあるところです。この部分を外部資金、あるいは国のプロジェクトの資金で補っていかなければならないというのが、われわれ国立の研究機関、あるいは国立大学など、みんなが抱えている課題と考えていただければと思います。

これが学術振興会から来ている科学研究費で、ここの直接経費に対して、間接経費 30%が、われわれの研究の真水の部分を支えていくということになります。件数的に、承継職員 50 人ぐらいの組織なのですけれども、かなりの部分を取ってきています。近年、少し件数が増えましたけれども、最近、大型科研費が取れるようになっているということで、ここは伸びています。

財務的に申しますと、外部資金の獲得状況で、近年この受託研究といわれている部分が、非常に大きく伸びています。現在国のプロジェクトを 3 件ほど受託していることがあって、ここが非常に大きく伸びた部分です。そのうちの 1 つは、リスクセンターが抱えている地震、それからもう一つは、ものづくりセンターが抱えている富岳の事業、それから、今日お話しします大学統計教員の育成というものです。その 3 つの事業が、非常に大きなシェアを占めていることが、実情です。それ以外にも、もちろん民間との共同研究を含めて、少しずつ増えてきているということが実態です。

民間との共同研究自体については、かなり昔に比べて圧倒的に増えているのですけれども、コロナ禍で若干伸び悩んでいるというのも実情だということを申し上げておきたいと思います。それから、学術指導は、ご説明が遅れましたけれども、コンサルテーションに当たるものを近年行っているということです。民間に対してのコンサルテーションです。

それで、今日の本題です。第 4 期中期目標・中期計画に関わることです。ここは、概略的に書いていますけれども、先端的な基幹研究の推進以外に、共同利用・共同研究というところで、NOE 形成事業の役割です。その応用分野との共同研究・共同支援体制を充実させるスパコンの提供です。これも、既に成功して、うまくいっています。それから、後にあります高分子物性データベースです。このようなところが、非常に大きいです。それから、人材育成系のものがあって、統計エキスパート、データサイエンスの教員、それから、医療統計です。このようなところが、中期目標・中期計画に書き込まれているところです。

この中期計画・中期目標をもう少し詳しく述べますと、非常に重要なこととして、基幹研究系を改組します。先端データサイエンス研究系を設置します。新研究系に、高等研究センターをつくるのですけれども、ここに、マテリアルズインフォマティクス相当の、現在、NOE 活動でやっている共同研究拠点を含む、そのようなものをつくるということを、既に公約しているところです。これを支えるコーディネート機能も強化します。研究を支援する機能も強化して、できるだけ教員の研究専従率を上げていくことを、赤字に書いたようなこ

とを構想しているところです。

先ほどのものづくりセンターに関しては、かなり具体的な KPI を配置しています。これは、後ほど詳しく述べる機会がありますけれども、高分子物性データベースというのも、現在、吉田センター長が、産学連携の中で、富岳プロジェクト等で実践しています。

登録数を、第 4 期までに、12 万件以上のデータベースをつくっていきます。これは、世界最大となりますけれども、このようなことを行うことを公約しているところです。

これに関わって、現在、われわれは、組織改革の概算要求を行っています。先端データサイエンス研究系、現在の基幹研究系を改組して、先端データサイエンス研究系をつくって、その中に、統計的機械学習研究センターと、マテリアルズインフォマティクスの研究拠点となるセンターをつくっていきます。ある意味で、機械学習を横軸に、マテリアルズインフォマティクスのような先端的な学術を縦軸にして、これを融合する、まさに現在の NOE の中の非常に核となるものを、まず第 1 段階として先端データサイエンス研究系の中に入れます。今まで研究系とセンターの兼務というものを所員には掛けていたわけですけれども、兼務というよりは、ここ自体が、専任体制で行えるような形のものをつくっていくことを構想しています。これによって、国際連携、あるいは産官学の連携というものが立ち上がってくると同時に、先ほどのような大規模データベースを構築するということが可能になると考えているところです。

これは計画になりますけれども、もしこの概算要求が認められても、あるいは認められなくても、われわれは、やるつもりでいるのです。先端データサイエンス研究系をつくり、この中に、分野横断型のセンターと、特定領域型のセンターを配置し、それを結合するような、いわゆる、バーチャルラボラトリというものを配備しながら進めていきます。これを徐々に拡充して、統数研の共同研究の機能、NOE の機能も含めて、強化していきたいと考えているところです。

これをつくること、特にバーチャルラボラトリをつくることによって、機械学習、人工知能と先端学術分野、まず第 1 弾は、マテリアルズインフォマティクスですけれども、そのようなものを融合させる場を積極的につくります。そこで、このようなものを支援するエンジニア、いろいろなシステムをつくる、データベースをつくって実装するエンジニアとマネジメントを行っていただくような URA 的な方々、コーディネータを配置することで、ここを加速的に推進しようという、そのような構想です。

もう一点申し上げますと、産官学連携強化や、この NOE 型のところ、先ほどの統計的機械学習と、マテリアルズインフォマティクスというのは、その 1 例ですけれども、これにつきまして、今後できるだけ既存の NOE 型研究組織を、方法論型と分野・領域型に再編して、先ほどあるように研究組織との連携機能を強化することを徐々に進めていきたい。これによって、連携協定等を強化、促進していくことも、われわれは中期計画の中でコミットメントした次第です。

既に統数研は、国内研究協力として、これだけの研究協定を結んでいます。この中で、紫

で書いてあるものが、いわゆるNOE型の連携といわれているものです。一般協定とは違って、ネットワークをつくっていくということを主たる目的として行っているものです。これは国内です。

それに、国外のものというものも、以前に比べると、統数研は非常に海外との連携協定の強化を進めているところです。国外に関しても、この連携協定というものは、先ほどありましたNOE型というものを中心にして、進めていくことで、ヨーロッパやアジアの連携が進んでいます。最近、少しアメリカ、北米に対する連携というものが進んできたという状況です。

もう一点、人材育成の件も、今日、少しご意見を頂戴したいと思っている案件です。一つは、統計エキスパートです。これも、文部科学省の事業として受託したものですけれども、大学等における統計学の研究を担う若手中堅人材等を育成することにも、コミットメントしているところです。それと同時に、これに付帯する事業というか、こちらのほうが先にあったのですけれども、データサイエンス共同利用基盤施設という、私ども統数研に隣接する施設で立ち上げたデータサイエンス教員養成パイロット事業も、本格事業化するということです。この種のKPIを、いわゆる大学統計教員を35名以上、実際につくっていくということ、それから、参画していただく機関というものを強化することを考えています。

また、医療健康データ科学研究センターの中で、やはり大学のネットワーク形成が、第3期以降、進んでいるということで、ここにおいても人材育成を行っていきます。

それから、産業界のためのリーディングDATといわれているものもあって、これは、統計思考院というところで行っているのですけれども、このようなことも強化していくという構想を持っています。

一応、今日、中期計画に関連して、ご意見を頂戴したいことを中心に話させていただきました。この後、各論を各担当が話すことになります。私のほうの総論的なことは、以上として、ご質問等を頂ければ、幸いです。

【司会】椿所長、ご説明ありがとうございました。顧問の先生方のご質問、ご意見がありましたら、発言いただきたいのですが、NOE関係、また、統数研の組織全体は、普通の企業などに比べると、大変複雑な部分があります。ここは、よく分からなかったという点がありましたら、ぜひここでご質問いただければ、ありがたいです。伊藤先生、よろしくお願いします。

【伊藤顧問】伊藤です。所長、どうもありがとうございました。1つ教えていただきたいのは、所長のほうから、今回、URAや、あるいは研究を支援する部門を強化するというお話がありました。これは、何か新たな組織をつくるというご提案でしょうか。

【椿所長】現在、私どもは、支援組織は設けていまして、URAも自走化したので、これまでROIS、情報・システム研究機構に付いていたURAを、今年度から2名、あるいは、実情、2.5から3名なのですけれども、私ども、統数研で直接雇用する形にはなりました。こ

れをどのように改組していくかということはあるのですけれども、先ほど申し上げましたように、現在、統計的機械学習、あるいはマテリアルズインフォマティクスというのは、産学連携研究が加速度的に増大している分野で、それのいろいろなインターフェースの形成に、研究者に負担が掛かっているということです。今回の新研究系は、所長直下、つまり、運営企画本部で、少し事務やURAが配置しているところに設置して、できるだけ両センター長のマネジメントの負担を減らします。もちろん全研究系に対して、それなりの支援はするのですけれども、重点的にこの部分の研究支援機能を強化します。そのためには、実は、コーディネータ2名をさらに付けていただきたい、URAを付けていただきたいという概算要求、組織改革要求を現在行おうとしているところです。

【伊藤顧問】分かりました。以前に伺った時に、今もお話がありましたように、上位組織のほうにURAがいるというお話でしたので、やはりそれでは少し足りないかなと思って、ご質問させていただいたのですが、大変安心しました。どうもありがとうございました。

【椿所長】どうもありがとうございます。

【司会】どうもありがとうございます。他、いかがでしょうか。

【山形顧問】慶應の山形ですが、よろしいでしょうか。

【司会】山形先生、よろしくお願いします。

【山形顧問】椿所長、どうもありがとうございました。全貌がよく分かりました。少し最後の時に質問する時間があるかどうか分かりませんので、初めに質問させていただきます。NOEのいろいろなプロジェクトなど、いろいろな予算でサポートされているものと思いますが、例えば、ものづくり、あるいは医療という現場の問題解決型のセンターと、機械学習という手法に特化したセンターがあります。統計数理研究所の特徴というのは、多分、最先端の統計数理手法を研究するということかと思いますが、同時に、社会の問題を解決するというのが、今の日本全体の科学技術政策の方向かと思います。ですから、やはり大学や、このような国の研究機関も含めて、問題解決型で技を磨くといいますか、実際に刀を研いで、切ってみるということが必要かと思います。

そのようなことで、一般的に統計的機械学習を磨いていけば、何でも切れる刀ができるということなのか、あるいは問題解決型で、やはり踏み込んでいくことで初めて、役に立つツールができると考えるのか、その辺りはどのようなお考えでしょうか。

【椿所長】先ほどNOEの中には、横断的な、ある意味で、メソドロジーに関する、横断的なメソッドに関する研究センター、つまり、統計的機械学習は、どちらかというと、理論研究としては、横断的に利活用できます。それから、そのようなものの利活用の成果を受ける場としてのマテリアルズインフォマティクス、あるいは医療健康というものがあるわけです。横断的な部分で出てきている統計的機械学習研究センターなどは、どちらかというと科学研究費の競争的資金や何かの、そのようなところから来ている部分が大変ありますし、マテリアルズインフォマティクスは、逆に産学連携的な資金が流れています。そのような、それぞれの資金の特徴があります。

いずれにしても、山形先生のおっしゃったように、われわれは、第1弾としては、この統計的機械学習という、現在、人工知能の、数理的な基盤、基礎となっている部分、ここを横軸にして、こちらに先端学術、マテリアルズインフォマティクス、あるいは医療、生物、環境も含めて、今後、この場と理論とを形成するような、そのような融合のバーチャルラボラトリを少しずつ作りたい。バーチャルラボラトリは、統数研に限らず、いろいろな分野の先生方が常に集う場という、そのような場の形成を、第4期以降、順次、行っていきたい。バーチャルラボラトリを、いろいろな縦軸、横軸の構造体の中でつくっていきたいと、今、構想しているところです。

バーチャルラボラトリというのは、あまり今まで文部科学省もやった経験がないということで、いろいろなご意見を頂戴しているところです。私どものような、割と小型の研究所が、それなりに社会に対する貢献をするためには、やはりつないでいくという機能を強化する、NOEの機能をやはり強化していくことしか、なかなか考えにくいのではないかというのが、今、私どもの構想の基盤的なアイデアとなっているところです。

【山形顧問】ありがとうございました。基礎研究と社会実装の間の死の谷を埋めるということで、非常に重要な機能ですね。

【椿所長】先生がおっしゃるとおりです。そのような意味で、バーチャルラボが機能するように、システムデザインしなければならないと検討しているところです。

【山形顧問】無事に渡れるように、私たちも連携したいと思います。よろしくお願いします。

【椿所長】ぜひよろしくお願いします。

【蒲地顧問】蒲地ですけれども、よろしいでしょうか。

【椿所長】よろしくお願いします。

【蒲地顧問】ありがとうございました。全体的な流れは、ある程度つかめたと思うのですけれども、今回、人材育成ということを非常に強調されていたかと思います。人材育成は、24ページあたり、その後にもありますし、また、今日のご発表の中でも、後ほどあると思います。この人材育成ということに関しましては、最終的には、現場なり、研究機関なり、そのようなところで実際に統計数理を使って働く方たちの専門家を育成するということもあります。それから、次にそのようなエキスパートを育てる側の方たちも、また育てなければならないという、二重にあって、なかなか大変で、時間がかかるものではないかと思います。これは、今期の中期計画だけではなくて、まだ今後も続いていくものだと考えてよいでしょうか。

【椿所長】私どもの構想としては、現在、国のプロジェクトとして5年間行うことは開始しているのですけれども、これを、やはり5年で終わらせないで、第2期、さらに5年続ける。あるいは、データサイエンス共同利用基盤施設のほうは、運営費交付金化していますので、永続的に続けることができます。ですから、このような仕組みを第4期の末までには確立して、サステナビリティというか、永続的な活動にできればとは考えているところです。

それから先生がおっしゃったように、これはあくまで大学で教える教員と共に、それらの大学や大学院に教育システム、教材や教育方法も提供するという形で、そこで、さらに大学院レベルの統計に、あるいはデータサイエンスに明るい修了生というものは、われわれは孫弟子という言い方をしてしまうことが、よくあるのですけれども、そのようなものを育成するということが最終的な狙いになってきます。日本においては、中高の先生方を育成するという教育学はあるのですけれども、なかなか大学、大学院レベルの教員を育成するのは、実は、初めての試みといわれています。後ほど説明がありますけれども、研修部で非常に頑張っていただいています。ぜひこれをうまく進めて、立ち遅れている日本の統計データサイエンスを少しでも進展させたいと考えているところです。

【蒲地顧問】そのような意味では、人材育成というか、組織だって、組織自体も考えていくということにもなるわけですか。

【椿所長】はい、そのとおりです。そのために、大学統計教員育成センターというものを昨年 1 月に設置しました。これは、後ほどセンター長から少し説明させていただきます。

【蒲地顧問】分かりました。あともう一点です。これは、11 ページ、科研費のデータがあったと思うのですけれども、今年度は、件数は減って、大型研究が通ったから、予算としては増えているので、それは、それでよかったと思います。けれども、若手研究が 9 件ぐらいだということで、基盤 C が 15 件などあります。例えば、私が以前いた JAMSTEC などでは、老獪といいますか、長年科研費を取ってきたり、あるいはそのような審査員になったり、そのような経験のある人間が選ばれて、若手が応募する時にもっと具体的なアドバイスを応募の書類を見ながら議論して、それを何かもっとよくして、採択件数を増やしていくという戦略的な取り組みを行っています。統数研のほうでも、そのような戦略的な試みというものは何かされているのでしょうか。

【椿所長】統数研でも、必ず若手の方々が出すものに関しては、副所長級が、あるいは主幹がレビューをするということになって、指導しています。この件数が減っているのは、確かに非常に心配はしています。ただ、先ほど申し上げましたように、承継職員数 50 名の研究所で、43 件ということになっている。それでも、われわれは、件数的には少し心配している部分が多々あるところです。ですから、先生がおっしゃったように、実は統計科学という科研費が細目になって、通りにくくなったということがあります。統計科学の審査員は、統計学ではなくて、ほとんど情報の先生方が入っているという実態があって、なかなか苦戦しています。それによって苦戦したということは、分析的には、そのように思っているのですぜひこれから少し、件数という意味で立て直しを図らねばと思っているところです。

【蒲地顧問】分かりました。ありがとうございました。

【上田顧問】すみません。理研の上田ですが、よろしいでしょうか。

【椿所長】よろしくお願いします。

【上田顧問】ご説明、どうもありがとうございました。マテリアルズインフォマティクスというのは、確かに日本が強い領域で、一方、統数研は、機械学習、数理統計が強い機関でも

あるので、そこで、このようなバーチャルラボを展開されるというのは、非常に良いことだと思います。一方、例えば、NIMS（国立研究開発法人物質・材料研究機構）なども、データベースは結構、整備されていると思うのですけれども、プログラムや技術の面では、個人的な印象ですけれども、不十分だと思います。その中で、この第 4 期中期計画の中で、このバーチャルラボというのは、日本のリーダーシップを取るような規模感でつくっていくのか、あくまでも統数研の中でとどまるのかというのは、どのような見解なのでしょうか。

【椿所長】これについては、後ほど吉田センター長から詳しく話があると思いますけれども、基本的には、世界最大のデータベースがほぼできる形になってきて、そのために、もう統数研にとどまらず、非常に大きなコンソーシアムができています。それから、統数研の所内に、既に三菱ケミカルを含めて、各民間のサブ研究所というものが出来上がってきているという状況が進んでいます。

私どもとしては、ものづくりセンターができてから、加速度的にそのことが進捗して、しかも富岳の CPU のかなりの部分を使えるということも認めていただいているということです。いわゆる NIMS は、実際の実マテリアルのデータベースをつくるわけですけれども、統数研は、スーパーコンピュータを使った計算機科学によって、計算機理論で第 1 原理に基づくデータベースとして、世界最大、恐らく今までと 1 桁規模が違うものをつくりつつあるという、そのような状況になっています。

【上田顧問】分かりました。非常に素晴らしい計画だと思います。ありがとうございます。

【椿所長】ありがとうございます。山下先生、各論も含めて話して、本当は、私より吉田センター長に答えていただくほうがいいかと思われるのですけれども。

【司会】皆さんから貴重なご意見をありがとうございました。各論に入る前に、もう核心を突いたご意見を頂いたような気がしますけれども。

【椿所長】核心を突かれた、そうですね。本当にありがとうございます。

【司会】もう少し掘り下げて説明させていただければと思います。

International Strategic Advisor

第4回 統計数理研究所 NOE形成事業 顧問会議

国際戦略アドバイザーからの助言への対応

統計数理研究所 副所長
川崎 能典

資料２

国際戦略アドバイザーからの助言への対応について

川崎能典 / 副所長(研究企画・評価担当)

令和5年5月16日

第4回NOE形成事業顧問会議

大学共同利用機関法人 情報・システム研究機構
統計数理研究所

● スライド 1

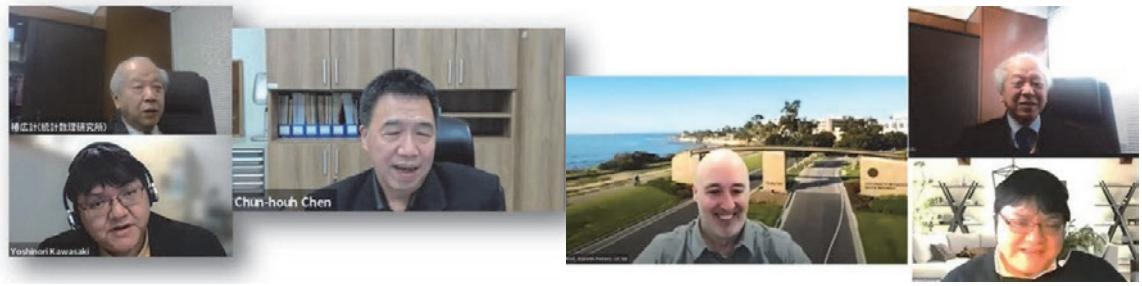
国際戦略アドバイザー: 制度概要

- 国際アドバイザリーボード(ISMの制度)
  - 5名程度、海外の学識経験者に委嘱
  - 任期(2年)中の来訪、助言レポートの執筆
- 国際戦略アドバイザー(ROISの制度)
  - ISMのアドバイザリーボードから1名をROISの国際戦略アドバイザーに推薦
  - 本部で審査、旅費・謝金等ROISが負担
  - 来訪、ヒアリングに基づくActivity Reportの執筆

● スライド 2

国際戦略アドバイザーとのミーティング

- Chun-houh Chen所長 台湾中央研究院統計科学研究所 [2022年1月20日, オンライン]
- Gareth W. Peters教授 カリフォルニア大学サンタバーバラ校 [2023年2月10日, オンライン]
  - 第4期中期計画について国際的見地から助言



● スライド 3

ROIS 第4期中期目標・中期計画

- 計画P28 評価指標(2)
  - 国際戦略アドバイザーからの助言に基づき、戦略企画本部において研究所等のベンチマーク機関の候補となる海外中核研究拠点を新たに第4期末までに5機関以上選定し、研究所等に提示する。
- 計画P43 評価指標(3)
  - 各研究所で組織化されている、外国人有識者を含むアドバイザリーボードからの報告・助言を踏まえた、研究所運営に関する取組等の検証を年1回以上実施し、その対応・反映状況を毎回公表する。
- https://www.mext.go.jp/a_menu/koutou/houjin/detail/1386151_00009.htm

● スライド 4

項目	指摘事項
教育及び研究の質の改善について	先端データ科学研究系の設置に向けて、既存のデータ科学研究系はどうするのか? (†)
NOE形成事業について	連携先の国際的組織、国際共同利用・共同研究に係るスキームに対する予算支援を、複数年の予算支援とともに提示すべき。
NOE形成事業について	既存のNOE型研究センターを方法論型と分野型に再編するのはよいことである。しかし対象となる方法論及び分野はあるのか?
スパコンについて	ISMはすでに世界で最先端のスパコン機能を有する研究機関であり、国内外の産官学連携を行ってきている。関連するスパコン設備の高い利用率を維持することが必須であり、様々な側面で利用率の継続的なモニタリングは行うべきである。(†)

(†)のついた指摘事項は「対応すべき案件」、無印は「検討すべき案件」と認識

● スライド 5

項目	指摘事項
共同利用・共同研究の促進について	戦略的に優先順位の高い共同研究事業に関して、少なくともその事業に係る活動／成果の要旨等、何らかの評価が必要ではないか? (†)
共同利用・共同研究の促進について	新たな国際共同研究・共同利用事業の選定に関する基準はあるか? (†)
国内外大学・研究機関との連携について	ISMは国内外組織とMoUを最も多く締結している学術機関の一つであり、潜在的連携先とMoUを締結し続けるべきである。但し、研究者交流、国際研究の提案数、事業数、研究論文発表数といった様々な側面から、過去と現在のMoUを評価すべきである。
研究者交流について	若手研究者の国際交流は、時空間分布で図示するべきである。過去の構造を示し、今後の配置を計画する上で、MoU／国／研究者ネットワークを評価したノードリンク図またはヒートマップが適していると思われる。

(†)のついた指摘事項は「対応すべき案件」、無印は「検討すべき案件」と認識

● スライド 6

項目	指摘事項
女性研究者の比率の向上について	全研究系において女性研究者の比率が低い。STEMサイエンスでは既知の問題であるが、ISMにおいても女性研究者を増やすことを検討すべきである。
研究者・研究業績の可視化について	ISMで行われている素晴らしい研究へのアクセスとインパクトを高めるために、ISMの主要な研究者一人一人の研究がより可視化されること、すべての教員がGoogle Scholarのプロフィール（そうすればH-index, I-index, 被引用率が容易に維持できる）、Research Gate, SSRNのプロフィールを積極的に活用することを含めて提案する。それにより、より大きなインパクトが示され、どの研究でどの研究分野にインパクトを与えているか、どの論文にそれが引用されているかなどがより認識されるようになるだろう。(†)

(†)のついた指摘事項は「対応すべき案件」、無印は「検討すべき案件」と認識

● スライド 7

結語

- 中期計画P28, P43を見据えた作業は、第4期初年度を終えて、対応/検討事項をピックアップした段階。
- 対応すべき事項(†)は、中期計画の実施に向けて進む中で自ずと対応状況が明らかに。
- ISMアドバイザリーボード、ROIS国際戦略アドバイザーからの新たな助言を踏まえ、指摘事項は増補、対応の状況は毎年更新。
- 大規模な国際外部評価は令和8年度(中期5年目)に実施予定。

● スライド 8

● 国際戦略アドバイザーからの助言への対応について

【司会】それでは、本研究所が 2021 年度、22 年度に招聘した国際戦略アドバイザーという方々がいらっしゃいます。その方たちのご助言というものを、いったんまとめました。いわゆる世界の先生方から頂いた意見について、その対応の方針について、研究企画担当の副所長、川崎のほうからご説明させていただきます。資料 2 をご覧いただければと思います。では、川崎先生、お願いします。

【川崎副所長】ご紹介いただきました、研究企画ならびに評価を担当しています副所長の川崎です。今日は、少し本筋という感じではないかもしれませんが、国際戦略アドバイザーからの助言への対応ということで、簡単にご報告させていただきたいと思います。

まず制度の概要をご説明するところから始めたいと思います。統数研は、研究所の制度として、まず国際アドバイザリーボードというものを設けていまして、5 名程度、海外の学識経験者に委員を委嘱するという形で、任期は 2 年ということにしています。その間にご来訪されれば、ベストですけれども、そうではなくても、最近は、いろいろと便利になってきていますので、オンラインミーティング等を通じて助言レポートを執筆していただいて、それを生かしながら、われわれも、その研究、それから、業務の体制をいろいろと変えていく、助言を生かしていくということでやっています。

一方、機構本部には、国際戦略アドバイザーというのがありまして、これは、構造的には、統数研のアドバイザリーボードを 5 名程度お願いした中から、1 名を機構本部の国際戦略アドバイザーというものに推薦申し上げて、これは、機構本部が決めることなわけですけれども、本部で審査の上、認められれば、旅費、謝金等を本部が負担してくれるというものです。これは、来訪していただいて、あるいはヒアリングに基づいて、Activity Report というものを、その任期内にご執筆いただくという、そのような制度になっています。ですので、統数研の制度と、機構本部の制度と 2 本あるのですけれども、本日は、この国際戦略アドバイザーとして、2021 年度、それから、22 年度、過年度 2 年度分のところで、国際戦略アドバイザーの方々から頂いたご助言に関して、簡単に取りまとめたものをご紹介申し上げたいと思います。

まず、2021 年度の末に近いところですが、台湾アカデミアシニカの統計科学研究所所長でいらっしゃいます Chun-houh Chen 先生に、オンラインでしたけれども、いろいろとミーティングを持たせていただきまして、その結果として、非常に長大なレポートを頂いています。それから昨年度は、カリフォルニア大学サンタバーバラ校の Gareth Peters 先生が、夏に来訪されて、短いミーティングを持ったのですけれども、それだけでは、少し足りませんので、結局、年が明けまして、2 月にオンラインで再度ミーティングを持たせていただきました。結果、大変、参考になるレポートを頂いています。この国際戦略アドバイザーの先生方には、第 4 期中期計画について、国際的な見地から助言を頂くというところが 1 つの主眼ですけれども、特に、機構本部の第 4 期中期目標・中期計画におきましては、関連する

事項が大きくは 2 つあります。

1 つは、計画番号 28 番の評価指標 2 というものです。これは、国際戦略アドバイザーからの助言に基づいて、戦略企画本部において、研究所等のベンチマーク機関、要するに、研究分野、それから、サイズ感等、似通った国際機関を選定して、そことの比較において、パフォーマンス等をチェックするということは、求められつつあります。ということですけれども、そのようなベンチマーク機関の候補となる海外の中核拠点等をいろいろとリコメンドしていただいて、それを機構本部で第 4 期末までに取りまとめるという、このような機構全体の目標が 1 つあります。

それから、計画の 43 ですけれども、こちらは、このような外国人有識者を含むアドバイザリーボードからの報告・助言を踏まえて、研究所運営に関する取り組み等の検証、これは、年 1 回以上実施して、その対応・反映状況を毎回公表するという、このような目標が上がっています。これは、全研究所が機構傘下で一斉にというよりは、機構本部で、どこかの研究所が毎年このようなことをやるということでした。統数研は、幸いにして、過去 2 年度にわたって、このような立派な先生方からご意見を頂く機会がありましたので、それを取りまとめて、昨年度末に公表という形にした次第です。

本日は時間も限られていますし、現状は、指摘事項の中から、このようなことに、われわれとして対応していく中で体制を強化していこうという事項を取り上げたという、第 4 期の初年度が終わった段階で、まだそのようなところですので、詳細の読み上げはしません。しかし、まずは、先ほど所長から説明のあった基幹研究系の再編に関する質問や、あるいは国際共同研究に関する予算支援に関するご助言・ご質問、それから、NOE 型研究センターの再編等に対するご質問、それから、スパコンの利用、特に利用率に関する、モニタリング等に関するご助言、このようなものも頂いています。

それから、共同利用・共同研究の促進についてということでは、その優先度の高い共同研究というものは、予算配置の大きなものという意味ですけれども、そのようなものに関しては、何かしらそのような評価というものを考えていく必要があるのではないかということです。それから、あとは、国際共同利用・共同研究ということを推進していくに当たっての課題の選定基準如何ということや、あとは、MOU というのは、これまで同様、締結を推進していくべきであるけれども、一方では、パフォーマンスチェックも怠らぬようにという、そのような助言も頂いています。それから、このページの 4 つ目は、やや特殊なというか、エキセントリックな視点ですけれども、これは若手に限らないと思いますが、研究者の国際交流というものを可視化するということは、非常に suggestive であるから、やってみてはどうかと。実際、これは Chun-houh Chen 先生からのアドバイスだったのですけれども、確かに提示してくださった図をレポートに含めていただいたりもしています。

それから、次のページのこちらは、所長からもご説明がありましたけれども、女性研究者の採用ということでは、近年、統数研は苦戦しています。これは、海外の大学でも、STEM 分野においては同じなのだけれどもという前置きがありながら、しかしながら、統数研にお

いても女性研究者を増やすことを検討すべきであります。これに関しては、機構長も変わりまして、新機構長の下で、機構全体として何か大きなアクションの方向性のようなものが出る可能性がありますので、そのようなものを、歩調を合わせながら進めていくということになろうかと思います。

最後は、これは丹念に統数研の教員の業績状況等をアドバイザーの方々が調べてくださった結果の助言と思いますけれども、なかなか統数研の教員一人一人は、業績の可視化という意味では、Google Scholar のプロフィールや、Research Gate、SSRN のプロフィール、このようなものがきちんとそろっている教員ばかりでないというご指摘と思います。そのような形で業績の可視化ということを、教員一人一人が心掛けるべきであるという、そのようなご助言を頂いているところです。実は、他にもたくさん助言いただいているのですが、われわれとして積極的に対応していきたいというものを、ある種セレクティブに選んだものがこれらです。

まだ第 4 期中目・中計期間の初年度が終わったところですので、このような助言で、われわれとして積極的に対応していく、あるいは対応すべきであると考えているものは、逐一といっても、毎年というレベルになると思いますけれども、リバイズしながら、われわれの体制強化につなげていきたいというところです。時間も限られていますので、私からは、以上とさせていただきたいと思います。どうもありがとうございました。

【司会】ありがとうございます。少し時間が押していますので、後に一括して質問を受け付けたいと思います。進行を変更させてください。

Network Of Excellence

第4回 統計数理研究所 NOE形成事業 顧問会議

NOE形成事業の概要

統計数理研究所 副所長
NOE推進室長
山下　智志

NOE（Network Of Excellence）形成事業の概要について

副所長／NOE推進室長
山下　智志

大学共同利用機関法人 情報・システム研究機構
統計数理研究所

● スライド 1

NOE

NOE形成事業とNOE型研究センター

【NOE (Network Of Excellence) 形成事業】
方法論の立場から、**各研究領域の発展**を図り、**新しい科学的方法論（第4の科学）の確立**、**異分野融合**、**新研究領域の創成**、そして**新しい共同研究スタイルの確立**を目指す

【NOE型研究センター】
- ✓ 喫緊の具体的な社会的課題の解決に向けて統計数理と個別科学分野の接点にあたる部分を活動の場とする
- ✓ NOE（Network Of Excellence）の中核として「ハブ」の役割を果たす

リスク解析戦略研究センター（Risk Analysis）
統計的機械学習研究センター（Statistical Machine Learning）
ものづくりデータ科学研究センター
（Data Science for Creative Design and Manufacturing）
医療健康データ科学研究センター
（Medical and Hearth Data Science）

● スライド 2

NOE

NOE形成事業とNOE型研究センター２

NOE事業とNOE型研究センターの違い： 1対1ではないことに注意

例）

リスク科学NOE

リスク解析戦略研究センター（所内）
リスク研究ネットワーク（所外）

調査科学NOE

データ科学研究系調査科学グループ（所内）
社会データ構造化センター（DS施設）

【NOE】
・リスク科学NOE
・統計的機械学習NOE
・次世代シミュレーションNOE
・医療健康データ科学NOE
・調査科学NOE
・ものづくりデータ科学NOE

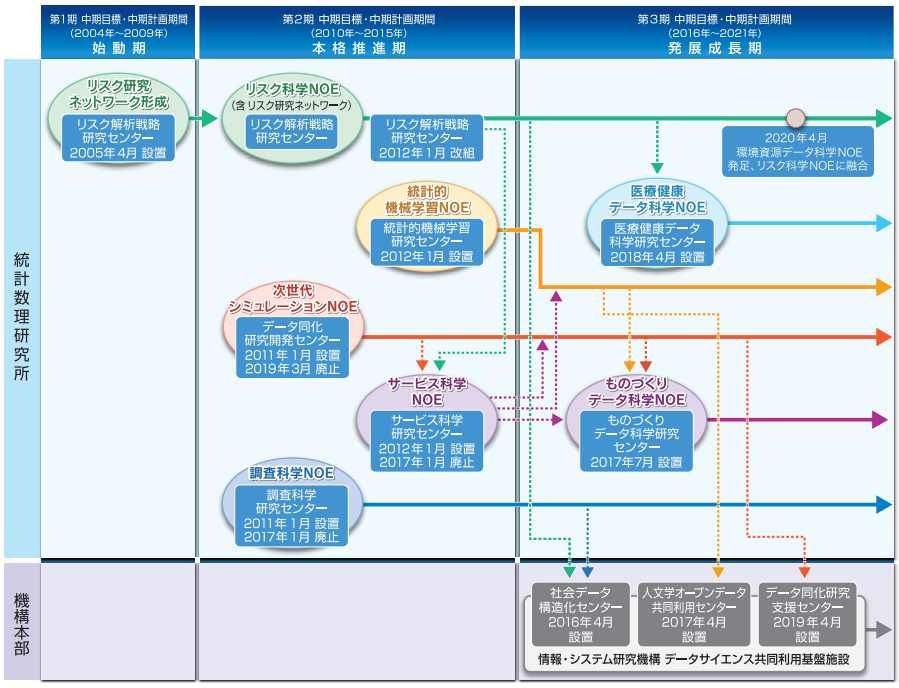
【研究所内のNOE型研究センター】

リスク解析戦略研究センター
統計的機械学習研究センター
ものづくりデータ科学研究センター
医療健康データ科学研究センター

● スライド３

NOE

ＮＯＥ形成事業の遷移



● スライド４

NOE

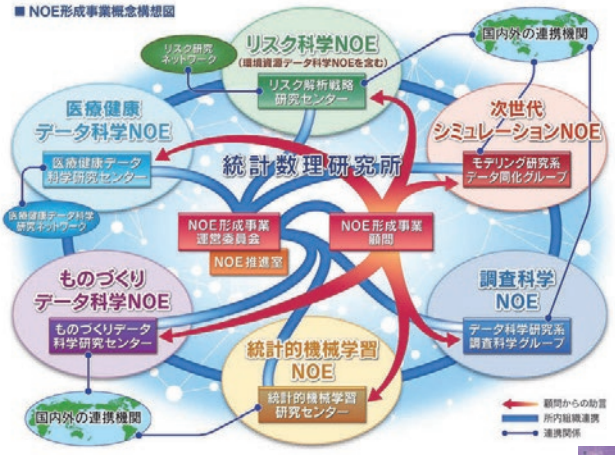
統計数理研究所ＮＯＥ形成事業

コミュニティから支持：大学共同利用機関として大学が求める要望を反映した**ネットワーク拠点・ハブ**の役割を果たす

- 現代社会の複雑な問題の解決には大規模データの活用と融合研究の実現が不可欠。
- 大規模データの利用技術を確立しつつある統計数理科学が中心となる分野横断型の**NOE**（**N**etwork **O**f **E**xcellence）構築が必要

個別NOEを総合的に推進してシナジー効果を発揮／NOE構築を通して新しい共同研究システムの確立を目指す

■ NOE形成事業概念構想図



https://www.ism.ac.jp/noe/project/intro.html

NOE形成事業顧問（2021.6.1−2023.5.31）

（継続）計算科学振興財団　チーフコーディネータ	伊藤　聡
【新規】理化学研究所　革新的知能総合研究センター副センター長	上田　修功
（役職指定継続）日本製薬工業協会　会長	岡田　安史
（継続）株式会社オーシャンアイズ　取締役	蒲地　政文
【新規】京都先端科学大学　人文学部　学部長・教授	佐藤　嘉倫
（役職指定継続）日本銀行金融研究所　所長	副島　豊
【新規】慶應義塾大学大学院システムデザイン・マネジメント研究科　教授	山形　与志樹

令和元(2019)年12月17日第3回顧問会議開催
各NOE推薦顧問6名が全員出席

平成24(2012)9月24日
第1回顧問会議開催

平成27(2015)6月19日
第2回顧問会議開催

※開催年度は運営指針の参考として
「NOE形成事業顧問会議報告集」を毎回発行
次回第4回顧問会議は2023年5月16日に開催予定

● スライド５

NOE事業の予算概要

第４中期計画
当初の概算

医療健康人材育成事業
概算要求⇒ミッション実現分
約6000万

富岳プロジェクト
約5000万

企業との共同研究費
約6000万

Ster-E
約3000万

所長裁量経費
1億3000万

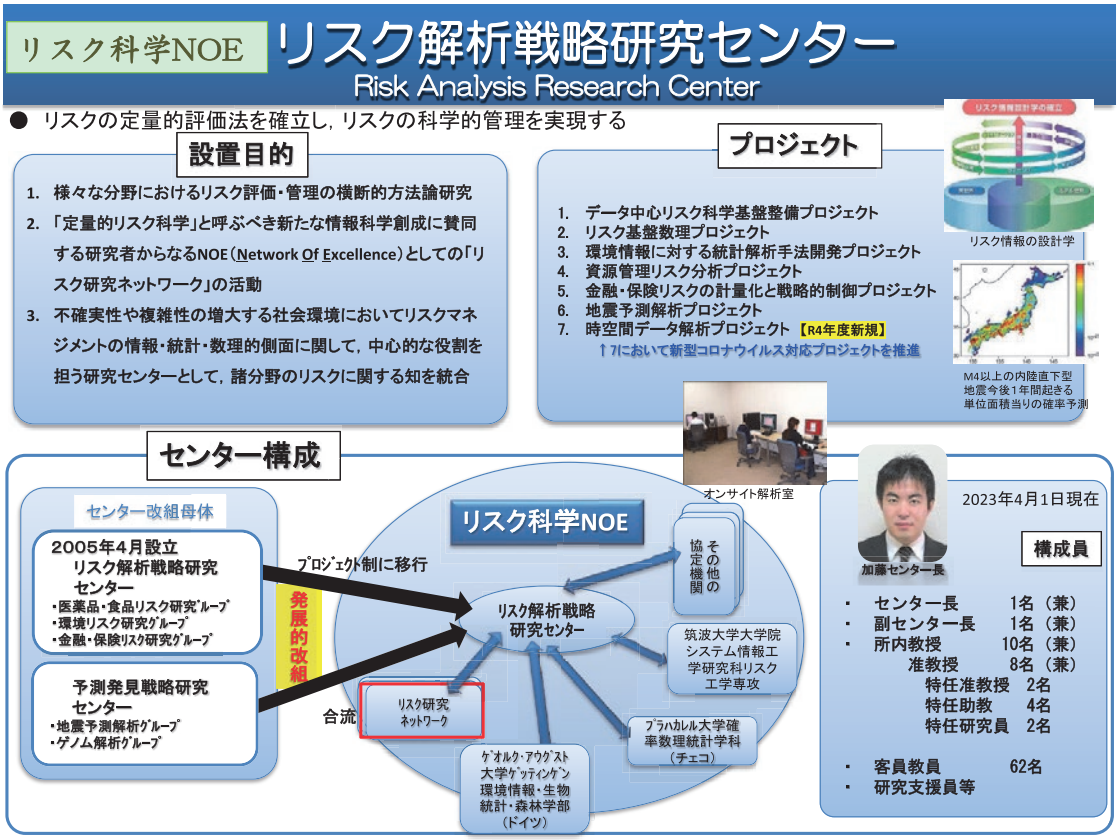
DS施設ミッション実現分
9000万

教員個人科研費・JST

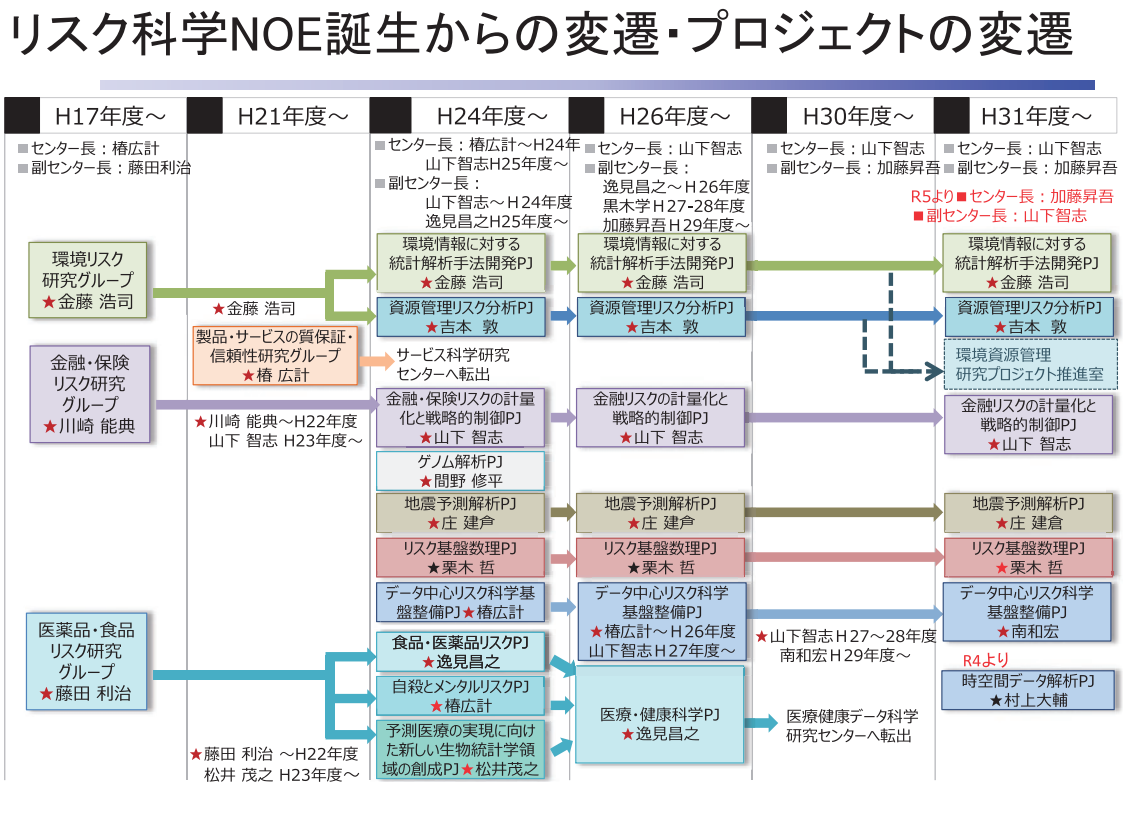
科研費・国プロ
間接経費
自己収入
運営交付金

- 医療健康データ科学NOE
- ものづくりデータ科学NOE
- リスク科学NOE
- 統計的機械学習NOE
- 次世代シミュレーションNOE
- 調査科学NOE

● スライド６


リスク科学NOE
リスク解析戦略研究センター
Risk Analysis Research Center
● リスクの定量的評価法を確立し，リスクの科学的管理を実現する
設置目的
1. 様々な分野におけるリスク評価・管理の横断的方法論研究
2. 「定量的リスク科学」と呼ぶべき新たな情報科学創成に賛同する研究者からなるNOE（Network Of Excellence）としての「リスク研究ネットワーク」の活動
3. 不確実性や複雑性の増大する社会環境においてリスクマネジメントの情報・統計・数理的側面に関して，中心的な役割を担う研究センターとして，諸分野のリスクに関する知を統合
プロジェクト
1. データ中心リスク科学基盤整備プロジェクト
2. リスク基盤数理プロジェクト
3. 環境情報に対する統計解析手法開発プロジェクト
4. 資源管理リスク分析プロジェクト
5. 金融・保険リスクの計量化と戦略的制御プロジェクト
6. 地震予測解析プロジェクト
7. 時空間データ解析プロジェクト【R4年度新規】
↑7において新型コロナウイルス対応プロジェクトを推進
リスク情報の設計学
M4以上の内陸直下型地震今後1年間起きる単位面積当りの確率予測
センター構成
センター改組母体
2005年4月設立
リスク解析戦略研究センター
・医薬品・食品リスク研究ｸﾞﾙｰﾌﾟ
・環境リスク研究ｸﾞﾙｰﾌﾟ
・金融・保険リスク研究ｸﾞﾙｰﾌﾟ
予測発見戦略研究センター
・地震予測解析ｸﾞﾙｰﾌﾟ
・ゲノム解析ｸﾞﾙｰﾌﾟ
プロジェクト制に移行
発展的改組
合流
リスク科学NOE
リスク解析戦略研究センター
リスク研究ネットワーク
協定その他の機関
筑波大学大学院システム情報工学研究科リスク工学専攻
ﾌﾟﾗﾊｶﾙﾚ大学確率数理統計学科（チェコ）
ｹﾞｵﾙｸ・ｱｳｸﾞｽﾄ大学ｹﾞｯﾃｨﾝｹﾞﾝ環境情報・生物統計・森林学部（ドイツ）
オンサイト解析室
加藤センター長
2023年4月1日現在
構成員
・ センター長 1名（兼）
・ 副センター長 1名（兼）
・ 所内教授 10名（兼）
准教授 8名（兼）
特任准教授 2名
特任助教 4名
特任研究員 2名
・ 客員教員 62名
・ 研究支援員等

● スライド 7


リスク科学NOE誕生からの変遷・プロジェクトの変遷
H17年度～
H21年度～
H24年度～
H26年度～
H30年度～
H31年度～
■センター長：椿広計
■副センター長：藤田利治
■センター長：椿広計～H24年
山下智志H25年度～
■副センター長：
山下智志～H24年度
逸見昌之H25年度～
■センター長：山下智志
■副センター長：
逸見昌之～H26年度
黒木学H27-28年度
加藤昇吾H29年度～
■センター長：山下智志
■副センター長：加藤昇吾
■センター長：山下智志
■副センター長：加藤昇吾
R5より■センター長：加藤昇吾
■副センター長：山下智志
環境リスク研究グループ ★金藤 浩司
★金藤 浩司
環境情報に対する統計解析手法開発PJ ★金藤 浩司
資源管理リスク分析PJ ★吉本 敦
製品・サービスの質保証・信頼性研究グループ ★椿 広計
サービス科学研究センターへ転出
金融・保険リスク研究グループ ★川崎 能典
★川崎 能典～H22年度
山下 智志 H23年度～
金融・保険リスクの計量化と戦略的制御PJ ★山下 智志
ゲノム解析PJ ★間野 修平
地震予測解析PJ ★庄 建倉
リスク基盤数理PJ ★栗木 哲
データ中心リスク科学基盤整備PJ ★椿広計
医薬品・食品リスク研究グループ ★藤田 利治
食品・医薬品リスクPJ ★逸見昌之
自殺とメンタルリスクPJ ★椿広計
予測医療の実現に向けた新しい生物統計学領域の創成PJ ★松井茂之
★藤田 利治 ～H22年度
松井 茂之 H23年度～
環境情報に対する統計解析手法開発PJ ★金藤 浩司
資源管理リスク分析PJ ★吉本 敦
金融リスクの計量化と戦略的制御PJ ★山下 智志
地震予測解析PJ ★庄 建倉
リスク基盤数理PJ ★栗木 哲
データ中心リスク科学基盤整備PJ ★椿広計～H26年度
山下智志H27年度～
医療・健康科学PJ ★逸見昌之
★山下智志H27～28年度
南和宏H29年度～
医療健康データ科学研究センターへ転出
環境情報に対する統計解析手法開発PJ ★金藤 浩司
資源管理リスク分析PJ ★吉本 敦
環境資源管理研究プロジェクト推進室
金融リスクの計量化と戦略的制御PJ ★山下 智志
地震予測解析PJ ★庄 建倉
リスク基盤数理PJ ★栗木 哲
データ中心リスク科学基盤整備PJ ★南和宏
R4より
時空間データ解析PJ ★村上大輔

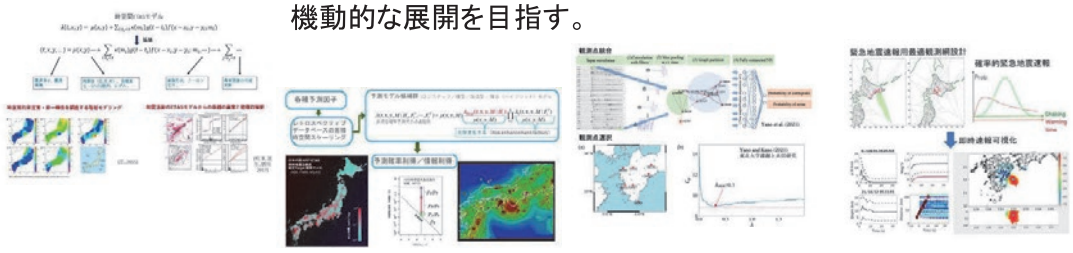
● スライド 8

情報科学を活用した地震調査研究プロジェクト（STAR-Eプロジェクト）

長期から即時までの時空間予測とモニタリングの新展開

研究代表者 庄 建倉
モデリング研究系
（兼）リスク解析戦略研究センター　教授

- 時空間ETAS モデルを始めとした長期・中期・短期という多項目の確率予測から想定された地震学的シナリオに対する配分確率のもと、地震動予測や緊急観測網の機動的な展開や首都圏観測網などの効果的な利用を目指す。
- とくに複数の予測をオンラインで実装し、それらの合わせ技（多項目複合確率予測）として確率利得を高めることを目指す。
- オンラインの確率予測の出力は、緊急地震速報システムの信頼性を向上させるための事前情報として使用し、また想定された地震学的シナリオの配分予測確率のもと地震動予測や緊急観測網の機動的な展開を目指す。

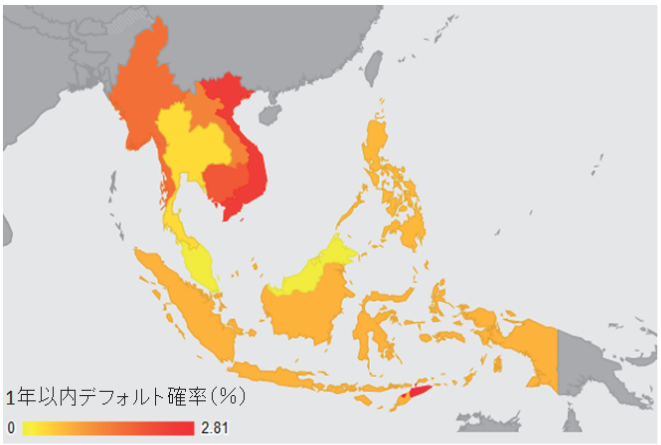


https://star-e.ism.ac.jp/

● スライド 9

ソブリンデフォルトリスクの計測と格付システムの開発

国際協力銀行（JBIC）の海外融資のリスク計量化・格付システム
　2011年開発開始、2013年運用開始、2014年～モニタリング、2016年バックテスト
国際協力機構（JICA）の海外融資のリスク計量化・格付システム
　2014年開発開始、2015年からモニタリング



シンガポール　0%
マレーシア　0.03%
ブルネイ　0.21%
タイ　0.34%
フィリピン　0.71%
インドネシア　0.77%
ラオス　1.36%
ミャンマー　1.62%
カンボジア　1.97%
ベトナム　2.42%
東ティモール　2.81%

・テクニック的には、スパースモデリングとデータ構造化（欠損値補完、異常値補正、最適セグメンテーション）のミックス
・クロスバリデーションでAR値0.7台の予測精度

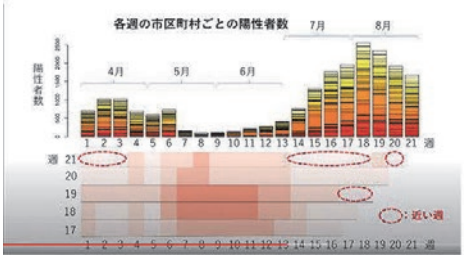
● スライド 10

時空間データ解析プロジェクトで継続推進中　【2020年3月発足】

新型コロナウイルス対応プロジェクト

このプロジェクトで開発した統計数理の技術が、新型コロナウイルスに限らず、
今後発生するであろう新たな感染症対策の中核技術となることを目指して

メンバー
PL：松井知子（教授）
椿広計（名誉教授）
船渡川伊久子（准教授）
村上大輔（助教）
鈴木和幸（客員教授）
久保田貴文（客員准教授）
下野寿之（外来研究員）



4月から8月までの各週の陽性者数をマルコフ連鎖でモデル化
各週のモデルの近さを検定することによって、週ごとの特徴を分析

市区町村ごとに時系列で陽性者数の推移をグラフ化すると、
都全体のグラフからは見つけにくい現象を捉えられる

人口密度が陽性者数に及ぼした影響」の分析

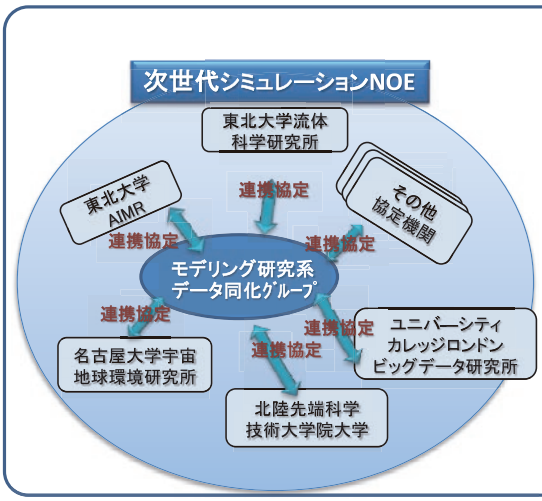
● スライド 11

次世代シミュレーション NOE

モデリング研究系データ同化グループ
Data Assimilation Group, Department of Statistical Modeling

大量の観測データに基づいた超高次元状態変数の逐次ベイズ推定など，数値シミュレーションと観測・実験データをつなぐ基盤技術であるデータ同化の研究を推進する。各応用分野における計算アルゴリズムの開発や超高並列計算機システムへの実装を通して，リアルタイムに未来予測が可能な次世代シミュレーションモデルを構築することを目指す。

“データ同化グループ”における
次世代シミュレーションNOE活動



データ同化グループ構成員

2023年4月1日現在

- 所内教授　2名（内，１名兼）
- 特任教授　1名（兼）
- 客員教員　11名
- 研究支援員等

※機構本部データサイエンス共同利用基盤施設
データ同化研究支援センターと兼務

プロジェクト

- データ同化の基盤技術開発および応用研究
- 状態空間モデルによる動画解析技術開発
- パンデミックの予測，状況解析と流行対策策定
- 大規模地理空間データの統計モデリング
- シミュレーションによる宇宙天気現象の予測技術開発と応用研究

● スライド 12

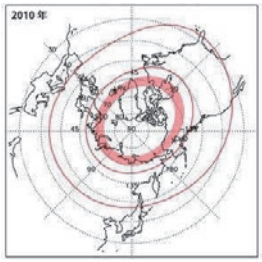
プレスリリース　（2021年9月8日）

広 NOE

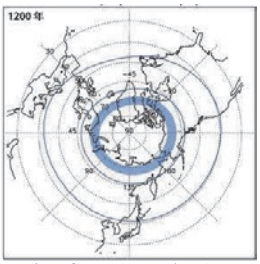
オーロラ帯の過去3000年間の変化を再現

国立極地研究所 大学共同利用機関法人 情報・システム研究機構　統計数理研究所 大学共同利用機関法人 情報・システム研究機構

オーロラ帯の位置を求める計算手法について、近年の観測データを用いて統計的に検証し、その手法を地磁気モデルに応用するという方法で、過去3000年のオーロラ帯の変化を連続的に再現することに成功



↑北極域における↑
2010年のオーロラ帯



↑本研究で明らかになった↑
1200年時点のオーロラ帯

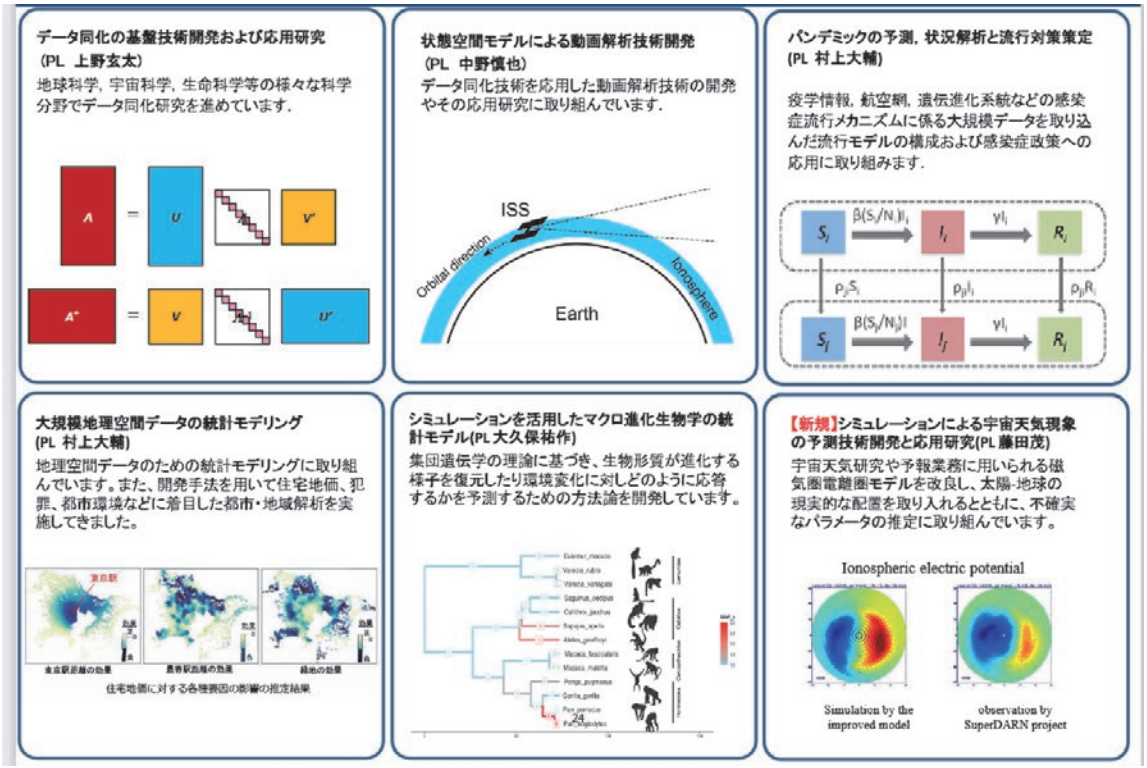
★鎌倉時代の歌人・藤原定家が『明月記』の
“１２０４年２月に京都からオーロラが見えた”記述情報との整合★

将来、激しい太陽活動が起こった場合、オーロラが広範囲に発生し、その影響による電力障害等が懸念

このような***宇宙災害の被害地域や規模を想定するための基礎をなす***成果

藤原定家画像：
https://ja.wikipedia.org/wiki/%E8%97%A4%E5%8E%9F%E5%AE%9A%E5%AE%B6
明月記画像：
https://www.kosho.or.jp/products/detail.php?product_id=232693159

● スライド 13



● スライド 14

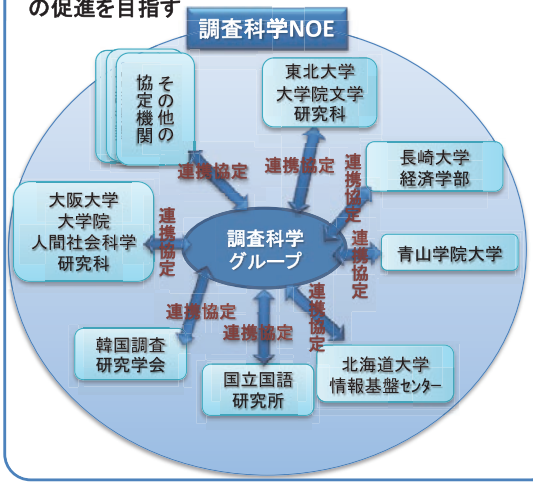
調査科学
NOE

データ科学研究系調査科学グループ
Survey Science Group, Department of Statistical Data Science

調査法をはじめとした多様な条件・環境下での統計データ取得法の設計と，収集デザインに即した統計的推論の研究，調査票情報の個人情報秘匿措置，ならびにそれらの応用に関する研究を進める。データの取得や解析・公開に関わる方法論上の研究に取り組むだけでなく，さまざまな領域における複雑な現象の解明に資する実用的研究を目指す。

“調査科学グループ”における
調査科学NOE活動

● **社会調査研究の成果を基盤とし，その更なる発展，調査科学NOE(Network Of Excellence)構築，連携，社会貢献の促進を目指す**



調査科学グループ構成員

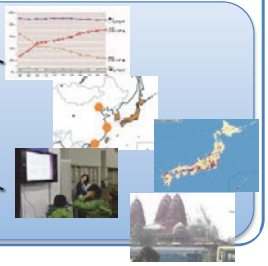
2023年4月1日現在

- 教授　1名
- 准教授　2名（内，1名は兼）
- 特任助教　3名（兼）
- 客員教員　9名（兼）
- 研究支援員等

（兼）は機構本部データ共同利用基盤施設社会データ構造化センターと兼務

プロジェクト

- 日本人の国民性調査プロジェクト
- 国際連携調査プロジェクト
- 連携研修調査プロジェクト
- 社会調査情報活用プロジェクト
- 意識の国際比較調査プロジェクト



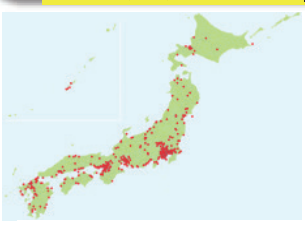
● スライド 15

日本人の国民性調査

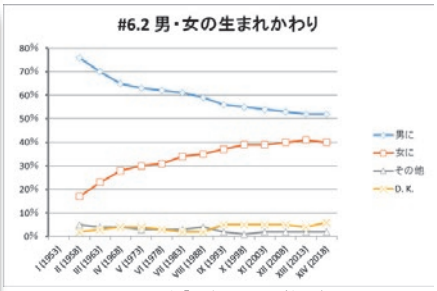
- 1953（昭和28）年から５年に１度実施
- 現在，第13次調査まで実施済み
  - ・世界的にも最長期間調査
  - ・アメリカ GSS(General Social Survey) は1972年開始
    http://www.norc.uchicago.edu/GSS+Website/
  - ・国際比較調査に発展
- 教科書，マスコミ等でしばしば引用
- 3つの意義
  - （１）社会学的意義
  - （２）調査手法の研究開発
  - （３）調査データ分析手法

2013（H25）年に第13次調査を実施
2014（H26）年１０月30日にプレスリリース@文部科学省記者会
2018(H30)年10～11月に第１４次調査を実施:2021年10月に結果を公表
“ https://www.ism.ac.jp/survey/index_ks14.html”

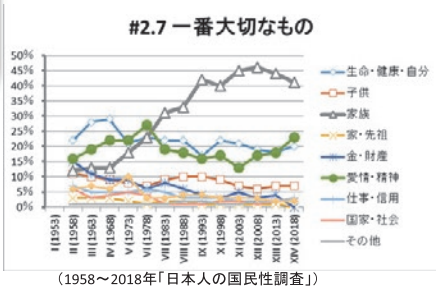
2023（R5）年秋に第15次調査を実施予定



第13次調査の400調査地点



（1958～2018年「日本人の国民性調査」）



（1958～2018年「日本人の国民性調査」）

● スライド 16

統計的機械学習NOE

統計的機械学習研究センター
Research Center for Statistical Machine Learning

● データからの推論を対象とする統計科学と，アルゴリズムを対象とする計算機科学を基盤とする機械学習分野の研究センターを設置することにより，国内研究コミュニティーへの貢献，共同利用機関としての価値向上，国内外の機関との連携等を通じた世界的研究機関としての認知度向上を意図する

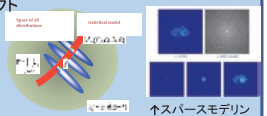
設置目的

1. 統計数理的観点からの，機械学習の理論・方法論の研究と，音声，画像，自然言語等における機械学習の応用研究の推進
2. 最適化と統計的推論との融合による新しいデータ科学の基盤的研究の推進
3. 情報幾何，カーネル法といった独自性の高い研究のさらなる展開
4. 機械学習分野の国内研究コミュニティーの活性化
5. 研究所の国際化と国際的認知度のさらなる向上

プロジェクト

- ◆ 情報幾何と機械学習プロジェクト
- ◆ カーネル法の理論と応用プロジェクト
- ◆ 最適化推論プロジェクト
- ◆ スパースモデリングの深化と応用プロジェクト
- ◆ メディアデータの統計的解析プロジェクト
- ◆ 都市インテリジェンス研究プロジェクト
- ◆ 位相的統計理論の構築とその応用
- ◆ 天文学とデータ科学プロジェクト
- ◆ 深層構造モデルプロジェクト

機械学習とは⇒経験に基づいて自動的に学習を行うシステムに関する研究分野。インターネットサービス，音声・画像などの情報処理，脳科学，ロボティクスなど広範な応用を持つ。


↑情報幾何と機械学習

↑スパースモデリングの深化と応用

センター構成

センター設立母体

- 新機軸創発センター
  - ○関数解析的推論研究ｸﾞﾙｰﾌﾟ
  - ○最適化推論研究ｸﾞﾙｰﾌﾟ
  - ○音情報解析研究ｸﾞﾙｰﾌﾟ
- 予測発見戦略研究センター
  - ○遺伝子多用性解析ｸﾞﾙｰﾌﾟ
- その他の関連分野の教員

発展的改組 合流

統計的機械学習NOE

統計的機械学習研究センター

連携協定：
- その他の協定機関
- ユニヴァーシティ カレッジ ロンドン 計算統計的機械学習センター
- オーストラリア マッコーリー大学保険数理・ビジネスアナリティクス学部
- ノルウェー科学技術大学（NTNU）電気工学通信学部
- マックスプランク生物学サイバネティクス研究所実証推論研究系
- ノルウェー産業科学技術研究所（SINTEF）情報・コミュニケーション技術部コミュニケーションシステム

福水センター長

2023年4月1日現在

構成員

センター長	1名（兼）
副センター長	1名（兼）
所内教授	9名（兼）
准教授	5名（兼）
助教	4名（兼）
特任助教	2名
客員教員	12名
研究支援員等	

● スライド 17

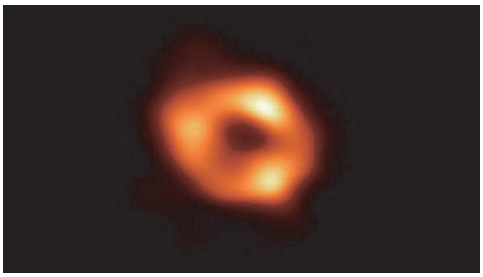
プレスリリース （2022年5月13日）

広 NOE

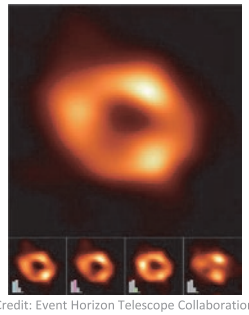
天の川銀河中心のブラックホールの撮影に初めて成功

池田教授が参加するEHTプロジェクトがブラックホールに関する新たな発表

史上初の天の川銀河中心のブラックホールいて座A*の画像


（Credit: Event Horizon Telescope Collaboration）

- ✓ この天体がブラックホールであるということを初めて視覚的に直接示す証拠。地球上の8つの電波望遠鏡を繋ぎ合わせて地球サイズの仮想的な望遠鏡を作るイベント・ホライズン・テレスコープ（EHT）によって撮影されたもの
- ✓ ブラックホールは光を放たない完全に漆黒の天体であり、そのものを見ることはできないが、周囲で光り輝くガスによって、明るいリング状の構造に縁取られた中心の暗い領域（「シャドウ」と呼ばれる）としてその存在がはっきりと映しだされる
- ✓ 今回新たに取得された画像は、太陽の400万倍の質量を持つブラックホールが作り出す強力な重力によって曲げられた光を捉えたもの


（Credit: Event Horizon Telescope Collaboration）

⇐ 天の川銀河中心のブラックホール画像ができるまで

イベント・ホライズン・テレスコープ（EHT）・コラボレーションは、EHTの観測から得られた画像を合成し、天の川銀河中心の巨大ブラックホールであるいて座A*の代表的な画像（上段）を作成

- この画像は異なる画像化手法を用いて得られた何千もの画像を平均することによって得られたもの
- これらの画像はどれも観測データとはとても整合的で平均画像には異なる画像に共通する特徴が残り、たまにしか現れない特徴は見えなくなる
- 画像は似たような特徴ごとに分類すると4つのグループに分かれ、下段の画像は4つのグループごとに平均した代表画像
- 3つのグループの画像にはリング構造が見られるが、リングの周りの明るさの分布が異なる
- 4つ目のグループはデータとは整合的であるもののリング状ではない画像を含んでいる
- 棒グラフはそれぞれのグループに属する画像の相対的な数を示す

何千もの画像が最初の3つのグループに属しているが、4つ目の最小のグループには数百個の画像しかない。棒グラフの高さは、平均画像へのそれぞれのグループの画像の相対的な寄与率を示す。

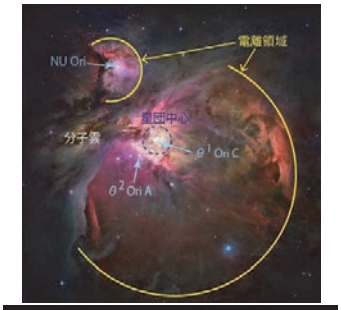
● スライド 18

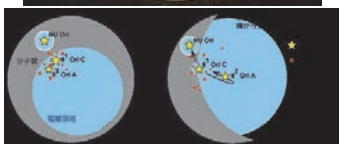
プレスリリース　（2022年6月8日）

広 NOE

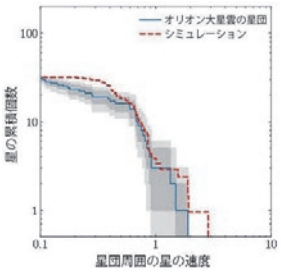
新しい高精度シミュレーションが明らかにした星団形成の現場

服部助教が参加する科研費「銀河シミュレーションで探る星団起源ブラックホール連星の宇宙史（課題番号：19H01933）」他の研究プロジェクトの一環として発表





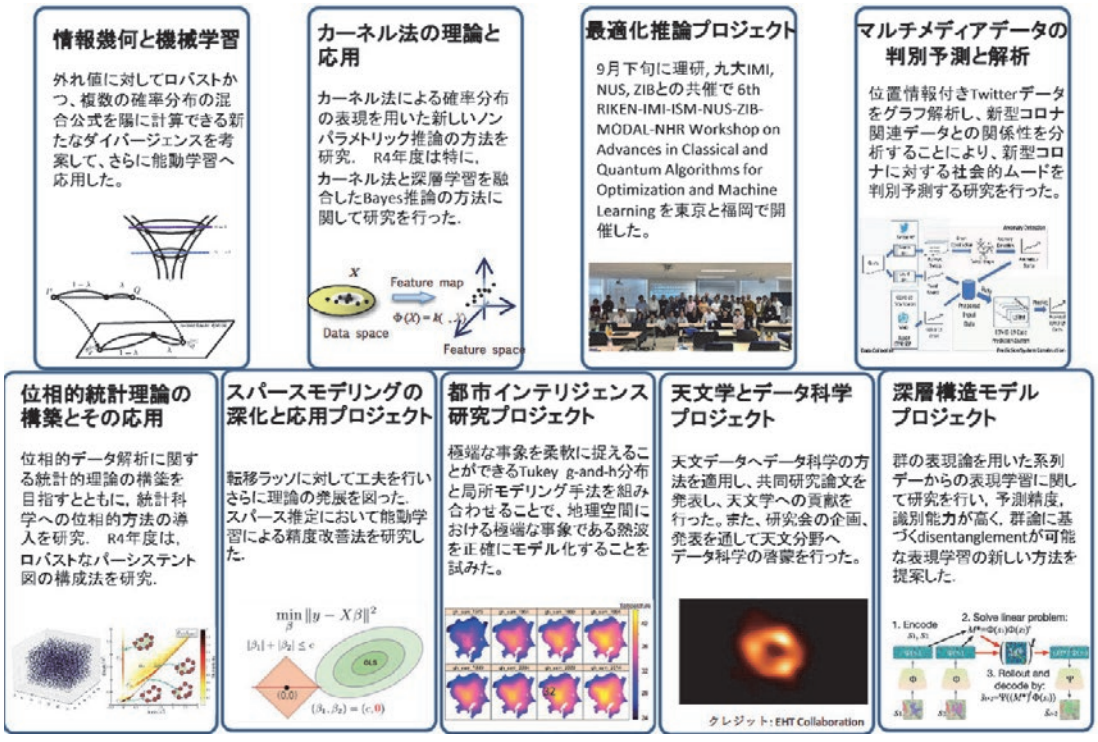
- オリオン大星雲の中心部にある星団から互いの重力による加速で弾き出された大質量星が、オリオン大星雲の電離領域の形成のきっかけとなった可能性を、スーパーコンピュータを用いた数値シミュレーションと観測データから示しました。
- 星団内の星の重力による運動を従来の手法より高精度で解く星団形成シミュレーションコードを新規開発し、それを用いることでオリオン大星雲内の星の運動を再現することができました。



- ✓ 本研究は星一つ一つの運動を追う新しい星団・銀河シミュレーションを行う「SIRIUS」プロジェクトの一環です。
- ✓ 今後は、この新規開発コードを用いたより大規模なシミュレーションを行い、未だ形成過程の解明されていない大質量星団の形成過程を明らかにしていくことが期待されます。

図1：a）オリオン大星雲。θ1 Ori Cはオリオン大星雲で最も大質量で明るい星であり、星団の中心部に存在する。NU Oriとθ2 Ori Aも大質量星であり、図2のように、重力相互作用によって星団中心から弾き出された星であると考えられる。黄色で囲われた赤～ピンクの部分は高温の電離水素で満たされた電離領域である。一方、星団の左上の黒っぽい部分は、低温の水素分子ガスが集まっている領域（分子雲）である。分子雲の中では、新しい星が生まれつつある。（画像クレジット：NASA, ESA, M. Robberto (Space Telescope Science Institute/ESA) and the Hubble Space Telescope Orion Treasury Project Teamを改変）
b）オリオン大星雲の構造の模式図（左：図1aのように正面から見た図、右：断面の予想図）。観測者から見て星団の奥側に分子雲があり、手前側に電離領域が広がっている。θ2 Ori Aは星団手前の電離領域の中、NU Oriは星団より奥の分子雲の中に半分埋もれている。

● スライド 19



● スライド 20

ものづくりデータ科学NOE

ものづくりデータ科学研究センター
Data Science Center for Creative Design and Manufacturing

データ科学の先進技術を結集し，ものづくりに革新をもたらす科学的方法論を創出する．設計と製造を対象とするデータ科学を本センターに集約し，”データ科学駆動型スマートものづくり技術” の社会普及を促進する．

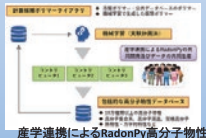
設置目的

1. ものづくりに革新をもたらすデータ科学の学術基盤・革新的アルゴリズムの創出
2. グランドチャレンジ・オープンイノベーション型プロジェクトを推進
3. データ科学のユニークな切り口から「次世代ものづくり」の在り方に対して新しいビジョンを示す
4. データ科学の先進技術の産業応用を促進
5. 産学連携による価値共創の場を形成

「データ科学駆動型スマートものづくり」

【戦略分野】製造，バイオ，マテリアル，化学，エネルギー等
【コア技術】機械学習，データ同化，シミュレーション，最適化，システム制御，ハイパフォーマンス・コンピューティング等
【主要プロジェクト】
データ駆動型材料研究の基盤技術の開発と実践，XenonPyプロジェクト，RadonPyプロジェクト，高分子物性データベース共同開発事業

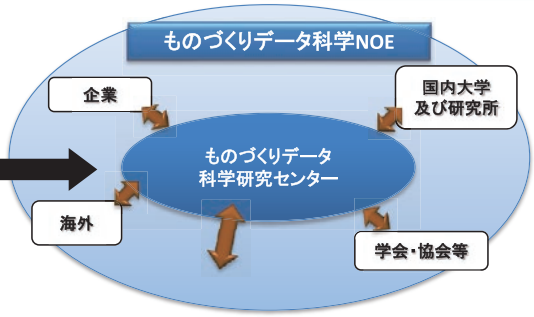
機械学習による新規準結晶及び形成・安定化ルールの発見


産学連携によるRadonPy高分子物性データベースの共同開発

センター構成

センター設立母体

データ同化研究開発センター，統計的機械学習センター，サービス科学センター等の研究者を参集



吉田センター長

2023年4月1日現在

構成員

- センター長 1名（兼）
- 副センター長 1名（兼）
- 所内教授 3名（兼）
  - 准教授 1名（兼）
  - 助教 1名（兼）
  - 特任助教 2名
  - 特任研究員 8名
- 客員教員 3名
- 研究支援員等

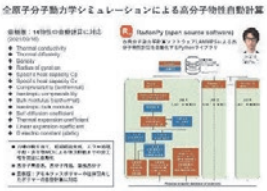
● スライド 21

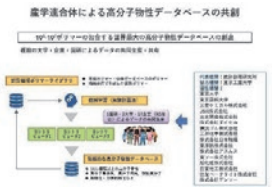
「富岳」成果創出加速プログラム

「データ駆動型高分子材料研究を変革するデータ基盤創出」

- 全原子分子動力学シミュレーションに基づく高分子物性自動計算システムRadonPyを用いて、世界最大の高分子物性オープンデータベースを構築する膨大な計算量の壁を乗り越えるため、多数の企業・大学・国研からなる連合体を形成し、データベースを共同で開発する。
- 超並列計算機におけるRadonPyシステムの拡張・高度化を図ると共に、「富岳」の計算資源を活用して10万種類以上の高分子骨格を包含する前人未踏の高分子物性データベースを創出する。

研究代表者 吉田亮
データ科学研究系　教授
（兼）ものづくりデータ科学研究センター長





https://fugaku100kei.jp/fugaku/promoting_research.html#r3-08

● スライド 22

プレスリリース 日本経済新聞，日刊工業新聞他掲載　（2021年7月20日）

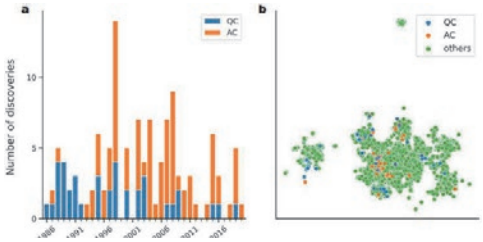
広 NOE

機械学習で準結晶を形成する化学組成を同定 ～準結晶の安定化メカニズムの解明に向けた第一歩

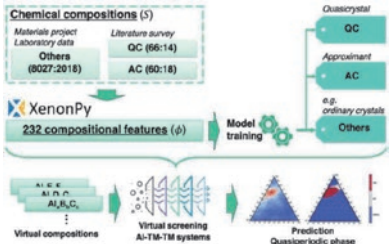
大学共同利用機関法人 情報・システム研究機構 統計数理研究所　東京大学 THE UNIVERSITY OF TOKYO　東京理科大学 TOKYO UNIVERSITY OF SCIENCE

- 統計数理研究所、東京大学、東京理科大学の共同研究グループ機械学習のアルゴリズムを駆使してこれまでに見つかった準結晶の組成パターンを読み解き、新しい準結晶の化学組成を予測できることを実証
- 機械学習のブラックボックスモデルに内在する入出力のルールを抽出することで、準結晶相の形成に関する法則を解明

この法則は五つの単純な数式で表れ、これらが準結晶研究において長年求められてきた**物質探索の設計指針になる可能性**あり



これまでに発見されたアルミニウム合金の安定準結晶
(a) 準結晶（QC）と近似結晶（AC）の発見数の年推移
(b) これまでに見つかった準結晶・近似結晶・周期結晶を含むその他（others）の組成パターンの2 次元平面への可視化。



準決勝の組成予測：機械学習（3クラス分類問題）のワークフロー。モデルの入力変数は化学組成（S）、出力は準結晶（QC）、近似結晶（AC）、周期結晶を含むその他（"Others"）を表すクラスラベル

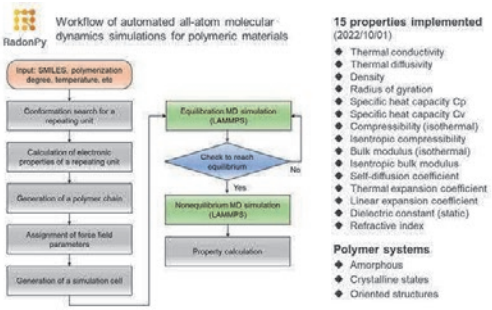
● スライド 23

プレスリリース 日本経済新聞，日刊工業新聞他掲載　（2022年１１月９日）

広 NOE

全原子古典分子動力学法による高分子物性計算を全自動化する ソフトウェアRadonPyをリリース ～高分子材料物性大地図の作成に向けた第一歩～

RadonPy



↑ *高分子物性自動計算ソフトウェアRadonPyの概要。高分子の繰り返し単位とその化学構造、重合度、温度等の計算条件を入力とし、分子モデリング、電荷計算、平衡・非平衡MD計算、平衡化完了の自動判定、収束に失敗した場合の再スタートのスケジューリング、ポストプロセス段階での物性計算等、MDシミュレーションの全工程を完全に自動実行する。ファーストリリースで公開されたバージョンは、熱物性や光学特性を含む15種類の物性を自動計算できる。*

全原子MDシミュレーションによる高分子物性計算を全自動化する世界初のオープンソースソフトウェアRadonPy

⇒高分子の繰り返し単位とその化学構造、重合度、温度等の計算条件を入力とし、分子モデリング、電荷計算、平衡・非平衡MD計算、平衡化完了の自動判定、収束に失敗した場合の再スタートのスケジューリング、ポストプロセス段階での物性計算等、MDシミュレーションの全工程を完全に自動実行する

RadonPyを用いて10万種類以上の分子骨格を包含する高分子物性オープンデータベースを創出するプロジェクトを推進し、産学の垣根を越えてRadonPyとデータベースの共同開発に取り組む（*2022年11月現在、統数研、3大学、19企業がコンソーシアム加盟、プロジェクトに参画*）

【ねらい・目標】
- ✓ **高分子材料物性の大地図の作成**
- ✓ **データを生産・蓄積する過程で、複数物性の同時分布とパレートフロンティア、さらにフロンティアを形成する新材料解明**
- ✓ **脱炭素・循環型社会やサーマルマネジメントに資する新材料の創製**
- ✓ **産学の組織の垣根を超えたデータの共創というモデルケースを社会に発信**

● スライド 24

医療健康データ科学NOE

医療健康データ科学研究センター Research Center for Medical and Health Data Science)

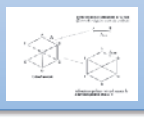
我が国におけるデータサイエンスの英知を結集し、健康科学分野（基礎医学、臨床医学、公衆衛生・疫学、保健・看護学、スポーツ科学）の先端的データサイエンス研究および高度専門統計教育を実施する。

設置目的

1. 健康科学分野における高度専門教育を実施し、医学アカデミアの先端的研究の発展を後押しする
2. 全国規模の産学官の教育研究コンソーシアム「健康科学研究ネットワーク」を設立
3. 健康科学におけるデータサイエンスの深化を目指す国内随一の学術拠点を形成
4. 画期的な高度専門教育の成果をさまざまな分野における研究リテラシーの向上に援用

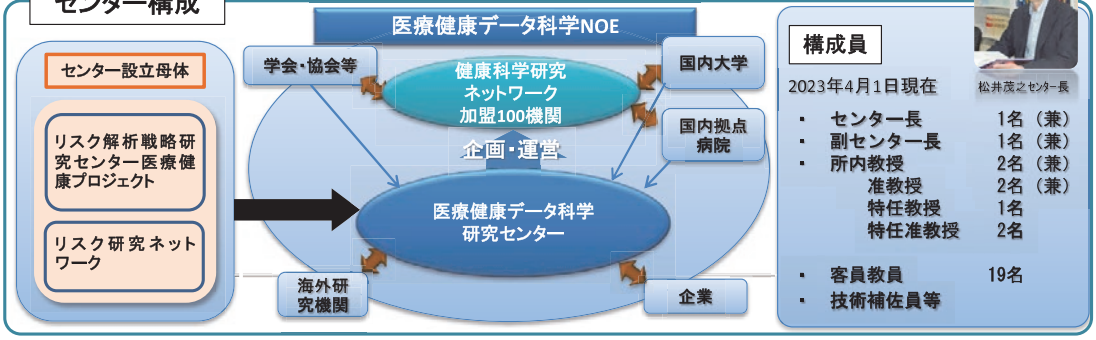
「高度データ分析専門教育事業」

【教材開発】共用教材クラウド，系統的教育コース，OJTプログラム
【教育コース】健康科学数理，公衆衛生・疫学，臨床研究，先端医療データ解析，スポーツ科学等
【コンサルテーション】現場の研究者・実務者を対象-高度な統計科学の実践-
【主要プロジェクト】医療健康データ基盤整備と計算機技術，臨床研究・臨床試験とエビデンス統合の方法論，医療健康科学基盤数理，マルチモーダル生体信号データの時空間解析，機械学習とビッグデータ解析の数理



松井茂之センター長

センター構成



構成員

2023年4月1日現在

センター長	1名（兼）
副センター長	1名（兼）
所内教授	2名（兼）
准教授	2名（兼）
特任教授	1名
特任准教授	2名
客員教員	19名
技術補佐員等	

● スライド 25

医療・健康科学分野における大学等とのネットワーク形成に基づく統計科学の高度な実践力・研究力を備えた統計専門家人材の育成

現状・背景・課題

- 国際水準の研究計画・遂行には統計科学の高度な実践が不可欠
- 近年、多層オミクス・医用画像、電子診療データ解析による疾患解明・診断法開発等、実践の場は飛躍的に拡大
- 個別領域での問題解決には統計科学の高度な専門性と研究力が必須

現状の課題

統計科学の高度な実践力・研究力を備えた「統計専門家」の不足

■ 欧米：医学系大学院/病院単位で数十名以上．国際学術誌での圧倒的な研究報告数

■ 日本：認定統計家は全国で80名程度．現在、人材育成・配置の途上

試験統計家認定者（2021.4現在）*
責任統計家：31名，実務統計家：51名

AMED生物統計家育成支援事業
2017-20年度 計60名の修士課程修了見込み
全国の大学病院等に配置予定

現場の統計専門家（特に若手・中堅）の実践力，研究力向上が喫緊の課題．全国レベルでの効果的な支援が求められる！

⇒ 統計数理研究所の統計研究・統計教育の実績とノウハウを活用

目的・手段・方策

大学・附属病院の統計専門家、医療健康科学研究者（主に、若手・中堅研究者）を対象とした専門コース、OJT，コンサルテーション等

主要プロジェクト

① 短期専門コース・集中講義
国際水準の実践基準，先端統計・機械学習の短期コース（webセミナー等），集中講義

② OJTプログラム
統計科学の実践・研究に関するOJT（小数若手研究者に対する統数研でのOJT）

③ 広域コンサルテーション
全国展開の医療健康ネットワーク（後出）内での統計の実践・方法論研究に関するコンサルテーション（先端事例・技術共有）

④ 遠隔教育
e-learning 教材の開発と運営・管理

⑤ コミュニティ・アクティビティ
国内外シンポジウム、研究集会、先導的海外研究機関の視察・情報収集

課題解決の母体となる組織・活動

統数研医療健康データ科学研究センター

- 所内10名，客員17名
- 第3期中期概算要求事業の実績
「健康科学分野における大学等との連携ネットワーク形成に基づく統計教育者人材育成および研究・専門性の強化」
医療健康科学研究者に対する系統的教育コース、e-learning教材開発等
- 医療健康データ科学研究ネットワーク
関連学会，大学研究機関，病院，企業から94団体が参画（2021.4現在）

評価指標5項目の内、3項目が「目標値に対して実績値が著しく進捗（判定A)」

次期に向け、統計専門家を対象に加えたより広範かつ充実した統計教育事業を計画！

見込まれる成果・効果

◆ 質の高い医療・健康科学研究実施の「要」となる統計専門家を短期間で育成

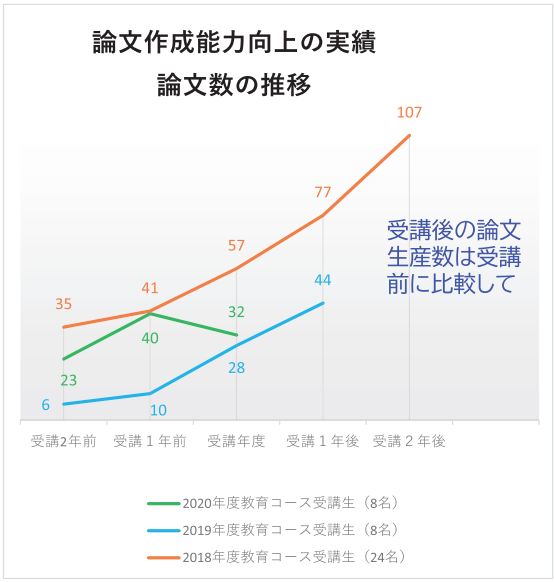
◆ 我が国の医療・健康科学研究の水準向上に直接的に貢献

◆ 疾患機序解明，AI医療含む先端医療開発等の医療・健康科学分野のデジタルトランスフォーメーション推進のための統計・データ科学の基盤構築・強化

*一般社団法人日本計量生物学会による認定

● スライド 26

これまでの活動成果とコンソーシアムメンバー



IQVIA ソリューションズ ジャパン株式会社
愛知医科大学
秋田大学医学部公衆衛生学講座
アステラス製薬株式会社
一般社団法人国際栄養食品協会（略称 AIFN）
一般社団法人 日本計量生物学会
一般社団法人日本統計学会
一般社団法人日本糖尿病・生活習慣病ヒューマンデータ学会
一般社団法人ヘルシィエイジング学会
一般社団法人ヘルスデータサイエンス学会
一般社団法人ヘルスデータサイエンティスト協会
一般社団法人臨床疫学研究推進機構
岩手医科大学衛生学公衆衛生学講座
大分大学医学部附属病院総合臨床研究センター
大阪医科薬科大学医学研究支援センター医療統計室
大阪公立大学大学院医学研究科産業医学
岡山大学生殖補助医療技術教育研究センター
鹿児島大学病院臨床研究管理センター
株式会社エクサウィザーズ
株式会社エヌ・ティ・ティ・データ経営研究所
情報未来イノベーション本部デジタルコグニティブサイエンスセンター
株式会社 Open Health Initiative
株式会社健康保険医療情報総合研究所
株式会社 ニッセイ基礎研究所
川崎市総合リハビリテーション推進センター
北里大学病院
岐阜大学医学部附属病院 先端医療・臨床研究推進センター
京都大学大学院医学研究科社会健康医学系専攻
京都府立医科大学大学院医学研究科生物統計学
杏林大学医学部 IR 室
熊本大学病院
久留米大学バイオ統計センター
群馬大学医学部附属病院臨床試験部
慶應義塾大学医学部医療政策・管理学教室
慶應義塾大学医学部衛生学公衆衛生学教室
慶應義塾大学ウェルビーイングリサーチセンター
慶應義塾大学大学院健康マネジメント研究科
慶應義塾大学病院臨床研究推進センター生物統計部門
公益財団法人 神戸医療産業都市推進機構 医療イノベーション推進センター
高知大学医学部公衆衛生学教室
神戸大学医学部附属病院
公立大学法人 名古屋市立大学
国立がん研究センター研究支援センター生物統計部
国立国際医療研究センター
国立精神・神経医療研究センター
国立大学法人 佐賀大学医学部附属病院 臨床研究センター
国立大学法人 東京医科歯科大学 統合イノベーション推進機構 医療イノベーション推進センター
国立大学法人 浜松医科大学 医学部附属病院 臨床研究センター
国立長寿医療研究センター先端医療開発推進センター
埼玉医科大学大学院医学研究科社会医学研究系
産業医科大学病院臨床研究推進センター
自治医科大学公衆衛生学教室
島根大学医学部附属病院医療情報部
順天堂大学医学部附属順天堂医院 臨床研究・治験センター
臨床研究支援室
聖マリアンナ医科大学予防医学教室
千葉大学医学部附属病院
地方独立行政法人 大阪府立病院機構 大阪国際がんセンター がん対策センター
地方独立行政法人 東京都健康長寿医療センター
筑波大学 医学医療系 生物統計学
筑波大学人間総合科学研究群公衆衛生学学位プログラム
DOT ワールド株式会社
帝京大学大学院公衆衛生学研究科
東海大学大学院医学研究科
東京医科歯科大学 M&D データ科学センター生物統計学分野
東京医科歯科大学救急災害医学講座
東京医科歯科大学大学院 医療政策情報学分野
東京慈恵会医科大学
東京大学大学院医学系研究科精神保健学／精神看護学分野
東京大学大学院医学系研究科生物統計情報学講座
東京理科大学工学部情報工学科寒水研究室
東京理科大学 データサイエンスセンター
統計数理研究所 医療健康データ科学研究センター
統計数理研究所 リスク解析戦略研究センター
東邦大学医学部社会医学講座医療統計学分野
東北大学大学院医学系研究科
特定非営利活動法人 JORTC
鳥取大学医学部
富山大学 第一内科
長崎大学大学院医歯薬学総合研究科公衆衛生学分野
長崎大学病院 臨床研究センター
名古屋大学医学系研究科生物統計学分野
奈良県立医科大学附属病院臨床研究センター
日本計算機統計学会
日本行動計量学会
日本製薬工業協会
日本保険年金リスク学会
兵庫医科大学医療統計学
兵庫医科大学 医学部公衆衛生学講座
広島大学病院 広島臨床研究開発支援センター
広島大学医学部
福岡大学病院
福島県立医科大学
藤田医科大学 医学部公衆衛生学
北海道公立大学法人 札幌医科大学
北海道大学環境健康科学研究教育センター
北海道大学大学院医学院社会医学講座
北海道大学病院 医療・ヘルスサイエンス研究開発機構
ミリマン・インク
明治大学科学リテラシー研究所
リアルワールドデータ株式会社
和歌山県立医科大学附属病院臨床研究センター

2023年3月現在
医療健康データ科学研究
ネットワーク加盟
100機関

● スライド 27

● NOE 形成事業の概要について

【司会：椿所長】それでは、次は、NOE 形成事業の概要説明で、まさに司会進行役の山下副所長が、NOE 推進室長として説明するので、一時、私が司会進行役をさせていただきたいと思います。それでは、資料 3 をご覧いただければと思います。それでは、山下副所長、ご説明をよろしくお願いします。

【山下副所長／NOE 推進室長】NOE は全部で 6 個あり、6 個のうち、今回、3 つ個別に説明させていただくのですが、残りの 3 つについて、一体どのようなことをやっているかを中心に、NOE 事業について説明させていただきたいと思っています。

まず NOE という言葉ですが、Network Of Excellence という英語の略です。われわれ統計数理研究所は、大学共同利用機関といって、大学に対するサービス、もしくはコネクションの中心になるというのが、元々の役割として持っています。それを具体化する学術ネットワークのハブとなるように、いろいろな制度、人、施設、そのようなものを整えるというのを、NOE 形成事業と言っています。話が見えにくくなる原因の 1 つが、NOE 形成事業と NOE 型研究センターというのは、必ずしも 1 対 1 でつながっていない。NOE は、先ほど言ったとおり、6 個あるのですけれども、研究所内には、NOE 型研究センターというのは、4 つしかありません。4 つ挙げますと、リスク解析、機械学習、ものづくり、医療健康です。

さて、何がどうなっているかというと、もう少し整理させていただきたいのです、リスク科学 NOE という NOE 事業があり、この内容を見ると、所内にリスク解析戦略研究センターというセンターがありまして、あと所外にリスク研究ネットワークという組織を結び付けるためのコア事業、コアネットワークがあります。だから、ネットワークのほうは、所外の人が委員長になっていただいています。そのような、所内のもの、所外のものを集めて、リスク科学 NOE というものができています。

例えば、調査科学 NOE というのは、これは、所内にはセンターというものがありません。では、担っている人たちは誰かというと、データ科学研究系という、先ほど言った、横糸の 1 グループが、この NOE を担っています。所外の DS 施設という情報・システム研究機構に存在する 1 つの施設の中に社会データ構造化センターというセンターがあり、その中で主に研究しているという構造になっています。

それで、NOE は全部で 6 個あるけれど所内の NOE 型研究センターは 4 つしかなく、調査科学 NOE と次世代シミュレーション NOE がこれにあたります。この 2 つについては、この DS 施設の中で活動しているので、統数研のこの資料の中には出てこないという仕組みです。NOE は統数研のものなのですけれども、センターは、統数研のものではない形が含まれて、複雑になっていると考えてください。

DS 施設とはわれわれの上部機関である情報・システム研究機構が持っている 1 つの研究施設です。

歴史的な変遷を見ると、リスクの NOE から始まりました。この時の初代センター長は、

今の椿所長です。それが発展しまして、次世代シミュレーション、調査科学、それで、少し遅れまして、機械学習、サービス科学というものが出来上がりました。その後、世の中のニーズの変遷と、あとは、いわゆる情報・システム研究機構の組織改革、もしくは、また、所内でのいろいろな人材の関係も含め、NOE がいろいろと変わっています。いろいろな変遷の結果、DS 施設の中にあるものと、研究所本体に残った NOE 型センターというものが存在します。いわゆる現在の形になったとご理解いただければと思います。

それで、6 個の NOE は、研究系という横糸によって、それぞれ結び付いています。2 つは DS に、4 つは統数研に残っているわけですけれども、その全体を NOE 形成事業運営委員会、NOE 推進室というもので支援しまして、また、一方、この会議、顧問の先生方からご意見を頂き、6 つの NOE の運営の参考にさせていただくというのが、NOE 形成事業の全体像です。

さて、NOE の予算がどのようになっているかという質問が、前回の顧問会議で寄せられましたので、新たに作った図です。この図自体は会計担当の私が、大まかに入れた数字なので、正確ではありません。NOE の運営予算というものは、1 つは、テーマごと、トピックごとにお金を頂いている直接経費です。例えば、医療健康の場合は、概算要求で獲得できた予算が、現在は、ミッション実現分という運営費交付金に変わっていますけれども、その 6,000 万が、ほぼつぎ込まれています。あとは、ものづくりの場合は、富岳プロジェクトや企業との共同研究、そのようなものが活動の中心になって、予算的にはそのようなところから出ています。リスク科学の場合は、地震の STAR-E という国プロと、所長裁量経費から来ていまして、調査科学と次世代シミュレーションは、これは DS の施設なので、DS の予算を使うことができているということです。

ちなみに、所長裁量経費というのは、科研費や国プロからの間接経費、自己収入、運営交付金ということで、基本的には、使途に制限のない予算を NOE に投下しています。これをやることで、全体で 4 億近く投入しているのですが、統数研の予算規模は、全体で 17 億ぐらいです。運営交付金が 13 億です。そのような規模の中で、4 億以上と、統数研にとって重要ポイントということになっています。

それでは、簡単ですが、各 NOE の活動状況について解説させていただきます。まずリスク科学 NOE なのですけれども、活動の中心は、所内にありますリスク解析戦略研究センターです。リスクセンターは、大変規模の大きいセンターでありまして、また、一番古いセンターです。それで、現状、7 つのプロジェクトをやっています。まず、リスク科学共通の問題点を取り上げるリスク数理、リスクに関する数学です。あとは、リスクに関するデータベースの議論という、分野共通型のテーマと、あとは、環境、資源、金融、地震、時空間ということで、テーマ、分野限定型の研究が 5 つということで、全部で 7 つの活動を続けています。

この 7 つというものは、常に新陳代謝していまして、このような変遷をたどっています。消えたものもありますが、例えば、一部のプロジェクトは、センターとして独立しました。

そのようなものも存在します。現在、センター長は加藤先生で、私が副センター長ということで、センター長、副センター長は、今年から入れ替わったことになっています。

具体例として幾つか紹介します。まず、地震予測です。緊急地震速報システムの信頼性向上のために、実務的な貢献をしているチームです。他、例えば、金融チームの活動例です。これは JBIC（国際協力銀行）と JICA（国際協力機構）との共同作業なのです。ソブリンリスク、国の破綻リスクについてのモデルを作って、それで毎年チェックしている、モニタリングしているという活動をしています。それから、新たにできた時空間データ解析プロジェクトでは、いわゆるパンデミックリスクについて、感染の広がりのモデルについて研究しているところです。

次に、データ同化グループです。データ同化グループの特徴は、シミュレーション技術です。要素間に一種の方程式を使いまして、あとは、観測値を境界条件にして、それで連立的に説いていくような形の解き方をすることを、データ同化手法と言いますが、それを使ったことによる方法論的な研究グループです。応用先ですけれども、特に有効に使われている部分は、気象や時空間に関するようなことが多いです。例えばですけれども、オーロラの状況というものが、実は古典の情報からオーロラの状況というものが再現できるという話であったり、地理データの応用であったり、あとは、物理的な空間を意味しているわけではないのです。遺伝子情報の解析であったり、あとは、そのような実証的分析よりも、理論的分析が含まれているという形が、データ同化、次世代シミュレーション NOE の状況です。

調査科学グループに関しては、後ほど特別に時間を設けていますので、ここでは省略させていただきます。国民性の調査等、統数研の看板事業を行っているということです。

機械学習センターにおいても、後ほど説明がありますので、省略させていただきます。近年ブラックホールの撮影への貢献や、あとは新しい星団の形成など、どちらも宇宙なので、似たような話かなと思われがちなのですけれども、こちらは画像解析で、こちら側は、物理シミュレーションです。だから、内容としては、全然違うのですけれども、そのようなものが研究として成功しているということです。

ものづくりの解説についても、後ほど行いますので、今回、ここでは省略させていただきたいと思います。先ほど所長からの説明がありましたように、化学合成、有機物、無機物、両方ですけれども、そのようなもの物理実験からシミュレーション実験への大きなかじ取りをしているというセンターです。

最後に、医療健康データ科学研究センターです。これは、前の第 3 期中期で概算要求が通って、事業化されたものです。医療の研究そのものも、もちろん含まれるのですけれども、重点としては、医療の中での統計学者の育成ということです。なぜこのようなことをしているかと言いますと、医療関係の論文というものは、必ず統計の専門家というものをグループの中に入れて、統計的なデータを正しく扱っているという証明が必要になってきています。それでなければ、査読が通らないというのが現状です。これが、なかなか日本は達成できなくて、世界での論文順位というものが、30 位や 40 位など、先進国ではないレベルまで下が

ってしまっていました。

そのようなものを解決するために事業を立ち上げて、医療の先生方の中で統計的な先生方を育てるということを、集中講義やOJTプログラム等で行っていました。今期から新しく広域コンサルテーションというものを行いまして、臨床において、どのような点が統計的問題になっているかというのを、実際コンサルを行って、それをまとめて、マーケットの皆さんが自由に利用できる情報をきちんとストックしていこうという活動をしています。

実際、われわれの活動に協力していただいている機関というものは、全部で100機関を超えてきています。われわれが行った研修生がどれぐらい論文を出せるようになったかというもののデータなのですが、受講年度を真ん中に置いています。1年前、2年前と、受講後の1年後、2年後というものを比較すると、明らかに受講することによって、論文の生産数が増えたということが見て取れます。これが、因果関係で、われわれが原因で、他は原因ではないようなことは、決して言えるわけではないとは思うのですけれども、数字的には、成果がきちんと出ていると認識しています。

以上、簡単でしたがわれわれのNOE活動全体についてと、本日発表の用意をしていないNOEについて紹介させていただきました。ご質問等がありましたら、よろしくお願いします。

【司会：椿所長】以上で説明が終わりました。NOE形成事業の概要でした。いかがでしょう。顧問の皆さま方、ご質問があれば、よろしくお願いします。

本日、機械学習やものづくりについて、あるいは調査については、後ほど説明があるのですけれども、他のNOEに関する質問等々があれば、それから、先ほどの川崎からの説明も含めて、何かご質問があれば、よろしくお願いします。

【伊藤顧問】すみません。伊藤ですけれども、よろしいでしょうか。

【司会：椿所長】はい、伊藤先生。よろしくお願いします。

【伊藤顧問】ご説明、どうもありがとうございました。活動内容というよりは、予算のことに関してお伺いしたいのですけれども、39ページのところで、全体の予算の話をしていただきました。先ほど椿所長のほうからも予算の話がありました。産業界からの外部資金の件なのですが、全体で6,000万という額は、非常に多くて、予算総額4億のうち15%ぐらいが、産業界から来ているということで、大変素晴らしいと思います。けれども、気になるのは、共同研究が全部で20件あって、その額が3,300万。ということは、1件当たり百数十万ということになるのかなと思います。共同研究ですから、統数研にとってもメリットもあることをおやりになっているので、このような共同研究をする時、金額の大小は、特に問うわけではないのですが、何か根拠があって、すなわち積み上げ的な作業によって、額を決められているのでしょうか。というのは、今、国の機関が保有する施設や技術を広く共用するという方向に動いていて、共用することによって、利用者が、どのような価値を得るのかという、その価値をきちんと把握した上で、金額は考えるべきではないかという議論がありま

す。このような議論もあるのでこの金額は、どのようにお考えになっているのかということをお伺いしたいです。

【山下副所長／NOE 推進室長】金額については、契約がさまざまです。割と小さめのものについては、100 万、200 万レベルのものが、とても多いです。それは、研究者個人同士のつながりでの契約が多いです。に対して、例えば、先ほどの三菱ケミカルなどはこちらに研究室をお貸ししますという、そのような規模で組織と組織のきちんとした契約になります。モデル金額のようなものを設定しまして、その金額を払い込んでいただくという仕組みになっています。かなりいろいろなスキームがありますので、一概に説明できないのですけれども、大きいほうが、やはりレディメイドの制度の上に乗せたいという形です。

【伊藤顧問】それは、2 件ある研究部門という話ですね。

【山下副所長／NOE 推進室長】そうですね。

【伊藤顧問】分かりました。URA の制度も、さらに拡充していただけるということなのですけれども、統数研が持っている唯一無二の研究力というものは、とても大きな価値です。ですから、その価値を産業界の方にご説明いただいて、その上で産業界側も納得していただけるような形で、きちんと共同研究の経費というものを考えたほうがいいのではないかなと思います。これはコメントです。

【山下副所長／NOE 推進室長】ありがとうございます。

【司会：椿所長】どうもありがとうございました。他にいかがでしょうか。

【蒲地顧問】よろしいですか。

【司会：椿所長】蒲地先生、よろしくお願いします。

【蒲地顧問】この NOE 形成事業というのは、統数研がネットワークの拠点としての活動という意味があるということのご説明があったかと思うのですけれども、これは、今日は、どうしても概略のところでの紹介でしか、時間や枚数の都合などで、それは仕方がないことだと思います。今日の紹介していただいた研究成果は、素晴らしいものが幾つも並んでいるのですけれども、これは、統数研だけでできたものではなくて、ネットワークで、他の 100 以上あるところとの共同な、あるいは、つながりがあっての話だと思います。だから、そこのところをもう少し強調されるような紹介のされ方をしていただいたほうが、NOE という元々の意味からして分かりやすいかと。各センターに、センター構成ということで、どこと関わり合いはありますという図は出ていたのですけれども、それぞれのテーマに、どこと一緒にやって、出てきました。このような成果を得られましたと。

その中で統数研としては、このような役割をして、他は、どのような役割があったということが、少し分かるような、そうなると、資料が膨大になって、申し訳ないのですけれども、そこまでしなくてもいいとは思います。でも、何かそれらしい紹介の仕方をしていただいたほうが、NOE という元々のネットワークのハブとしての役割をやっているのだということが示せるのではないかという気がしました。

【山下副所長／NOE 推進室長】ありがとうございます。このようなブラックホールの研究

など、当然われわれだけでは、できるわけではありません。

【蒲地顧問】そうですよね。

【山下副所長／NOE 推進室長】ですから、このようなものの研究成果を見せると同時に、この時に使われた、いわゆる研究ネットワークを一気に説明すると、もっと NOE らしい説明になるということと理解しました。ありがとうございます。

【司会 : 椿所長】どうもありがとうございます。他は、いかがでしょうか。後半のほうでは、医療健康やリスク自体の説明はありません。後のほうで、もちろんコメントを頂いてもいいのですが、今、もしご質問等があれば、よろしくお願いします。

【小宮山顧問】日本製薬工業協会の小宮山です。

【司会：椿所長】小宮山先生、よろしくお願いします。

【小宮山顧問】今日は、医療健康データ科学のネットワークの話の詳細はないということなのですけれども、質問ではなくて、感想になってしまいます。コンサルタント事業、これを本年度というか、先ほど規定を拝見したのですけれども、3 月から施行になったということで、多分われわれの業界側でも、十分に周知されていないので、急いで宣伝しなければならないなと思いました。この統数研とのネットワーク、NOE とのつながりを広げていく、あるいは強化していく上で、とてもいいツールになるのではないかなと本日思いましたので、もう、すぐにでも宣伝、あるいは、どのようにしようかという話を始めたいと思います。ありがとうございました。

【司会：椿所長】どうもありがとうございました。松井センター長、よろしいですか。

【松井（茂）教授】コメントを頂きまして、ありがとうございました。センター長をしています松井ですけれども、おっしゃるとおりで、コンサルテーション事業は、初めての事業で、いろいろと調整などで少し時間がかかったのですけれども、3 月にようやく、年度内ぎりぎりに開始できる状態になったところです。年度明けの 4 月になってから、徐々に今、開始しているところです。おっしゃるように、まだ全然アピールの部分が十分できていません。6 月の末にネットワークのシンポジウムを開催しますが、そのテーマは、コンサルテーションということでちょうど案内したところです。そのシンポジウムの中で、コンサルテーションの背景や詳細も含め、ネットワークの事業として紹介することになります。基本的には、まずはネットワークに入っていただいている加盟団体の方々が対象になりますけれども、そこで、しっかりアピールしていきたいと思っています。

恐らくは、まだどのような中身かということは、なかなかイメージしにくい部分があると思います。例えば、このような具体的なテーマがありますなど、そのようなことも紹介しつつ、よりアピールの場を広げていきたいと考えています。製薬協のほうでも、ぜひアピールしていただければと思います。どうもありがとうございます。

【司会：椿所長】どうもありがとうございます。6 月末に、まさにこのコンサルテーションをテーマとしたシンポジウムがあるということですので、ぜひそれも小宮山顧問にもご参加いただければと思います。

【小宮山顧問】ぜひよろしくお願いします。
【松井（茂）教授】お願いします。
【司会：椿所長】他、いかがでしょうか。山形先生、よろしくお願いします。
【山形顧問】すみません。最後、移動前に少し質問をさせていただきたいのですが、実は、このコンサルテーションにも関わっているのですけれども、自治体さんや企業さんで、いろいろなDXということが進んでいますが、特に深刻なのは自治体で、いまだにファクスを使ったり、全然フォーマットが違ったり、データがきちんと整備されていないという現状があると思います。と同時に、AIの手法はChatGPTで、もうほとんど私たちも、正直言って、専門家知見が、それもChatGPTに入れると出てきて、IPCCも、もう要らなくなるのではないかというぐらい正確な評価が出るような時代になってきたわけです。そのような時に、あまりに格差が大きい現状で、どのようにしたらいいのかというのが、私には少し分からなくなってしまっています。ここで、さらにChatGPTの性能を上げるような機械学習の技を磨き続けるというのは、もちろん価値があると思うのですが、社会全体の最適化を考えると、むしろ、今までのやり方では、もう全く間に合いません。ですから、根本的にDXを考え直して、ChatGPTなどの活用も含めて、どのように社会にAIを取り込むかということを考え直さなければならないのではないかという気もします。これをやると、研究を放棄するようなことになりかねないという面もあるのですが、そのようなところのバランスというものは、どのように考えていらっしゃいますか。
【司会：椿所長】これは、いかがでしょうか。非常に重い課題を頂戴したように伺いましたけれども、NOE活動として何かコメントはありますか。これは、場合によっては、私ども統数研だけではなくて、ROIS、情報・システム研究機構全体に対する宿題と、今、山形顧問のお話は聞こえました。まさに、今、DX化に関しては、ROIS全体で、かなりいろいろなことを機構長も含めて考えていらっしゃると思いますけれども、いかがですか。これは、まさに今のわれわれの課題として受け止めて、山形先生から非常に宿題を頂いたというように思いますが。室長としては、山下先生としては、いかがですか。
【山下副所長／NOE 推進室長】少しテーマが大き過ぎて、NOE形成事業というより、もう一段階上の情報・システム研究機構としての態度が求められているのかなというイメージを持ちます。
【司会：椿所長】非常に重要な論点だと思います。
【山下副所長／NOE推進室長】さらにその上の文科省の学術関係の、いろいろな部署にも、それぞれきちんと考えてもらわなければならない課題かなという気はします。ただ、われわれは、われわれの組織で与えられた条件の中で、独自に考えていかなければならない面もありますので、マクロの方向と、逆に、先生1人ずつが、それこそ本当に人工知能というのが発達した現状を考えて、自分自身の研究態度をどのようにするかというミクロの面と、その両方の組み合わせで、やっと何か糸口が見えてくるのかなという感想を持ちました。とても仕方のない感想で申し訳ないのですけれども、私の頭に浮かんだものは、以上、そのような

感じです。

【司会：椿所長】顧問の山形先生、正直、私も、これは、われわれの研究所に限らず、非常に大きな課題、宿題を頂いたと思います。ですから、今日すぐに答えられるということではないのですけれども、ぜひ中期的に、かなりもうChatGPTがこれだけ速いスピードで世の中に入ってきたということからすると、その背後にある統計数理科学としても、かなり早急にいろいろな戦略を考えなければならないというのは、そのとおりだと思います。

【山形顧問】ありがとうございました。

【司会：椿所長】他、いかがでしょう。

【副島顧問】すみません。日本銀行の副島です。

【司会：椿所長】副島先生、よろしくお願いします。

【副島顧問】今の山形先生のご質問と、それから、最初のセッションであった人材育成というところが、少し関係しているのかなと思いました。金融研究所は、日本銀行のCenter of Excellenceです。それゆえ、Excellenceという言葉が示す機能や役割についてNOEのE（Excellence）について非常にシンパシーを感じています。われわれが課題として日々感じているのは、技術の実務問題への応用です。日本銀行の業務の現場を抱えている部署にこそ問題はあって、それに対して、金融研究所の研究を役立てていこうとしています。

けれども、そこでスキルセットとして必要になってくるのが、ITシステムに関する知識です。①業務知と、それから、②データサイエンティストなり、統計なり、あるいは計量経済学なり、フィナンシャルエンジニアリングなりという、そのような分析技術というところと、③これを業務に落とし込むITシステム、この3つの素養というものをそろえていなければ、なかなか研究が実務にまでたどり着きません。今の実務というのは、全てITシステムで動いているので、ITシステムの中に研究を落とし込んでいかなければならないわけです。しかし、そこにギャップがあります。

なぜなら、いまだにITインフラの開発というのは、昔のウオーターフォール型のオンプレミスで動くようなシステムの開発の仕方がほとんど主流です。クラウドとかCI/CD、アジャイルといった今どきの開発の仕方というところに対するなじみがありません。加えて、分析者は分析者で、ITの実装は、われわれの仕事ではありません、研究ではありませんと考えがちです。でも、今起こっていることというのは、これは、大規模自然言語モデルなどは典型だと思うのですけれども、ITシステムの研究と、それから、機械学習なり人工知能なりの研究はセットになって行われています。それらをシステム実装する時には、必ずインプリメンテーションというところに対するスキルセットがなければなりません。この3つのスキルのギャップというものを、どのように埋めていこうかというところが、とても悩ましい課題になっています。

DXが進まない一つの大きな原因は、ITインフラに関する一般の働く人の知識が圧倒的に不足していることにあります。恐らく、これは、研究者やアナリストも同様です。先ほど少しデータベースの話が出てきたと思いますけれども、では、その大規模データを構築する

としましょう。それをどのようなOS群、あるいは、コンピュータ群の中にどうネットワークを構築して、あるいはクラウドの上に構築していくかという話になった時に、必ずITの問題に直面します。なのでITシステムと研究や分析を切り離さないほうがいいのではないのかな、一番弱いところを研究対象の外に置かないほうがいいのではないかなと考えています。金融研究所では、最近は、IT インフラ自体も研究の対象なのであると、考え方を切り替えたアプローチで臨んでいます。その辺が DX のヒントになってくるのではないかと考えています。以上です。

【司会：椿所長】コメント、ありがとうございました。先ほど最初に申し上げました私どもの新研究系の構想の中には、実装を担当するリサーチエンジニアというもの自体の雇用、増員ということも必要だといろいろと意識はしていたところですけれども、むしろ、そのような分野自体の研究活動というものも、私どものような、ある意味で、理論と実践の研究機関の中で、非常に重要ではないかというご示唆と拝聴しました。

これは、NOE 推進においても、今のような実装やシステム、ICT との関係性というのは、今後、非常に大きな問題になってくるのだろうと思います。どうもありがとうございました。

だいぶ時間の進捗も遅れていて、気になっています。本来ですと、もう休憩の時刻が近いのですけれども、もう一件、進めさせていただくということでよろしいですか。それでは、もう一度、山下室長に司会を戻します。

Research Center for Statistical Machine Learning

第4回 統計数理研究所 NOE形成事業 顧問会議

統計的機械学習研究センター

NOE活動紹介・事業計画

統計的機械学習研究センター長
教授　福水　健次

資料６

統計数理研究所
NOE（Network Of Excellence）形成事業

統計的機械学習NOE活動紹介

統計的機械学習研究センター長　福水 健次

2023年5月16日

統計数理研究所 NOE形成事業

● スライド 1

統計的機械学習NOEの概要・趣意

機械学習

経験やデータに基づいて自動学習を行うシステムに関する研究分野で、データからの推論を扱う統計科学とアルゴリズムを扱う計算機科学を基盤とする。現在の人工知能の基盤的技術。

応用分野は多岐にわたる：ロボティクス・情報通信・インターネット上のサービス技術、の工学から脳科学・天文学などの自然科学まで。

統計的機械学習NOE　2012年1月発足。

目的

・NOE活動として、統計的機械学習分野の研究コミュニティの発展に貢献

・さまざまな研究プロジェクトを国内外の研究者と共同で推進し、価値の高い研究成果を生む。

活動の4つの柱

- 研究プロジェクトの推進
- 諸機関との協定締結とそれに基づく共同研究の推進
- 国際連携：国際共同研究、国際ワークショップの開催
- 機械学習コミュニティへの貢献による共同利用機関としての価値向上
  研究会，セミナーの主催・共催

● スライド 2

統計的機械学習NOE中核組織構成・人員配置

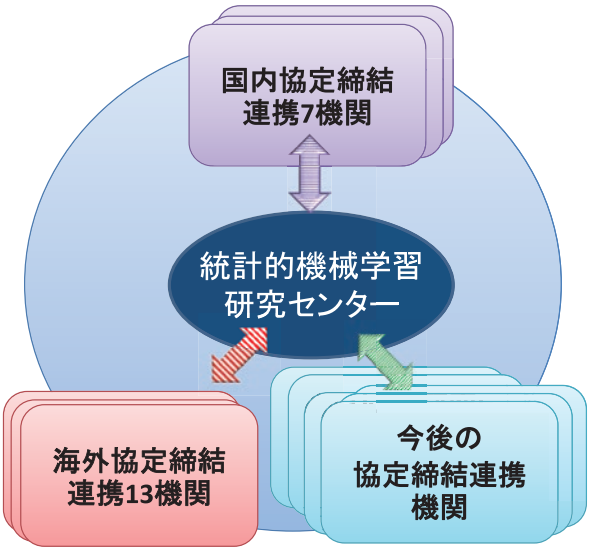
統計的機械学習研究センター（2023年5月現在）

センター長	福水健次	特任助教	坂本航太郎
副センター長	松井知子		中島秀人
教授	伊庭幸人	特任研究員	Zheng Ning
	伊藤聡		豊田祥史
	池田思朗		
	栗木哲	客員教授	Arthur Gretton
	間野修平		川野秀一
	藤澤洋徳		Konstantin Markov
	南和宏		後藤真孝
	日野英逸		藤澤克樹
准教授	持橋大地		品野勇治
	小山慎介	客員准教授	今泉允聡
	坂田 綾香		山田誠
	田中未来		小林景
	村上大輔		竹内努
助教	服部公平		塩田さやか
	白崎正人		本橋永至
	奥野彰文		
	Le Thanh Tam		

教員19名，特任教員/研究員4名，客員教員12名，支援員1名

● スライド 3

統計的機械学習NOE　他機関との連携状況

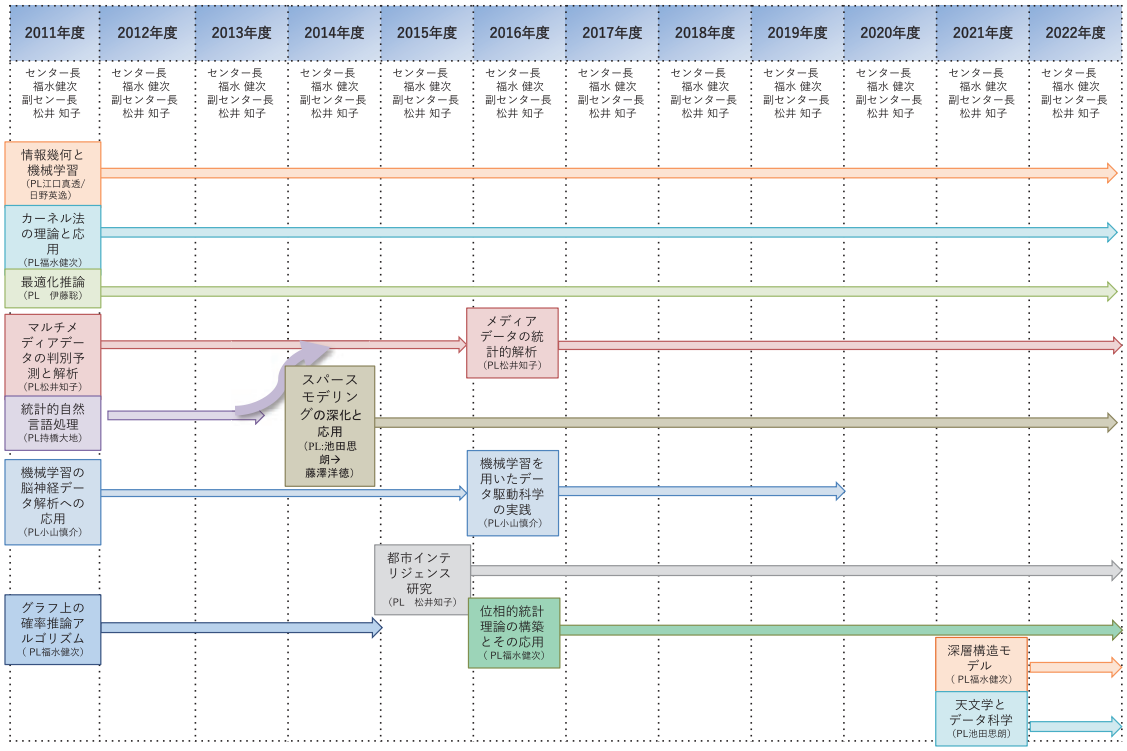


※2023年3月現在
計２０機関と協定締結・連携中

締結年月	協定機関名
2010年8月	Max Planck Institute for Biological Cybernetics（独）
2012年1月	ノルウェー産業科学技術研究所（SINTEF）
2012年2月	University College London, CSML(英)
2012年5月	ノルウェー科技大（NTNU）電気工学通信学部
2013年5月	東北大学原子分子材料科学高等研究機構
2014年1月	青山学院大学
2014年2月	トヨタ工業大学シカゴ校（米）
2014年2月	会津大学
2015年2月	University College London, Big Data Institute（英）
2015年2月	リスク研究所 ETH チューリッヒ
2016年1月	東京大学国際高等研究所カブリ数物連携宇宙研究機構
2016年3月	ボルト大学（ボルト大学）
2016年4月	九州大学・マス・フォア・インダストリ研究所
2016年6月	Zuse Institute Berlin （独）
2016年10月	京都大学情報学研究科
2019年1月	Jean Golding Institute, University of Bristol （英）
2019年3月	Universite´ Bretagne Sud（仏）
2019年3月	九州工業大学大学院生命体工学研究科
2020年8月	EURECOM（仏）
2023年3月	Laboratoire de Mathématiques Blaise Pascal, University of Clermont Auvergne （仏）

● スライド 4

統計的機械学習NOE誕生からの変遷・プロジェクトの変遷



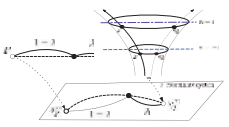
● スライド 5

現プロジェクトの研究活動紹介 1

理論基盤と方法論

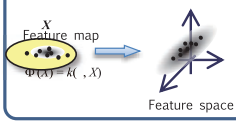
情報幾何と機械学習

確率モデルを幾何的対象として扱う「情報幾何」を用いて、機械学習で用いられる高度な学習アルゴリズムの統計的な性質を解明する.



カーネル法の理論と応用

カーネル法による確率分布の表現を用いた新しいノンパラメトリック推論の方法を研究. 検定やベイズ推論への応用.

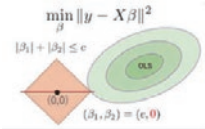


最適化推論Project

統計的機械学習の各領域を横断的に支えるための、大規模数値計算による新たな推論技術の開発。
理研, 九大IMI, NUS, ZIBらと共催の国際シンポジウムを毎年開催

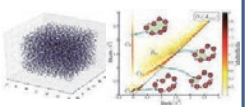
スパースモデリングの深化と応用プロジェクト

圧縮センシングやLASSOなどのスパースモデリングの理論と方法の研究. また, それをざまざまな実社会データに応用する.



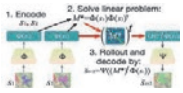
位相的統計理論の構築とその応用

位相的データ解析に関する統計的理論の構築を目指すとともに, 統計科学への位相的方法の導入を研究.



深層構造モデルプロジェクト

深層構造を持つモデルに対し, 関数解析, 代数構造, 微分方程式などの数理的アプローチによる理論的解明, モデルの構築, およびその実世界応用を目指す



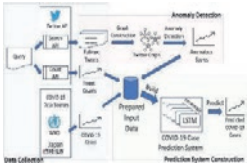
● スライド 6

現プロジェクトの研究活動紹介２

方法と応用

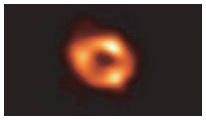
マルチメディアデータの判別予測と解析

音声・音楽、映像、テキストなどのマルチメディアデータから、判別予測の目的に応じて、有用な情報を発見するための研究開発。



天文学とデータ科学プロジェクト

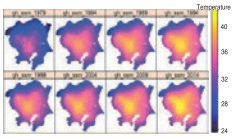
天文データへデータ科学の方法を適用し、共同研究論文を発表し、天文学への貢献を行う。また、研究会の企画、発表を通して天文分野へデータ科学の啓蒙も行う。



クレジット: EHT Collaboration

都市インテリジェンス研究プロジェクト

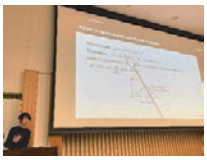
都市レジリエンス向上を目標として、環境・エネルギーや農業の状況解析からリスク管理、セキュリティ統合、都市レジリエンスボンド設計までを俯瞰的に行うための、統計数理／機械学習に基づく技術とその理論を研究開発。



● スライド７

ワークショップ・研究会など開催実績

【開催実績（2020年度以降）】

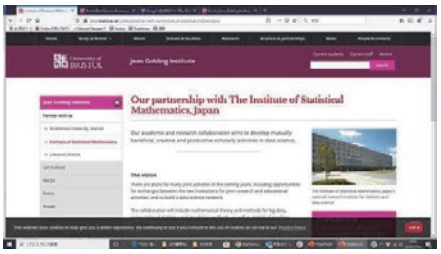


- 2020/9/3-10/8（計6回）JGI(Bristol)-ISM Joint Seminar Series （Online）
- 2021/4/21 AIMR-IFS-ISM合同ワークショップ（Online）
- 2021/11/05 統計数理研究所新型コロナウイルス対応プロジェクト
  主催公開シンポジウム「新型コロナウイルス関連データを解析する」（Online）
- 2021/9/27-30 5th ZIB-RIKEN-IMI-ISM MODAL Workshop on Optimization, Data Analysis and HPC in AI（ドイツ，ハイブリッド）
- 2022/9/16-22 6th RIKEN-IMI-ISM-NUS-ZIB-MODAL-NHR Workshop on Advances in Classical and Quantum Algorithms for Optimization and Machine Learning（対面）
- 2023/3/14-16 Workshop on Functional Inference and Machine Intelligence（ハイブリッド）
- 2022/10/3-5「天文学におけるデータ科学的方法」
- 2023/3/17 公開シンポジウム「COVID-19とデータ科学」（リスク解析戦略研究センターと共催）
- 統計的機械学習セミナー（計9回開催，オンラインからハイブリッドへ）
- 2022年度より「UEC-NAOJ-ISMアストロセミナー」オンラインセミナーを8回開催（池田）
- 2023年度から「統計物理と統計科学のセミナー」（坂田，伊庭）の共催

● スライド８

海外連携機関との協働

- 例）University of Bristol, Jean Golding Instituteとの研究連携
  - JGI： Bristol大のDSに関する中心組織.
  - 2019.1月 MOU締結
    Song Liu （元統数研特任助教）中心.
    2019.5.30-31 Bristol大で第1回共同WS
  - 2020 第2回WS（統数研開催予定）は中止
    → JGI(Bristol)-ISM Joint Seminar Series
    （オンライン．計6回）



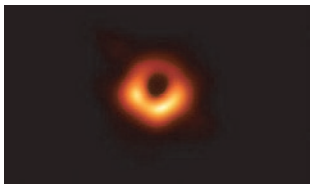
  - 2022 FIMI Workshop: U Bristol, UCL の学生によるオンラインポスターセッション
  - 2023 FIMI Workshop: Song Liu と3名の学生来訪.
  - 2024 FIMI Workshop: Bristol で開催予定
- Network的連携へ
  - RIKEN-IMI-ISM-NUS-ZIB-MODAL-NHR Workshop
  - Bristol Univ., Univ College London, EURECOM 等を中心としたFIMI WS
    （FIMI 2020は仏EURECOMで開催）
- 学生を含めた共同研究
  - U Adelaide (Sejdinovic), UCL (Gretton), Penn State U (Sriperumbudur)ら

● スライド 9

国内コミュニティへの貢献

- 国立天文台，天文学コミュニティとの連携
  - 国立天文台が採用し，最初の 5 年間統数研を勤務地とするテニュアトラック助教
    服部，白崎，助教
  - 池田教授をコアとした天文学・データ科学の協働
    - Event Horizon Telescopeによるブラックホール世界初の撮像
    - 銀河中心のブラックホール撮像

    ほか多数のプロジェクト
  - UEC-NAOJ-ISMアストロセミナー（月1回程度）



Credit: EHT Collaboration

- 統計物理と統計科学
  坂田准教授，伊庭教授主催の連続セミナー（ほぼ毎週）
- 機械学習コミュニティ
  - FIMI Workshop,
  - 統計的機械学習セミナー

● スライド 10

活動特筆事項等(2020年度以降)

【大型外部資金獲得等】

科研費・学術変革領域（A）「データ記述科学」（2022-2026 計画研究代表：福水）
JST CREST 「位相的データ解析」（2015-2021 主たる共同研究者：福水）
JST CREST 「数理知能表現による深層構造学習モデルの革新」（2020-2025 代表：福水）
JST CREST 「計測と情報の融合」（2017-2022 主たる共同研究者：日野）
NEDO 「確率深層コンピューティング」（2018-2022 福水）
NEDO「人工知能技術の導入加速化技術」（2018-2023 日野）
JST さきがけ「求解軌道のマクロ表現によるアルゴリズム制御理論の創出」（2019-2022 坂田）
JST さきがけ「解釈可能AIによるパターンダイナミクスの数理構造抽出と材料情報学への応用 」
（2021-2024 本武）

【プレスリリース・その他】

2022/06 新しい高精度シミュレーションが明らかにした星団形成の現場（服部）
2022/06 深層学習が優位性を発揮する特異データ空間の存在を証明（福水）
2022/05 天の川銀河中心のブラックホールの撮影に初めて成功（池田）
2020/05 統計物理のレプリカ対称性が解き明かす表現型進化の拘束と方向性
～進化における確率的表現型発現の重要性～（坂田）
2021/07 埋もれた暗黒物質の地図を掘り起こす
―観測・シミュレーション・人工知能のタッグで描くクリアな宇宙―（白崎，池田）
2020/12 工場の現場技術者の知見を反映・学習する不良原因解析AIを開発
～半導体工場等における解析結果の精査時間を大幅に削減～（藤澤）
2021/02 スーパーコンピュータで時間を戻して探る宇宙の始まり（白崎）

【受賞等】

2022.11　電子情報通信学会IBISML研究会賞　（有竹特任助教、日野教授）
2022.10　ICDIS 2022 で Best Application Paper Award（南和宏教授）
2022.9　日本OR学会第 12 回研究賞奨励賞（田中未来准教授）
2022.9　2022年度統計関連学会連合大会「優秀賞」（豊田祥史, 福水研究室）

● スライド 11

今後の計画

■ 共同研究・連携ネットワークの更なる発展
対面/ハイブリッドによる国内・国際的な人的交流，共同研究体制の展開

■ 体制の変更（2023年度中）
NOE型研究センター（兼務）から
先端データサイエンス研究系（本務）への移行

- 研究系内の**高等研究センター**（教員6-7名程度）：
機械学習コア分野の先端的な研究プロジェクト推進
- **バーチャルラボ**（共同研究拠点）：NOE活動の発展
  - 連携機関をコアとして，国内外・産学官のトップエッジ研究者が合流する仮想的な研究室
  - 共同研究のハブ：　統数研が保有する計算機設備、データ、滞在施設等を利用し統計的機械学習の先導的な共同利用研究を実施
  - 分野・組織・国境の垣根を超えた研究者らが学融合・新分野開拓・オープンイノベーションを推進

● スライド 12

● 統計的機械学習研究センター・NOE 活動紹介・事業計画について

【司会：山下副所長／NOE 推進室長】それでは、個別 NOE の紹介をさせていただきます。まず初めに統計的機械学習研究センターおよび NOE の活動紹介ということで、センター長の福水教授からご説明します。資料 4 をご覧ください。福水センター長、よろしくお願いします。

【福水教授】それでは、統計的機械学習 NOE の活動紹介を、統計的機械学習研究センター長の福水からさせていただきます。今まで既に所長からも、NOE 活動、個別の統計的機械学習研究センターに関してご紹介がありましたが、最近の活動をメインにして、もう少し詳しいご説明を紹介させていただきたいと思います。

まずわれわれの NOE 型研究センターですが、2012 年に発足して、もう 11 年ぐらいたっていますが、活動としては、一貫してこの 4 つの柱を掲げています。1 つは、独自の研究プロジェクトを推進していまして、NOE 活動ではありますが、まずわれわれとしての研究のコア、ハブをつくるための活動を行っています。もう一つは、NOE 的な活動としまして、諸機関との協定の締結と、それに基づく共同研究の推進を進めてきています。これに関しては、詳細を後でお話しします。それから、特にわれわれの機械学習 NOE は、他の NOE と比較した時に、目立った点として、国際連携に非常に力を入れています。後で話しますけれども、国際的な研究機関の多数と連携を結んでいて、それらの研究機関と共同研究、あるいは国際ワークショップの開催を進めています。もう一つは、機械学習コミュニティ、国内外のコミュニティへの貢献としまして、研究会やセミナー等を主催、共催しまして、共同利用機関としての価値の向上を図ろうとしています。

これが、現在の人員配置ですけれども、現在、教員は 19 名が所属しています。少し大所帯になっています。特任研究員、特任助教が 4 名いまして、客員 12 名という体制で研究開発を行っています。

先ほども少しご紹介しました他機関との連携です。MOU 等を結んだ他機関との協定締結ですが、現在、20 機関と協定を締結しています。そのうち、海外との協定が 13 機関ありまして、どちらかというと、国内よりも海外の機関と多くの協定を結んで、連携を深めているという形になっています。

これがプロジェクトの変遷でして、ここでは詳細をお話ししませんけれども、幾つかのプロジェクトに関しては既に終了しました。また、2 年ぐらい前に、新たなプロジェクト等も幾つか発足しているという形で、現在、9 つのプロジェクトを推進しています。

今日は、詳細をご紹介することはできませんけれども、簡単に、どのような研究をしているかに関して、プロジェクトのタイトルだけをご紹介します。9 つプロジェクトがあるのですが、6 つのプロジェクトに関しては、比較的、基盤的、あるいは方法論的なテーマを扱っています。1 つは、情報幾何という日本が研究として優れている、甘利俊一先生が構築された情報幾何の流れをくむ研究を行っているプロジェクトです。それから、私がテーマリーダ

ーを行っている、機械学習の 1 つの手法であるカーネル法を発展させていく研究プロジェクトです。それから、機械学習の基盤をなす数値最適化に関するプロジェクトです。それから、これも機械学習の 1 つの重要なテーマであります、スパースモデリングを扱う研究プロジェクトです。それから、数学的な概念であります、位相的な情報を取り込む手法をデータ解析に応用した位相的統計理論の構築とその応用と呼ばれています、位相的データ解析分野を研究するプロジェクトです。それから、これは深層学習に関わるプロジェクトで、私がテーマリーダーをしていますけれども、CREST の予算にリンクしまして、深層構造モデルプロジェクトというものを行っています。

比較的応用に近い分野としましては、1 つは、マルチメディアデータの判別予測と解析というテーマでして、マルチメディアデータからの推論を研究するプロジェクトです。それから、先ほど少しご紹介もありましたけれども、池田教授が中心になって、天文学とデータ科学のプロジェクトというものを非常に活発に進めていまして、このプロジェクトが 1 つ立っています。それから、最後に都市インテリジェンス研究プロジェクトといいまして、都市に関わるデータ解析、あるいは都市レジリエンス向上を目指した環境・エネルギー、あるいは農業も含めたような、非常に広い分野を扱うプロジェクトを 1 つ進めています。

NOE 活動の一環としまして、われわれは、ワークショップ・研究会等を共催、主催していまして、特にこれは 2020 年度以降の開催概要をまとめたものです。残念ながら、2020 年度以降というのは、コロナ感染がまん延した時期と重なっていまして、特に海外との提携を主眼に置いているわれわれの活動としては、極めて困難な状況にはありました。しかし、オンラインツール等を利用しまして、国際的な活動、ワークショップ・研究会というものも絶やさずに行っていこうという意識を持って、開催をしてきました。最初、2020 年度、21 年度というのは、主にオンラインで開催をしてきましたけれども、最近は、ハイブリッド、対面の開催が徐々に増えてきています。

その間の新しい試みとしまして、これは後でも少しお話ししますけれども、天文学分野のオンラインセミナーで、電通大、天文台、統数研の 3 機関共同で主催しているものが昨年度から開始されています。それから、今年度からですが、統計物理分野の研究者である坂田先生、伊庭先生が、統計物理と統計科学のセミナーを、高頻度で行うような活動を、新しい活動として行っています。

NOE 活動、外部機関との共同研究の活動内容に関して、幾つかの例をご紹介したいと思います。まず 1 つは、海外との連携の例をご紹介したいと思います。密接に海外の機関との共同研究を行っている例としまして、University of Bristol の Jean Golding Institute との研究連携というものが挙げられます。Jean Golding Institute というのは、Bristol 大学のデータサイエンスに関する、全学をまとめた中心組織でして、ここと、2019 年に MOU を締結しました。これは、実は、Bristol 大学の Department of Math.に統数研の元特任助教をしていた Song Liu という研究者が Lecturer として働いていまして、彼が、相手側の中心的な人物として MOU を結んでさまざまな連携活動を行っています。2019 年、Bristol 大

で、第 1 回の共同ワークショップを開催しまして、その後も共同ワークショップを開催するという計画でいました。しかし、残念ながら、コロナの状況がありまして、2022 年の第 2 回ワークショップは、統数研開催予定だったものが中止になりました。その代わりに、Bristol 大学と統数研とのジョイントセミナーというものを、オンラインで計 6 回、2 人または 1 人ずつスピーカーを立てて、企画・開催しました。

これは私が中心になっている国際ワークショップなのですけれども、2022 年、まだオンラインでの開催でしたが、University of Bristol と、やはり MOU 締結機関である University College London とのワークショップ時に、学生さんも含めたオンラインポスターセッションというものを開いて、若手も含めた交流というものを図る企画を行いました。大体、毎年 3 月にワークショップを行っているのですけれども、2023 年 3 月に、ハイブリッドで対面も含めた開催が可能になりまして、Bristol の学生 3 名を含めた来訪者がありました。来年度は Bristol 大学でワークショップを開こうと企画しています。

われわれは MOU 等を背景としたネットワークを形成しています。MOU というのは、統数研対相手側という 1 対 1 の関係、あるいは場合によっては、3 者というケースもあるのですけれども、MOU 自身は、1 対 1 の関係で結ばれることが多いです。それを越えて、ネットワーク的な連携というものを、ぜひとも推進したいと考えています。例えば、ワークショップなどの開催を通じて、統数研がハブ的な役割を果たすようなネットワーク形成を積極的に行おうとしています。ここに 2 つ例がありますが、この 1 つ目の、多くの機関名がついている、RIKEN-IMI-ISM-NUS-ZIB-MODAL-NHR Workshop というのは、5 機関、理研と九大 IMI、それから、統数研、National University of Singapore、Zuse Institute Berlin というドイツにある最適化数値計算の研究機関、MODAL-NHR という ZIB の中のセンター、多機関共同でのワークショップを毎年開催していまして、いろいろな場所でワークショップを順番に行っているような活動が行われています。

それから、これは Bristol 大学も含めた活動ですけれども、私が中心になっているワークショップですが、Bristol 大に限らず、University College London、それから、フランスにあります EURECOM、これも元々統数研で特任教員をやっていた方が、現在、ここで研究者をしているので、そのような方をコアにしてネットワーク的な活動として国際ワークショップを開いて、統数研が、そのハブになっていくということを構想しています。

もう一つは、若手研究者の育成という観点に立って、海外との共同研究においても、学生を含めたような共同研究というものをプランニングすることを考えています。実際に、University of Adelaide や University college London に長く共同研究をしている研究者がいます。Penn State University などとは、相手先の学生がこちらに来たり、こちらの学生を向こうに派遣したりという形で、学生を含めた共同研究というものを推進して、論文につながるような成果を上げていくという活動を行っています。

これは、国際的なものも含みますが、どちらかというと国内コミュニティとの連携です。先ほどブラックホール等の話で、天文台との連携ということをお話が、私もしましたし、山

下副所長のほうからもありました。天文のコミュニティとの連携というのは、通常の連携を超えて、極めて密接な連携が図られています。1 つは、国立天文台が採用した助教の方を、最初の 5 年間、統数研を勤務地として研究活動を行うというプログラムを、国立天文台と統数研が、双方で合意の下で行っています。実際に 2 名の助教の方が、天文台のテニュアトラック助教であるが、統数研でも助教としてアサインされて、統数研で勤務を行っているという活動を行っています。

特にこの天文学とのコミュニティとの共同研究においては、統数研の池田教授がコアになり、このような連携の国内での中心的な人物となって、天文学とデータ科学の協働が行われています。これはよくご覧になる図かと思いますけれども、このブラックホール直接撮像においては、池田教授がスパースモデリングの使用法において、本質的な貢献を行っていると聞いています。それから、このような活動の派生的な活動としまして、先ほどご紹介したセミナーや研究会等を活発に行っていらっしゃいます。

天文学以外にも、先ほどお話しした統計物理との連続セミナーや、また、国内コミュニティという点では、統計的機械学習セミナーというものを不定期に行っています。現在、ハイブリッドで、このようなセミナー、あるいは、ワークショップというものを行っています。

特筆事項としまして、外部資金の獲得状況、それから、プレスリリース、受賞等に関してまとめていますけれども、統計的機械学習研究センター、あるいは NOE 活動においては、大型の外部資金の獲得が非常に多く見られます。この 2 年間ぐらいのものをまとめてありますが、現在、科研費の学術領域の計画研究代表を私が務めています。これは終了しましたけれども、CREST の主たる共同研究者で、それから、これも私ですけれども、CREST の代表者を現在行っています。また、昨年度終了していますけれども、CREST の主たる共同研究者を日野先生が行っていました。NEDO に関しても、1 件は終了していますけれども、1 件は今年度まで継続しているものがあり、合計 2 件、この 2 年間で活動があります。若手の研究者の方の JST さきがけも、1 つは終了、1 名は転出してしまいましたけれども、2 件の獲得状況があります。プレスリリースや受賞等もここに書かれていますような実績があります。

所長のほうから最初にお話がありましたように、統計数理研究所は、今年度、改組をしまして、統計的機械学習研究センターも、その組織変更で大きく変わることになります。そのような改組によって、共同研究や連携のネットワークのさらなる発展を図りたいと考えています。

大きな変更点は、先ほどもご説明がありましたけれども、現在の NOE 型研究センターというのは、正規の教員が全員兼務の組織ですけれども、兼務から本務、先端データサイエンス研究系の下のセンターとして本務所属になるということです。ただ、先ほど言いましたように NOE のセンターには、現在、教員が 19 名いまして、全員が新研究系に所属するとはいきません。人数バランスがかなり偏ってしまいますので、現状では、研究系の中のセンターに所属する教員というのは、6〜7 名の計画です。概算要求の獲得状況にもよりますけれ

ども、このような人数が想定されています。もう一つは、先ほどご説明がありましたけれども、バーチャルラボと呼ばれている、こちらが、NOE 活動の拠点になるべき、外部との共同研究の拠点になるものだと考えています。

このような 2 層の体制によって、今後の統計的機械学習の NOE 活動というものを発展させていきたいと考えています。高等研究センターのほうは、研究系ですので、どちらかというと、研究プロジェクトの推進、機械学習の特にコアな部分に関する研究プロジェクトの推進を考えていまして、バーチャルラボのほうはより連携的な研究です。

コロナの状況において 1 つのいい点としては、オンラインでの共同研究、ミーティングで、人を集めるのが非常に楽になったという点があると思います。そのようなメリットを生かして、海外、あるいは国内の遠隔地に住んでいらっしゃる方も含めて、簡単に人が集まることによって共同研究、バーチャルラボというものを運営していきたいと考えています。機械学習のコアな部分だけではなくて、マテリアルズインフォマティクスを含めた、いろいろな応用分野、展開できる分野の人々を巻き込んだバーチャルラボというものを運営して、共同研究、連携活動というものを進めていきたいと考えています。私のほうからは以上です。ありがとうございます。

【司会】ありがとうございました。ただ今の福水センター長からの説明に関して、何かコメント、質問等ありましたら、よろしくお願いします。

【上田顧問】理研の上田です。

【司会】よろしくお願いします。

【上田顧問】ご説明ありがとうございました。学術的には、非常に顕著な成果が出ていまして、かつ大型研究も採択されて、非常に素晴らしいと思います。一方、今後の展開として、バーチャルラボの話がありました。先ほど IT システムというコメントもあったと思うのですけれども、これも、釈迦に説法だと思いますけれども、最近、大規模生成モデルが注目され、サイエンスの領域にもかなり影響を及ぼしつつあるというのは事実です。科学における大規模モデルというものを、やはりこのような NOE の中では、非常に重要なテーマではないのかと思います。少しその辺り、どのようなご見解なのか、恐らくこのバーチャルラボというところが、そのようなところを担うのではないかと思います。

バーチャルラボでの展開がないと、福水先生のところで考案された、いろいろとがった学術的な技術が、論文で止まってしまうというのが、心配されるわけです。それ自身大きな成果なのですけれども、やはり NOE の本質として、その辺りを上位レイヤーにどのように展開して、統合していくのかという観点では、大規模生成モデルの構築といいますか、それをサイエンスに融合していくというインフラをどのように構築するのかについて、何か見識があれば、お願いします。

【福水教授】ありがとうございます。あまり研究の中身に関してお話をしなかったので、その面の構想というのは、ご説明できていなかったのですけれども、おっしゃるように、深層

学習をはじめとする人工知能というのは、とても発展をしてきて、基礎研究のみならず、社会への展開というものが非常に重要になっていることは、もちろん認識しています。それに対して、コントリビューションしたいという思いは強く持っています。

大規模言語モデルは、現状、統数研でやるというのは難しいと考えています。1つは、データがないということもありますし、自然言語分野には優秀な研究者が統数研にも在籍していますけれども、1人でできるようなプロジェクトではないわけです。大規模言語モデルに関しては、少し統数研を超えたレベルでの何らかの動きというものが必要な分野だと思います。私が個人的にこれから進めたいと考えていますのは、もう少しサイエンティフィックな分野、数値計算や、シミュレーション、微分方程式に関わるような分野です。

これは、統数研がデータ同化等も含めて、非常に得意な分野であります。そのような、例えば自然現象の記述や、自然現象に限らずコンピュータグラフィックスや、工業分野など、シミュレーションを使う分野というのは、生産技術も含めて非常にたくさんあるわけです。そのような分野の基礎技術、人工知能的な技術を開発することによって、社会、あるいは産業、そのような分野へのコントリビューションをしていきたいと考えています。

先ほど上田先生からお話があったように、そのためにバーチャルラボというのは、さまざまな分野、われわれが今、統数研で持っている強い分野とは異なった分野の人たちを集めることができると考えています。例えば、数値計算、もっといえば、大型の計算機、あるいは、シミュレーション技術、そのようなものに明るい人たちを含めて、研究を発展させていきたいと考えています。

【上田顧問】ありがとうございます。誤解があってはならないので補足です。私自身も、統数研で、大規模生成モデルそのものを構築すべきという意見ではありません。現在のChatGPTにしても、トランスフォーマーなど、ベースに核技術があります。それにとって代わるような核技術の創出に期待しています。後半のPDEの話でも、ご承知かと思いますけれども、Physics-informed neural networksなどは、データなしにいわゆる深層学習で、原理的には任意の偏微分方程式を解けます。ただ、まだ学習の面でプアなところがたくさんあるので、その辺を、このようなところで開発していくと、いわゆる物理シミュレーションと機械学習、深層学習の融合という意味での発展にも非常に貢献できるので、そのようなことを意識した技術のコアをつくっていただければいいかなという意識です。決して大規模生成モデルを構築して欲しいということではありません。ただ、創出した核技術をバーチャルラボでどのように展開するのかについてはご検討お願いします。

【福水教授】それは認識しています。ありがとうございます。まさにおっしゃったことは、私もぜひやりたいことですので、いろいろなチャンネルで、またご説明する機会があるかと思います。

【上田顧問】ありがとうございます。以上です。

【司会】どうもありがとうございます。他、いかがでしょうか。

【蒲地顧問】よろしいですか。

【司会】蒲地先生、お願いします。

【蒲地顧問】2点ありまして、1つは、意見というか、感想のようなもので、2番目は、質問です。1番目は、先ほどの山下副所長の発表の中の私の質問で、NOEだけれども、あまりネットワーク的なことがよく見えないということを申し上げたのですが、今のご発表で、非常によく分かりました。ネットワーク的連携で、海外の連携機関との共同ということで、ワークショップをやられて、その後、共同で始まったテーマがあったりするということもあるし、国内でのコミュニティへの貢献というか、それは、もう具体的にハブとして役立っているということが、よく分かりました。それをもっと進めるという意味で、今後、バーチャルラボがネットワークの中心になっていくと考えればいいのだなという感じでお聞きしました。ありがとうございました。

2番目は質問なのですけれども、この社会への展開という機械学習、深層学習の社会への展開という時に、これは、先ほどの副島先生のご質問にも関係はあるかと思います。私が関係しているオーシャンアイズという会社は、ベンチャーでして、京都大学、滋賀大学の深層学習のグループと、それから、海洋研究開発機構の海洋データ同化やモデルのグループが合わさって、海の状況を再現して、それと衛星データなどを組み合わせて、水産への応用をしようと。水産で、いつどこに行ったら、どのような魚が取れるということを、早く言えば、機械学習で出していこうということをやっています。その時に一番問題になるのはデータの量で、機械学習を行われている中で、例えば、実験系など、そのようなところでは非常にデータはよく取れます。

だから、それでビッグデータの解析としてなるのですけれども、水産の場合われわれがはたと困るのは、漁業者に聞き取りをして、それをデータとしてもらうとなると、まず彼らは、IT関係は全く関係ないというか、まだ紙に書いて、自分で好きなように書いて、フォーマットも一定していないなど、そのようなことから始まります。ですから、非常になかなか大変で、ビッグデータにならないような解析を、実は、われわれは、やっています。あるいは、それらの紙に書かれたものを、人がデジタルデータに起こして、それを解析するという、非常に手間が掛かっています。そのような中で、このようなネットワーク、NOEのような活動、あるいはバーチャルラボのような活動で何かうまくいくようなことはどのようなものか、質問としては、データとしては、非常に取りにくいデータを扱わなければならない時に、何か示唆になるようなご経験など、そのようなものがありましたら、教えていただきたいなと思います。

【福水教授】ありがとうございます。大変難しい課題だと思います。われわれ機械学習の研究者は、どちらかというと、データが集めやすい分野を中心に研究しています。ところが、例えば、産業分野、工場などもそうですけれども、現場でデータを集めることが、それほどたやすいことではないというのは、重々認識しています。もちろんそこにIT的な、例えば、画像を使うなど、そのような解決策を持ち込むというのも、1つのアプローチではあるかと思いますけれども、ただそれだけで解決する部分ばかりではありません。ですから、データ

をどのように集めるかというところは、1つ大きな課題ではあるし、もっと言うと、これから将来性のある分野ではあると認識しています。これからやることが、まだあるということです。ネットワーク的な活動で、それが解決されるかというと、すぐにいいアイデアがあるわけではないのですけれども、いろいろな研究者、あるいは、例えば、現場をよく知っている研究者など、そのような方たちとのネットワークをつないでいくというのも、1つの重要な活動だと思います。そのようなさまざまな観点での研究者が集められる場というものを提供していきたいと考えています。ありがとうございました。

【蒲地顧問】分かりました。ありがとうございます。

【司会】どうもありがとうございます。かなり時間がオーバーしていまして、長引いているので、休憩は入れたいと考えます。少し短いですが、35 分に再開ということでよろしいでしょうか。

それでは、3 時 35 分に再開ということで、いったん休憩させていただきます。どうもご議論、ありがとうございました。またよろしくお願いします。

（休憩５分）

Data Science Center for Creative Design and Manufacturing

第4回 統計数理研究所 NOE形成事業 顧問会議

ものづくりデータ科学研究センター

NOE活動紹介・事業計画

ものづくりデータ科学研究センター長
教授　吉田　亮

資料 7

大学共同利用機関法人 情報・システム研究機構
統計数理研究所

ものづくりデータ科学研究センター

NOE活動紹介・事業計画について

吉田 亮　ものづくりデータ科学研究センター・センター長

RadonPy　Python library for polymer properties calculation using fully automated MD simulation
https://github.com/RadonPy/RadonPy

XenonPy　Python library for inverse materials design
https://xenonpy.readthedocs.io/en/latest/

● スライド 1

重点領域：マテリアルズインフォマティクス

(1) 2015年にJSTイノベーションハブ構築支援事業「情報統合型物質・材料開発イニシアティブ」（MI²I：日本のマテリアルズインフォマティクス学術創成のフラッグシップ事業，PL 伊藤聡（当時NIMS，NOE形成事業顧問会議 顧問））が発足し，統数研からは吉田・福水が参画

(2) 2017年7月センター発足（6年弱）

(3) “マテリアルズインフォマティクス” の隆盛期

(4) 社会的ニーズと人的資源の制約から，マテリアルズインフォマティクスを重点領域に設定

● スライド 2

研究体制　(1) 産学連携の戦略的活用; (2) 学際性; (3) 統数研の統計的機械学習の理論・方法論の技術移転

RadonPy

RadonPy コンソーシアム
3大学・25企業 138名

「富岳」成果創出加速プログラム
3大学・24企業 111名

物質・材料研究機構 (MOU)

機械学習・統計科学・第一原理計算・分子動力学・
高分子化学・触媒・固体物理・HPC等

所内教員（7名）
吉田 亮（センター長）
藤澤 洋徳（副センター長）
福水 健次
日野 英逸
中野 慎也
Stephen Wu
林 慶浩

客員教員
塩見 淳一郎（東大 教授）
森川　淳子（東工大 教授）
髙田正彬（東芝 主務研究員）

特任研究員（10名）
Chang Liu
丸山 豊
青木 祐太
篠田 恵子
草場 穫
南 俊匠
髙橋 愛子
野口 瑶
山田 寛尚
木村 薫

共同研究部門

三菱ケミカル株式会社
ISM-MCC フロンティア
材料設計拠点

JSR株式会社
JSR-ISM スマート
ケミストリーラボ

新学術領域「パイパーマテリアル」
田村隆治（東京理科大）

JST-CREST「フェイゾンエンジニアリング」
田村隆治（東京理科大）

JST-CREST「熱物性高分子インフォマティクス」
森川淳子（東工大）

JST-COI-NEXT「再生可能多糖類植物由来プラスチックによる資源循環社会共創拠点」
高橋憲司（金沢大）

企業との共同研究・学術指導
2017-2023 75社

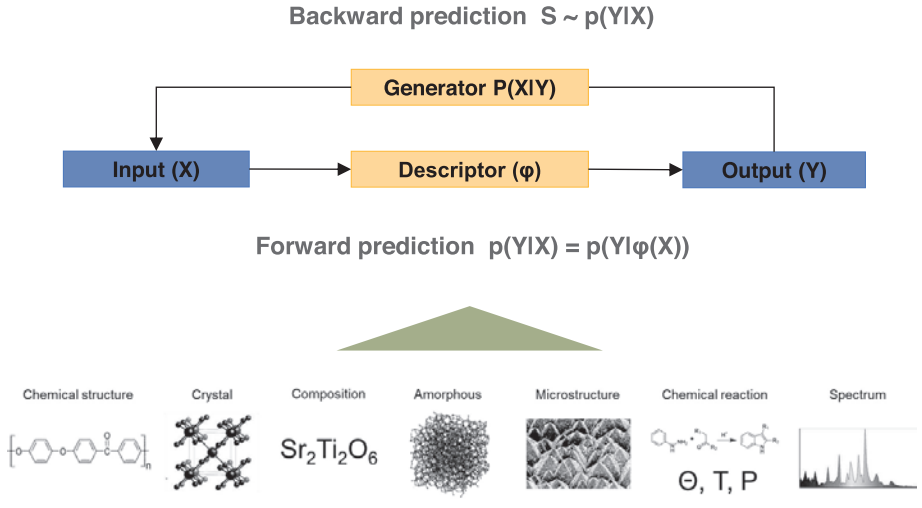
3

● スライド 3

XenonPy

データ駆動型材料研究の順問題・逆問題

Ikebata et al. Bayesian molecular design with a chemical language model. *J Comput Aided Mol Des* **31**, 379-391 (2017).
※ センター発足時に発表した原点となる論文（引用数 118）：言語モデルに基づく分子生成を提案した初の論文．Springer Nature トップ3% 引用数



● スライド 4

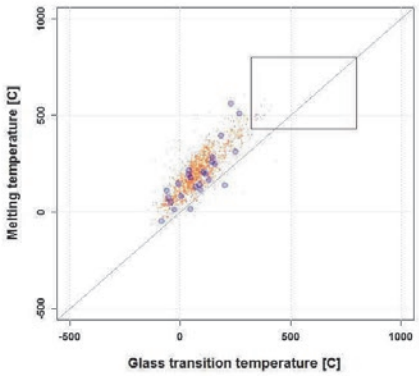
XenonPy

分子生成言語モデルを用いた高熱伝導非晶質高分子の発見

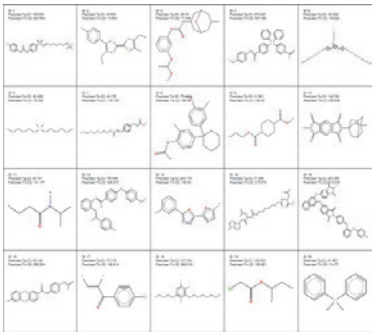
Wu et al. Machine-learning-assisted discovery of polymers with high thermal conductivity using a molecular design algorithm. *npj Comput Mater* **5**, 66 (2019).

※ 生成的機械学習で予測・発見された初の高分子材料（引用数 222, IF = 13.2 (2021-2022), Springer Nature トップ1%論文）

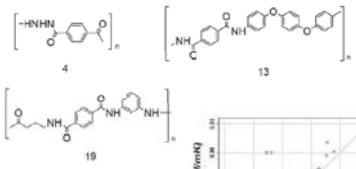
Refinement of polymer properties
Tg↗ and Tm↗



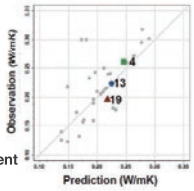
Designed molecules (monomers)
Generative models of chemical structures



Discovery of new polymers
Three thermally conductive polymers

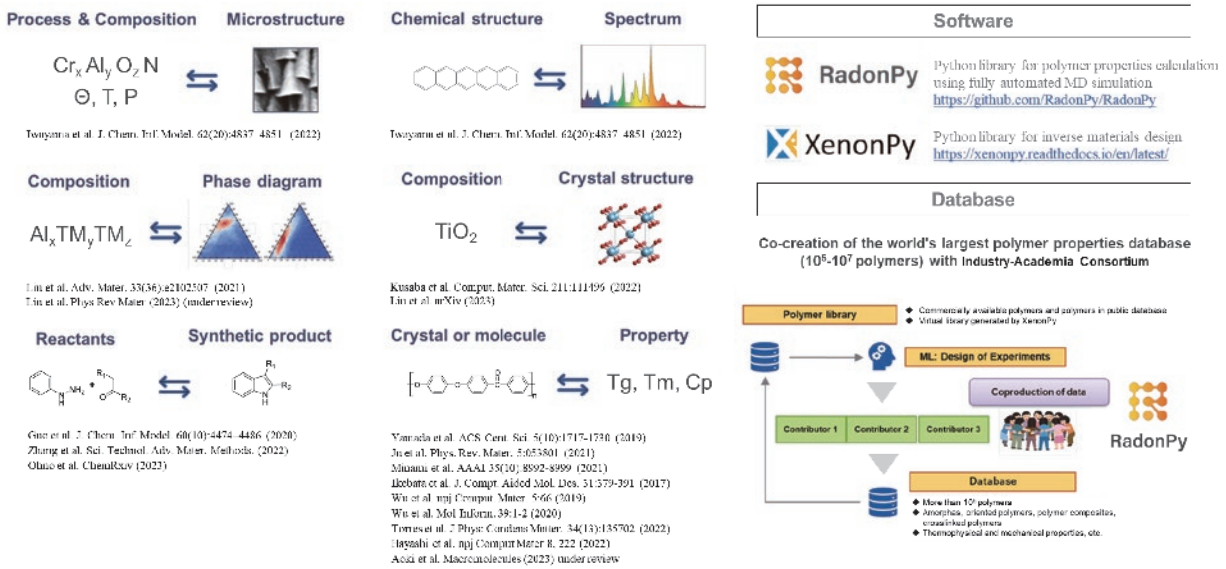


- Filmed polymers
- Crystallinity
- Glassy states
- Soluble in organic solvent
- TC = ~0.408 W/mK



● スライド 5

2017年～ものづくりデータ科学研究センター ～ 基盤技術・学術資源



● スライド 6

データ駆動型材料研究におけるデータ資源の不足

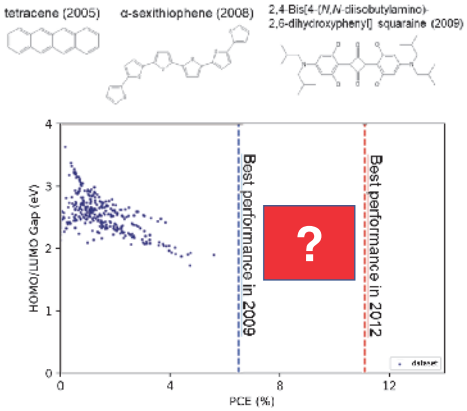
オープンデータの不足：技術的・文化的問題 ～ 短中期的解決は困難

- **PoLyInfo (literature survey)**
  ~100 properties for 18,015 polymers
  NO API
  https://polymer.nims.go.jp/
- **Polymer Genome (DFT)**
  ~10 properties for ~800 polymers
  NO API
  https://www.polymergenome.org/?m=hom
- **CROW (literature survey)**
  NO API
  http://www.polymerdatabase.com/
- **CRIPT** (not yet open)
  MIT + Citrine Informatics
  https://cript.mit.edu/

コスト	実験・シミュレーション・試料作製・物性評価に要するコスト
ニーズ多様性	研究者の興味や設計変数（材料種，試料の作製方法など）が多様 コモンデータを創出しようという動きが起きにくい．
インセンティブ欠如	競合相手に対する情報秘匿の意識 データを公開するインセンティブが研究者に働きにくい．

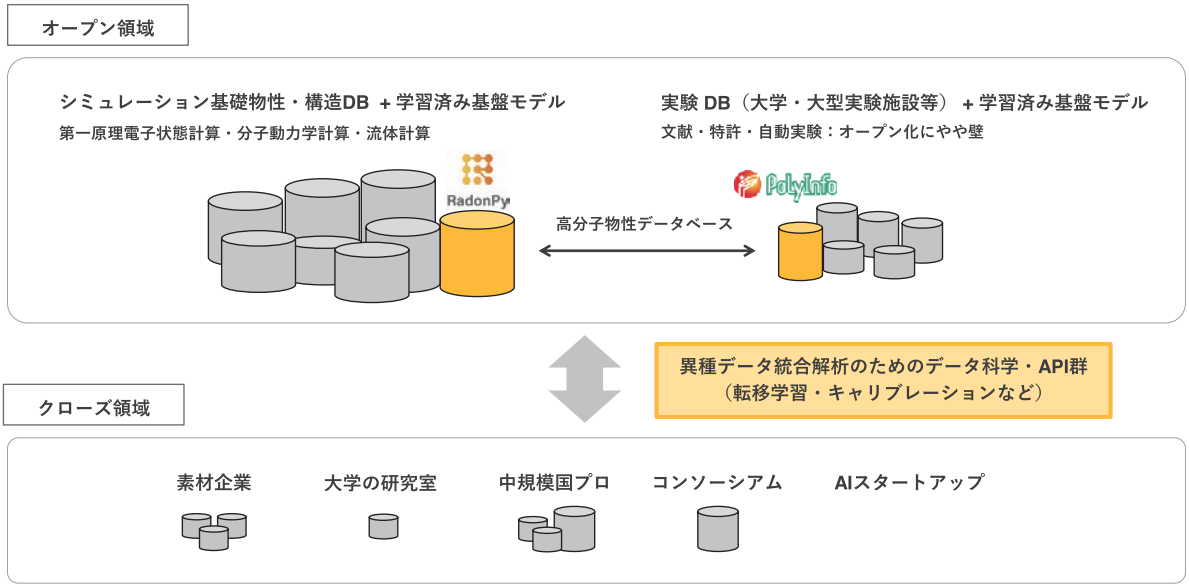
革新的な材料の周辺にはデータが存在しない

データ科学の内挿的予測の限界



● スライド 7

データ駆動型材料研究の在り方とセンターのミッション定義



● スライド 8

林慶浩
（統数研）

RadonPy: 全原子古典分子動力学法による高分子物性計算の全自動化

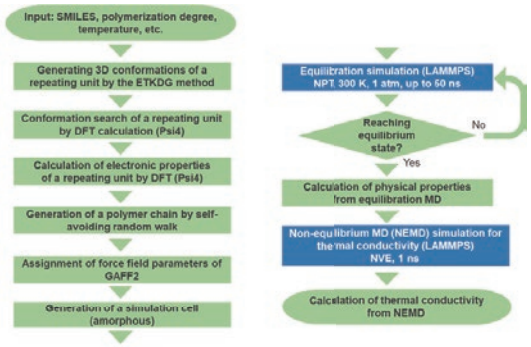
RadonPy

■ **2022年11月 論文発表（全原子分子動力学シミュレーションによる高分子物性計算を全自動化する世界初のソフトウェア）**
Hayashi et al. RadonPy: Automated Physical Property Calculation using All-atom Classical Molecular Dynamics Simulations for Polymer Informatics. npj Comput Mater 8, 222 (2022).引用数 5, IF = 13.2（2021-2022）

■ **ソフトウェア公開** GitHub: https://github.com/RadonPy/RadonPy

■ **プレス発表**
- 国内「全原子古典分子動力学法による高分子物性計算を全自動化する ソフトウェアRadonPyをリリース ～高分子材料物性大地図の作成に向けた第一歩～」https://www.ism.ac.jp/ura/press/ISM2022-08.html
- 海外（EurekAlert!）"Automated simulation software creates a world map of polymer properties" https://www.eurekalert.org/news-releases/975212

■ **メディア報道（参考資料⑩）**
日経新聞，日経産業新聞など その他：ウェブニュース 日本7件，海外7件



Nature Publishing Group の online attention スコアが同時期に出版された全論文中トップ3%にランクイン

This article is in the 96th percentile (ranked 14,317th) of the 444,910 tracked articles of a similar age in all journals and the 99th percentile (ranked 1st) of the 35 tracked articles of a similar age in *npj Computational Materials*

https://www.nature.com/articles/s41524-022-00906-4/metrics

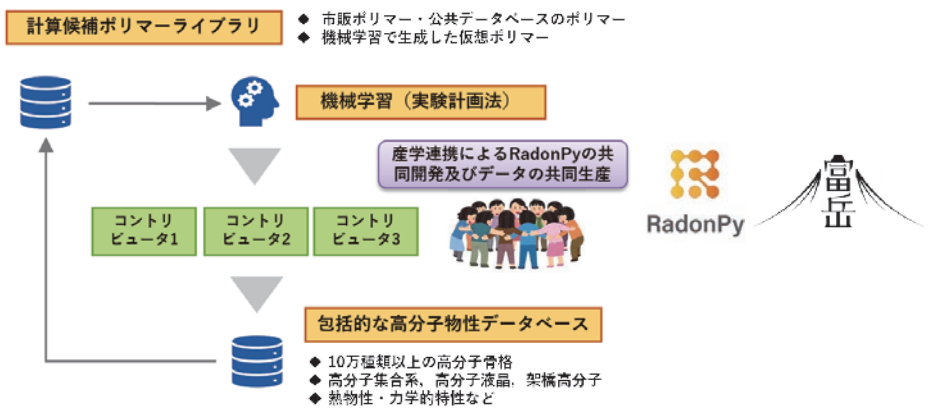
● スライド 9

産学連携による RadonPy 及び 高分子物性オープンデータベースの共創

2021- RadonPy共同開発コンソーシアム@統計数理研究所ものづくりデータ科学研究センター

10^5-10^7ポリマーの包含する世界初の体系的な高分子物性データベースの創出

■ **統数研・3大学・24企業（参画者数：約138名）が組織の垣根を越えてデータを共同生産・共有**
■ **「富岳」成果創出加速プログラム「データ駆動型高分子材料研究を変革するデータ基盤創出」（代表：吉田亮）**



● スライド 10

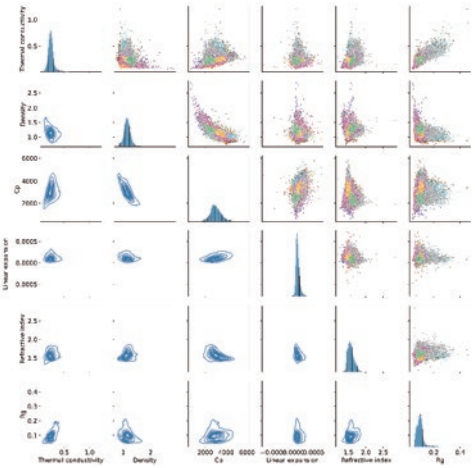
高分子物性の世界地図：広大な物質空間の全貌を明らかに

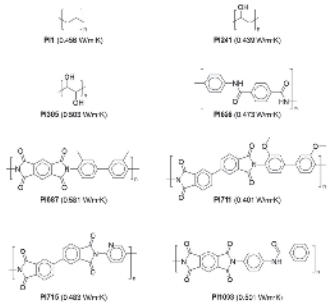
これまでに合成された高分子の分布
（PoLyInfo: 16,427）

計算が完了した47,500ポリマー
(● 未計算仮想ポリマー)

- 複数物性の同時分布の観測：パレートフロンティアの位置や構造的特徴が明らかに
- 特異な物性を発現する高分子の候補を同定
- 実験だけではこのような網羅的観測は不可能

Hayashi et al. npj Comput Mater 8, 222 (2022).





熱伝導率が0.5 W / (m・K)を超えるアモルファスポリマーが存在するという予想

● スライド 11

RadonPyの活用例：Sim2Real 転移学習

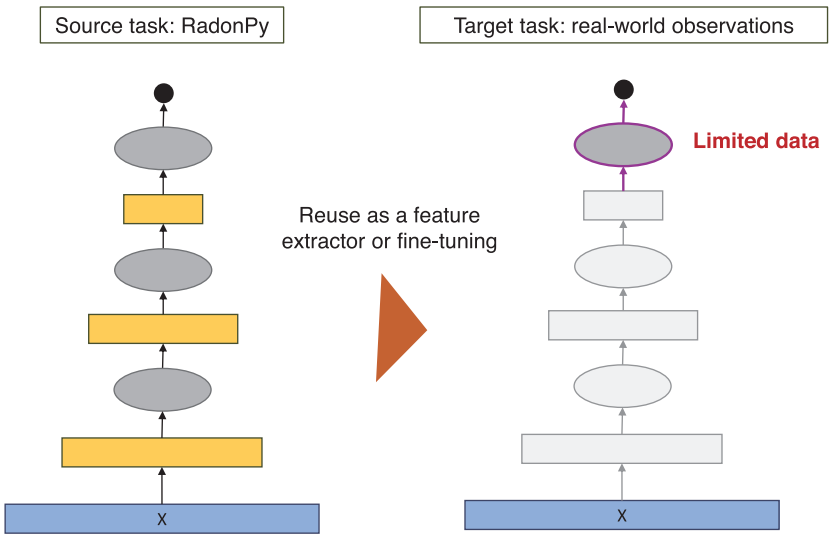
大量のシミュレーションデータと少数の実験データの統合解析（転移学習・ドメイン適応・マルチフィディリティ学習など）

Latest: 17 properties implemented
(2023/03/13)

- Thermal conductivity
- Thermal diffusivity
- Density
- Radius of gyration
- Specific heat capacity Cp
- Specific heat capacity Cv
- Compressibility (isothermal)
- Isentropic compressibility
- Bulk modulus (isothermal)
- Isentropic bulk modulus
- Self-diffusion coefficient
- Thermal expansion coefficient
- Linear expansion coefficient
- Dielectric constant (static)
- Refractive index
- Glass transition temperature
- Abbe's number

Polymer systems

- Amorphous/crystalline states
- Oriented structures
- Mixtures



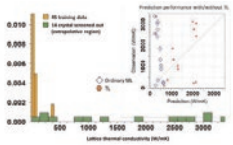
● スライド 12

Sim2Real 転移学習・キャリブレーションの方法と実践

- 転移モデルの外挿性を発見
- 訓練済みモデルデータベース

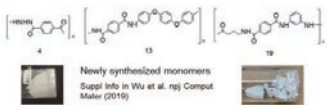
Yamada et al. ***ACS Cent Sci*** 5:1717-1730 (2019)

XenonPy



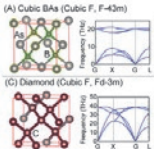
- 高熱伝導ポリマーの合成
- 28点のデータから予測モデル

Wu et al. ***npj Comput Mater*** 5:66 (2019)



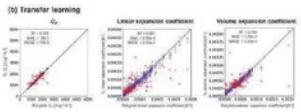
- 超高熱伝導を持つ無機結晶
- 外挿性の実証

Ju et al. ***Phys Rev Mater*** 5, 053801 (2021)



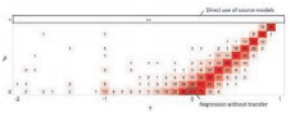
- 高分子物性のSim2Real転移
- 外挿性の実証

Hayashi et al. ***npj Comput Mater*** 8, 222 (2022)



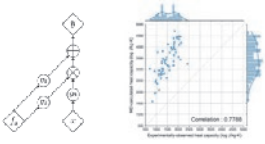
- ベイズ転移学習
- 転移学習の一般クラス

Minami et al. ***AAAI*** 35, 8992-8999 (2021)



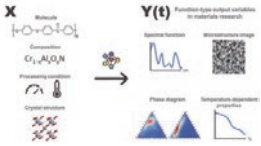
- アフィン転移学習
- L2期待損失最小化転移学習

MInami et al. ***arXiv*** (2022)



- 関数出力変数の回帰
- マルチタスク学習との関係性

Iwayama et al. ***J Chem Inf Model*** 62, 4837–4851 (2022)



- 高分子相溶性の外挿的予測
- 量子化学計算の活用

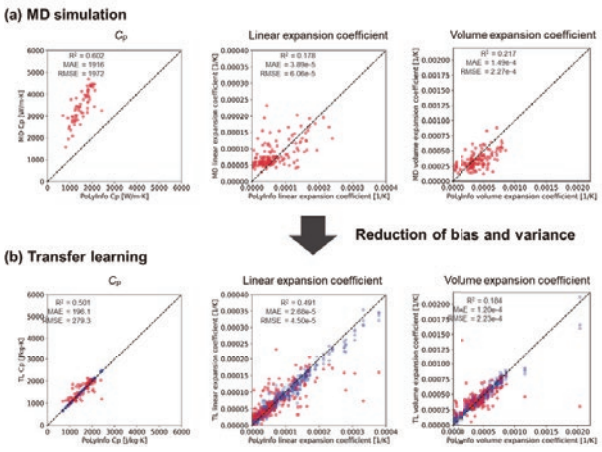
Aoki et al. ***Macromolecules*** (2023)

● スライド 13

不完全な計算と複雑な現実系の乖離を機械学習で埋める

実験値・MD 計算物性の間の乖離を補正

Hayashi et al. npj Comput Mater 8, 222 (2022)

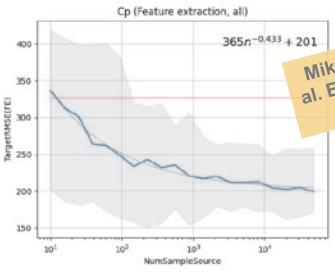


Sim2Real のべき乗則を観測

Predictive performance of a transferred model improves with $O(n^{-\alpha})$ for increasing computational data (n)

$$\text{test error} = Dn^{-\alpha} + C$$

Transfer gap ～ Value of computational database



Mikami & Fukumizu, et al. ECML PKDD (2022) のアイデアを導入

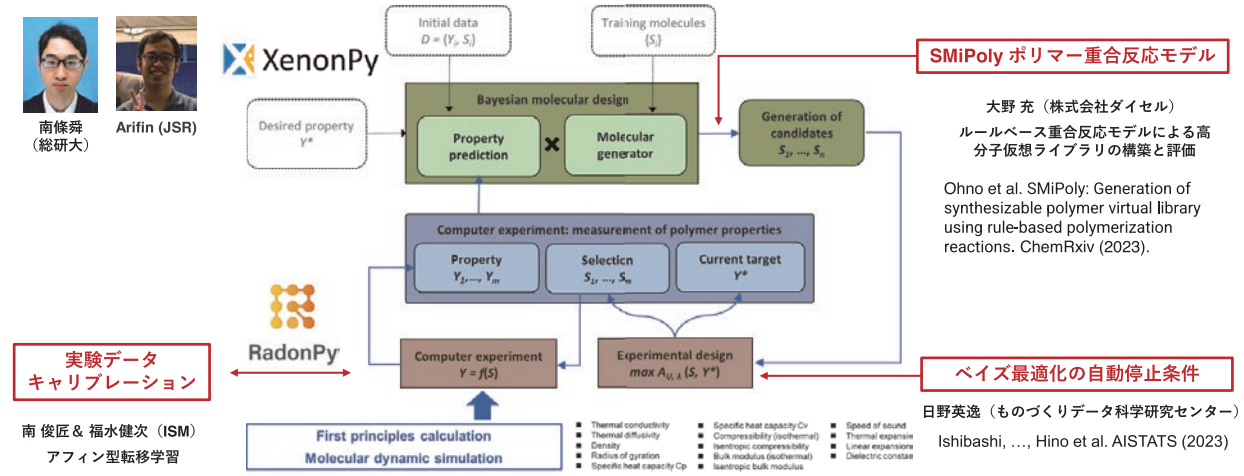
● スライド 14

SPACIER: RadonPy自動実験×キャリブレーション×生成AI×適応的実験計画法

世界初のポリマー物性自動計算×自動設計システム：センターの人員と技術を総結集

南條舜（総研大）

Arifin (JSR)



大野 充（株式会社ダイセル）
ルールベース重合反応モデルによる高分子仮想ライブラリの構築と評価

Ohno et al. SMiPoly: Generation of synthesizable polymer virtual library using rule-based polymerization reactions. ChemRxiv (2023).

南 俊匠 & 福水健次（ISM）
アフィン型転移学習

日野英逸（ものづくりデータ科学研究センター）
Ishibashi, …, Hino et al. AISTATS (2023)

Minami & Fukumizu et al. Transfer learning with affine model transformation. arXiv (2022) https://doi.org/10.48550/arXiv.2210.09745.

● スライド 15

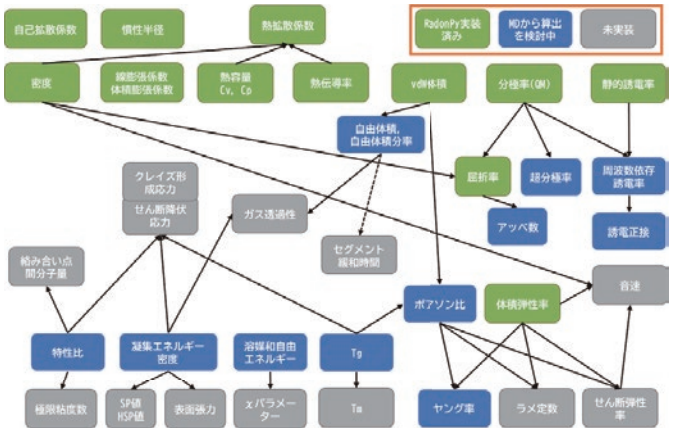
RadonPyの多機能化・高度化：コンソーシアム型オープンソース開発

産学連携コンソーシアムによるRadonPyオープンソース開発

- 多彩な高度専門人材による開発の加速
- 世界標準のソフトウェアに

■ ガラス転移温度（企業コントリビュータ）
■ 配向構造の物性（古屋 秀峰 東京工業大学）
■ 溶媒和自由エネルギー（山田 寛尚 東京薬科大）
■ レオロジー物性（企業コントリビュータ）
■ 量子化学計算（林 慶浩 統数研）
■ 架橋ポリマー（企業コントリビュータ）
■ 誘電特性（古屋 秀峰 東京工業大学）
■ 共重合体（南條 舜 総研大）
■ 高分子溶液
■ 物性の温度依存性・組成依存性
■ 低分子化合物の材料物性（企業コントリビュータ）
■ 生分解性ポリマー（篠田恵子 統数研）

計算対象物性の拡張計画



● スライド 16

年々増加する共同研究・産学連携

ユニークネス：データ科学の研究拠点 × 独自の材料データベース × 最先端の材料研究 の合流地点

大型競争的研究資金

2015-2020 JSTイノベーションハブ構築支援事業「情報統合型物質・材料開発イニシアティブ」（吉田・福水）
2015-2021 JST-CREST数理モデリング「ソフトマター記述言語の創造に向けた位相的データ解析理論の構築」（主たる共同研究者 福水）
2019- JST-CREST熱制御「高分子の熱物性マテリアルズインフォマティクス」（主たる共同研究者 吉田）
2019- 科研費 新学術領域「ハイパーマテリアルのインフォマティクスと hidden order の探索」（計画研究 吉田）
2019- 科研費 基盤研究A 「機械学習の先進技術による革新的機能性物質の発掘」（代表 吉田）
2020- NEDO超超プロジェクト「深層学習・転移学習等の適用による材料設計技術の高度化」（代表 吉田）
2021- 文科省「富岳」成果創出加速プログラム「データ駆動型高分子材料研究を変革するデータ基盤創出」（代表 吉田）
2021- JST未来社会創造事業「マテリアル探索空間拡張プラットフォームの構築」（日野・吉田）
2022- JST-CREST未踏物質探索「フェイゾンエンジニアリング：構造タイル組み換えに基づく新物質創製」（主たる共同研究者 吉田）
2023- JST COI-NEXT「再生可能多糖類植物由来プラスチックによる資源循環社会共創拠点」課題 6 高分子インフォマティクス・機械学習（代表 林）

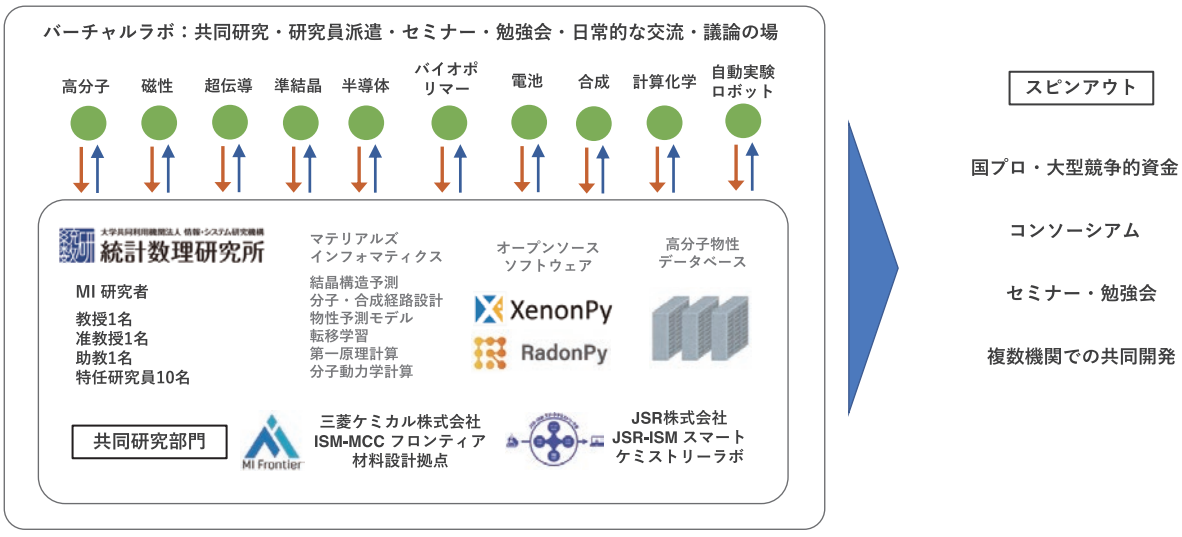
産学連携など

2019- 共同研究部門 三菱ケミカル株式会社
2020- 共同研究部門 JSR株式会社
2017-2023 企業との共同研究・学術指導契約 累計75件 ※センター設立時からパートナーシップを継続している企業がほとんど
2017- MOU 国立研究開発法人 物質・材料研究機構
2021- RadonPy高分子物性データベース共同開発コンソーシアム（3大学・25企業, 138名） ※パートナー企業との協力が大きな力に

● スライド 17

バーチャルなラボを運営する

- サイエンスの複合化が進むと，1研究グループで到達可能な領域が限られてくる．
- 「パートナーシップ」を築き，ラボの活動を外に開放：100-200名規模のバーチャルラボ（単独のグループではできない研究）



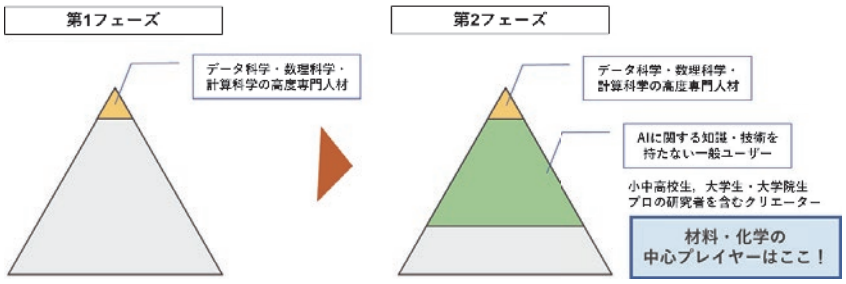
● スライド 18

短中期目標

1. RadonPy高分子物性データベースの公開
2. 材料創生：半導体準結晶，熱ダイオード，超高熱伝導ポリマー，生分解性ポリマー，高分子相溶化剤など
3. JST COI-NEXT「再生可能多糖類植物由来プラスチックによる資源循環社会共創拠点」（～2030年度）
   - ■ 高機能生分解性ポリマーの予測と合成
   - ■ 生分解性ポリマーの包括的データセットと自動合成装置の連携
   - ■ 既存プラスチックを代替

長期目標

1. 超巨大材料データの創出
2. 大規模基盤モデルの開発
3. データ駆動型研究の民主化



● スライド 19

参考資料

● スライド 20

参考資料①

論文・書籍・ソフトウェア等（1/5）

査読付き論文

1. [Editor's Choice] Zhang, Q., Liu, C., Wu, S., Hayashi, Y., Yoshida, R., A Bayesian method for concurrently designing molecules and synthetic reaction networks. Science and Technology of Advanced Materials: Methods, 2204994 (2023).
2. Zamengo, M., Wu, S., Yoshida, R., Morikawa, J. Multi-objective optimization for assisting the design of fixed-type packed bed reactors for chemical heat storage. Applied Thermal Engineering 218, 119327 (2023).
3. [Nature Publishing Group top 3% online attention score, IF = 12.256 (2022-2023), Citation = 5] Hayashi, Y., Shiomi, J., Morikawa, J., Yoshida, R., RadonPy: automated physical property calculation using all-atom classical molecular dynamics simulations for polymer informatics. npj Computational Materials 8, 222 (2022).
4. [IF = 11.021 (2022-2023), Citation = 6] Ma, R., Zhang, H., Xu, J., Sun, L., Hayashi, Y., Yoshida, Y., Shiomi, J., Wang, J-X., Luo, T. Machine learning-assisted exploration of thermally conductive polymers based on high-throughput molecular dynamics simulations, Materials Today Physics 28, 100850 (2022).
5. Iwayama et al. Functional output regression for machine learning in materials science. Journal of Chemical Information and Modeling 62, 4837–4851 (2022).
6. Kusaba, M., Liu, C., Yoshida, R., Crystal structure prediction with machine learning-based element substitution. Computational Materials Science 11, 111496 (2022).
7. Torres, P., Wu, S., Ju, S., Liu, C., Tadano, T., Yoshida, R., Shiomi, J., Descriptors of intrinsic hydrodynamic thermal transport: screening a phonon database in a machine learning approach. Journal of Physics: Condensed Matter 34, 135702 (2022).
8. Minami, S., Liu, S., Wu, S., Fukumizu, K., Yoshida, R., A general class of transfer learning regression without implementation cost. Proceedings of the AAAI Conference on Artificial Intelligence. 35(10):8992-8999 (2021).

● スライド 21

参考資料①

論文・書籍・ソフトウェア等（2/5）

10. [Citation = 33] Ju, S., Yoshida, R., Liu, C., Wu, S., Hongo, K., Tadano, T., Shiomi, J., Exploring diamondlike lattice thermal conductivity crystals via feature-based transfer learning. Physical Review Materials. 5:053801 (2021).
11. [IF = 32.086 (2022-2023), citation = 18, featured article] Liu, C., Fujita, E., Katsura, Y., Inada, Y., Ishikawa, A., Tamura, R., Kimura, K., Yoshida, R., Machine learning to predict quasicrystals from chemical compositions. Advanced Materials 33, 2102507 (2021).
12. Kusaba, M., Liu, C., Koyama, Y., Terakura, K., Yoshida, R., Recreation of the periodic table with an unsupervised machine learning algorithm. Scientific Reports 11, 4780 (2021).
13. Guo, Z., Wu, S., Ono, M., Yoshida, R., Bayesian algorithm for retrosynthesis. Journal of Chemical Information and Modeling 60, 4474-4486 (2020).
14. Toyoshima, Y., Wu, S., Kanamori, M., Sato, H., Jang, M. S., Oe, S., Murakami, Y., Teramoto, T., Park, C., Iwasaki, Y., Ishihara, T., Yoshida, R., Iino, Y., Neuron ID dataset facilitates neuronal annotation for whole-brain activity imaging of C. elegans. BMC Biology 18, 1-20 (2020).
15. [2019.01-2020.12 most downloaded article] Wu, S., Lambard, G., Liu, C., Yamada, H., Yoshida, R., iQSPR in XenonPy: A Bayesian molecular design algorithm. Molecular Informatics 39, 1900107 (2019).
16. [IF = 18.728, Citation = 204] Yamada, H., Liu, C., Wu, S., Koyama, Y., Ju, S., Shiomi, J., Morikawa, J., Yoshida, R., Predicting materials properties with little data using shotgun transfer learning. ACS Central Science 5, 1717-1730 (2019).
17. Takubo, N., Yura, F., Naemura, K., Yoshida, R., Tokunaga, T., Tokihiro, T., Kurihara, H., Cohesive and anisotropic vascular endothelial cell motility driving angiogenic morphogenesis. Scientific Reports 9, 9304 (2019).

● スライド 22

参考資料①

論文・書籍・ソフトウェア等（3/5）

18. [IF = 12.256 (2022-2023), Citation = 222, Springe Nature top 1% citation rank] Wu, S., Kondo, Y., Kakimoto, M., Yang, B., Yamada, H., Kuwajima, I., Lambard, G., Hongo, K., Xu, Y., Shiomi, J., Schick, C., Morikawa, J., Yoshida, R., Machine-learning-assisted discovery of polymers with high thermal conductivity using a molecular design algorithm. npj Computational Materials. 5:66 (2019).
19. Kawamura, Y., Koyama, S., Yoshida, R. Statistical inference of the rate of RNA polymerase II elongation by total RNA sequencing. Bioinformatics 35, 1877–1884 (2018).
20. [Citation = 118, Springe Nature top 3% citation rank] Ikebata, H., Hongo, K., Isomura, T., Maezono, R., Yoshida, R., Bayesian molecular design with a chemical language model. Journal of Computer-Aided Molecular Design 31, 379-391 (2017).

プレプリント

21. Ohno, M., Hayashi, Y., Zhang, Q., Kaneko, Y., Yoshida, R. SMiPoly: generation of synthesizable polymer virtual library using rule-based polymerization reactions. ChemRxiv (2023).
22. Liu, C., Tamaki, H., Yokoyama, T., Wakasugi, K., Yotsuhashi, S., Yoshida, R., Shotgun crystal structure prediction using machine-learned formation energies. arXiv (2023).
23. Wu, S., Yamada, H., Hayashi, Y., Zamengo, M., Yoshida, R., Potentials and challenges of polymer informatics: exploiting machine learning for polymer design. arXiv (2020).

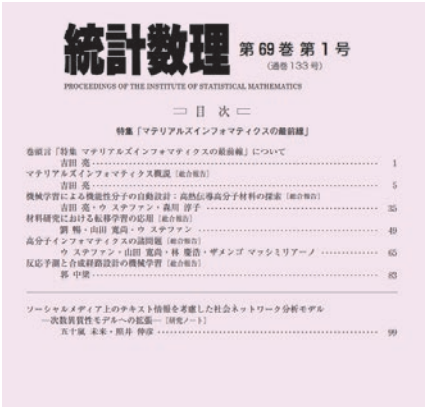
● スライド 23

参考資料①

論文・書籍・ソフトウェア等（4/5）

解説論文

24. 吉田 亮, 林 慶浩, “データ駆動型高分子材料研究におけるデータ資源の不足： シミュレーションと機械学習の役割”, 高分子, 第71巻8月号, pp. 382-384 (2022).
25. 吉田 亮, Chang Liu, “機械学習によるハイパーマテリアルの組成予測とその可能性”, 固体物理, 通巻669号, pp. 127-134 (2021).
26. 吉田 亮, “マテリアルズインフォマティクス概説”, 統計数理：マテリアルズインフォマティクス特集号, 第69巻1号, pp. 5-33 (2021).
27. 吉田 亮, ウ ステファン, 森川 淳子, “機械学習による機能性分子の自動設計: 高熱伝導高分子材料の探索”, 統計数理： マテリアルズインフォマティクス特集号, 第69巻1号, pp. 35-47 (2021).
28. 劉 暢, 山田 寛尚, ウ ステファン, “材料研究における転移学習の応用”, 統計数理： マテリアルズインフォマティクス特集号, 第69巻1号, pp. 49-63 (2021).
29. ウ ステファン, 山田 寛尚, 林 慶浩, ザメンゴ マッシミリアーノ, “高分子インフォマティクスの諸問題”, 統計数理： マテリアルズインフォマティクス特集号, 第69巻1号, pp.65-82 (2021).
30. 郭 中梁, “反応予測と合成経路設計の機械学習”, 統計数理： マテリアルズインフォマティクス特集号, 第69巻1号, pp.83-97 (2021).


統計数理 第69巻 第1号
(通巻133号)
PROCEEDINGS OF THE INSTITUTE OF STATISTICAL MATHEMATICS
目 次
特集「マテリアルズインフォマティクスの最前線」

● スライド 24

参考資料①

論文・書籍・ソフトウェア等（5/5）

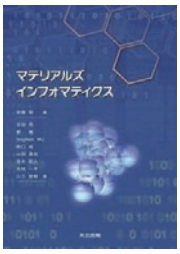
解説論文

31. 吉田 亮, “マテリアルズインフォマティクスによる新物質探索: 高分子材料の設計を中心に”, 週刊医学の歩み, 第276巻9号, pp.861-865 (2021).
32. 吉田 亮, 岩山 めぐみ, グォ チョンリャン, “材料研究の逆問題と機械学習”, 応用物理, 第90巻7号, pp.428-432 (2021).

書籍

33. 伊藤聡編 吉田他著「マテリアルズインフォマティクス」共立出版（2022）

    Pythonライブラリ XenonPy 演習ウェブサイト
    https://github.com/yoshida-lab/XenonPy/tree/master/mi_book



「マテリアルズインフォマティクス」
共立出版（2022）

ソフトウェア・ライブラリ

34. 高分子物性自動計算ライブラリRadonPy: https://github.com/RadonPy/RadonPy
35. マテリアルズインフォマティクスPythonライブラリ XenonPy: https://xenonpy.readthedocs.io/en/latest/
36. ルールベース重合反応仮想ポリマー生成器 SMiPoLy: https://github.com/PEJpOhno/SMiPoly
37. 分子・合成経路自動生成アルゴリズムSeq-Stack-Reaction: https://github.com/qi-zh/Seq-Stack-Reaction
38. 関数出力変数カーネル回帰: https://github.com/yoshida-lab/XenonPy/blob/master/samples/kernel_neural_network.ipynb
39. 結晶構造予測プログラムCSPML: https://github.com/Minoru938/CSPML

● スライド 25

参考資料②

受賞

1. 文部科学大臣表彰「若手科学者賞」：Wu Stephen「統計的機械学習によるデータ駆動型科学の新地平開拓研究」
2. 第２回「富岳」成果創出加速プログラム研究交流会「次世代研究者賞」：林慶浩「高分子物性自動計算システムRadonPyの開発と産学連携によるデータベース共創」

● スライド 26

参考資料③

プレスリリース・新聞報道等（1/4）

国内 プレスリリース

1. 2022年11月9日「全原子古典分子動力学法による高分子物性計算を全自動化する ソフトウェアRadonPyをリリース ～高分子材料物性大地図の作成に向けた第一歩～」https://www.ism.ac.jp/ura/press/ISM2022-08.html
2. 2021年7月20日「機械学習で準結晶を形成する化学組成を同定　準結晶の安定化メカニズムの解明に向けた第一歩」https://www.ism.ac.jp/ura/press/ISM2021-07.html
3. 2020年10月1日「統計数理研究所とJSR、共同研究部門を設置」 https://www.ism.ac.jp/ura/press/ISM2020-09.html
4. 2019年10月1日「物性予測タスク訓練済みモデルの包括的ライブラリXenonPy.MDLを公開 ～転移学習で材料インフォマティクスのスモールデータの壁を乗り越える～」 https://www.ism.ac.jp/ura/press/ISM2019-09.html
5. 2019年8月2日 「大量の欠損を含むデータからでも不具合の要因を特定する機械学習アルゴリズム(AI)を開発 ～品質低下や歩留悪化の要因を高速・高精度に特定し、製造現場の信頼性と生産性の向上に貢献～」
6. 2019年6月26日「機械学習の「記憶」を活用し、高分子の熱伝導性の大幅な向上に成功 ～少ないデータでも高精度な予測が可能に　高分子での材料インフォマティクス加速に期待～」 https://www.ism.ac.jp/ura/press/ISM2019-07.html
7. 2019年6月10日「統計数理研究所と三菱ケミカルの共同研究部門設置について」https://www.ism.ac.jp/ura/press/ISM2019-06.html
8. 2017年11月8日「統計数理研究所と物質・材料研究機構統合型材料開発・情報基盤部門との研究協力に関する協定締結」 https://www.ism.ac.jp/noesuisin/news/except/20171107_nims.html
9. 2017年6月30日「データ科学がもたらす「ものづくり」革新 ～ 統計数理研究所が新センターを設立 ～」https://www.ism.ac.jp/ura/press/ISM2017-03.html

● スライド 27

参考資料③

プレスリリース・新聞報道等（2/4）

海外 プレスリリース

1. 2022/12/22 Automated simulation software creates a world map of polymer properties https://www.eurekalert.org/news-releases/975212
2. 2021/07/29 Machine learning to predict new quasicrystals https://www.eurekalert.org/news-releases/923813
3. 2019/11/05 XenonPy.MDL -- Comprehensive library of pre-trained models for materials properties https://www.eurekalert.org/news-releases/894834
4. 2019/07/19 Successful application of machine learning in the discovery of new polymers https://www.eurekalert.org/news-releases/824893

● スライド 28

参考資料③

プレスリリース・新聞報道等（3/4）

新聞報道・雑誌・TV報道等

1. 2022年11月 日経新聞「統計数理研究所・東工大・東大・JST、全原子古典分子動力学法による高分子物性計算を全自動化するソフトウェアをリリース」
2. 2022年11月 日経産業新聞「高分子の性能、全自動で計算　統計数理研、富岳でデータ集積」
3. 2022年11月 日経新聞（ウェブ版）「高分子性能を自動計算 統計数理研、富岳でデータ集積」
4. 2022年9月 日刊工業新聞「書評「マテリアルズインフォマティクス」共立出版（2022）」
5. 2022年8月 日刊工業新聞「高分子ガラス転移点、スパコン「富岳」で1日1000種計算　三菱ケミがスキーム開発」
6. 2021年7月 マイナビニュース 統数研など、準結晶の相形成を予測する法則を機械学習を用いて発見
7. 2021年7月 OPTRONICS ONLINE 理科大ら，機械学習で準結晶の化学組成を予測
8. 2021年7月 日本経済新聞 統計数理研究所・東大・東京理科大、機械学習で準結晶を形成する化学組成を同定
9. 2021年7月 日刊工業新聞 統計数理研、ＡＩで準結晶の化学組成特定 新物質の探索加速
10. 2020年10月 日刊ケミカルニュース JSR 統数研と共同研究部門を設置、新規材料開発を加速
11. 2020年10月 ゴム報知新聞 ＪＳＲと統数研、機能化学品分野で共同研究部門設置
12. 2020年4月 BSフジ BSフジ「ガリレオX：材料開発を加速させる MI」
13. 2019年10月 週刊ダイヤモンド 第107巻41号 まだまだある！有望サイエンス＆ベンチャー
14. 日経Robotics 2019年10月10日 SEXY TECHNOLOGY「東芝らが「スパースモデリング」技術を改良、欠落が大きいデータでも効率的に学習、解釈もしやすく」
15. 2019年7月 日経産業新聞 少量データで材料開発

● スライド 29

参考資料③

プレスリリース・新聞報道等（4/4）

16. 2019年6月日本経済新聞 統計数理研究所と三菱ケミカル、共同研究部門「ISM-MCC フロンティア材料設計研究拠点」を設置
17. 2019年6月 日刊ケミカルニュース 三菱ケミカル 共同研究部門を10月設置で統計数理研究所と合意
18. 2019年6月 日経xTECH 統計数理研究所と三菱ケミカル、共同研究部門の設置で合意
19. 2019年6月 化学工業日報 物材機構など 少数データでMI 高分子熱伝導が大幅向上
20. 2019年3月 東芝レビュー Vol. 74 No. 2 「高次元・高欠測データのための回帰モデリング技術」
21. 2018年10月15日 日経新聞 9面「新材料開発、脱「経験と勘」　AIで自動化」
22. 2018年1月 情報・システム研究機構「Science Report 013」
23. 2017年12月 テクノロジストマガジン Vol.11 16-17面 「注目の大学研究室 vol.22データ科学の革新的アルゴリズムにより，未知の”新素材”を発見する．他の追随を許さない”ものづくり”革新を！」
24. 2017年11月20日 文教ニュース 53面「統計数理研究所 物材機構と研究協力協定」
25. 2017年8月7日 日経産業新聞 9面「直談 データ科学で新材料開発」
26. 2017年7月7日 科学新聞 2面「統計数理研が新センター設立」
27. 2017年7月4日 日刊工業新聞 23面「戦略目標にモノづくり データ科学研究拠点設立」

● スライド 30

● ものづくりデータ科学研究センター・NOE 活動紹介・事業計画について

【司会】皆さま、お戻りになりましたか。時間となりましたので、再開させていただきます。本日、紹介する NOE の 2 つ目に移りたいと思います。ものづくりデータ科学研究センター、ものづくりデータ科学 NOE の活動についての紹介です。説明は、センター長の吉田教授からお願いします。資料 5 に基づきますので、そちらをご覧ください。吉田センター長、お願いします。

【吉田教授】よろしくお願いします。センター長の吉田です。ものづくりデータ科学研究センターの活動紹介・事業計画についてお話しします。当センターは、2017 年 7 月に発足してから 6 年弱がたっていますが、その重点領域の一つとして、マテリアルズインフォマティクスにフォーカスして、研究活動を進めています。

その経緯についてご説明いたします。2015 年に JST のイノベーションハブ構築支援事業において、わが国におけるマテリアルズインフォマティクスの学術創成のフラグシッププロジェクトが立ち上がりました。このプロジェクトのリーダーが、顧問の伊藤先生ですが、発足して、国内のいろいろな物質科学、計算科学、そしてデータ科学の研究者がここに集められて、そこで私と福水先生が、ここに参画したのがスタートになっています。その 2 年後に、このセンターが発足して、現在、6 年弱がたっているという状況です。この間に、マテリアルズインフォマティクスの大きなムーブメントが起きて、非常に業界が盛り上がったということで、このような時勢の中で、社会的ニーズと、センターの人的資源、研究資源の制約から、マテリアルズインフォマティクスを重点領域に設定しました。

研究体制については、私たちのセンターは、比較的小規模のセンターでして、所内教員が 7 名で、特任研究員が 10 名という体制で進めています。このセンターの体制構築における戦略には、3 つ軸があります。1 つは、産学連携の戦略的活用ということで、多数の産学の研究者と連携して、共創していくことを目指します。2 番目が学際性です。当センターの人員のほとんどは、データ科学あるいは機械学習、統計数理以外の分野の、例えば、第一原理計算や分子動力学シミュレーション、高分子の化学などを専門とする研究者が活動の中心を担っています。もう一つは、統数研の中で生み出される、統計的機械学習などの理論や方法論、このようなものを私たちが拾い上げて、技術移転をしていくという形で、研究活動を進めています。この 3 つの軸について、具体的な例をもって、これからお話ししていきます。

私たちのデータ駆動型材料研究の基本的な考え方である順問題、逆問題のワークフローがここに示されています。例えば、物質の構造や組成が入力されて、その特性を予測します。入力と出力のデータを集めてきて、機械学習のモデルを作ります。そのモデルの逆写像を求めることで、所望の特性を持つ物質をどのようにつくればいいのかという逆問題を解きます。2017 年頃に発表したこの論文が当センターのその後の研究のコンセプト形成において中心的な役割を担ってきました。

2019 年に発表したこの論文では、先ほどの順問題、逆問題を解いて、高い熱伝導率を持

つアモルファスのポリマーを予測・発見しました。例えば、ガラス転移温度と融点という2つのターゲットの物性がありました。データを集めてきて、ポリマーの構造、組成を入力して、その2つの物性を予測します。そのモデルの逆問題を解いて、高いガラス転移温度、高い融点を持つようなポリマーは何なのかという計算をします。順方向のモデルを作る過程で、どのようなポリマーの特徴が、どのように特性に効いてくるかというのを学習させて、その逆問題を解くことで、どのような構造をどのように組み合わせれば、ガラス転移温度、融点が高くなるのかということになります。この論文は、私たちのセンターの中でも引用数の多い論文の一つでして、『Springer Nature』のトップ1%の論文にランキングされています。

コンセプトは非常にシンプルで、順問題、逆問題を解くというのは、材料に限らず、広い分野で、データ科学は、昔からやってきたことです。そのようなものをマテリアルの分野に、どんどん実装して、実践を重ねていくことで、このような研究成果を、6年ぐらいで出してきました。また、それらを実装したソフトウエアをつくり、あとは、この後お話しするデータベースをつくろうという取り組みが、私たちのセンターの活動のコアとなっています。

また、実践・実証に関しましては、センター発足当時は、まだMIで何ができるのかというのが分からない状況でして、一つ一つ新しい分野を開拓していきながら、実際に機械学習で予測した様々な物質を合成してきました。

例えば高分子の材料では、最近、極めて高い熱伝導率を有する高分子液晶を発見しました。また、準結晶・近似結晶という物質が大量に機械学習によって発見されています。その他にも、いろいろなコラボレーションの中で、いろいろな材料系を対象に実証研究を推進しています。

コンセプトは非常にシンプルですが、マテリアルズインフォマティクスの場合は、データ資源の不足という大きな問題を乗り越えなければなりません。ここに、今、現時点で存在する高分子物性のデータベースの一覧が記されていますが、データ量は、機械学習やデータ科学の他の応用分野からすると、非常に少ないです。オープンデータの不足という問題は、技術的な要因だけでなく、データ共有に関する文化的な問題も関わっているため、短中期的には解決が難しいことが予想されます。また、材料の研究の1つの特徴としては、革新的な材料を見つけたいわけですが、当然ながら、革新的な材料の周辺には、まだデータがありません。一方で、データ科学の予測というのは、基本的には内挿的であり、データが全くないような未踏領域の新物質、新材料を予測できません。すなわち、マテリアルズインフォマティクスの多くの本質的な問題は、データがないということに帰着します。

この現状を何とか突破したいと考えてきたわけですが、私たちのセンターでは、データ駆動型材料研究の将来像を定めて、それに向かって活動を進めています。今後、オープン領域において、シミュレーションをベースとした基礎物性や、物質の構造のデータベースが大量にできてくるでしょう。もう一つは、大学や大型実験施設等々から生成されるデータを収集した、中規模程度の実験データベースができていくことが予想されます。一方で、クローズ

ドな領域では、例えば 1 企業や大学の研究室などは、データが全然ないわけです。したがって、このようなオープン領域におけるビッグデータと、クローズド領域における非常にスモールなデータをうまく接続するようなデータ科学の方法論と、あとは、このデータベース、シミュレーションの大規模データベースを私たちのほうでつくろうと考えに至りました。

そこで、2020 年度後半ぐらいに、私たちのセンターの林慶浩助教が、全原子の古典分子動力学計算を全自動化、いわゆる MD シミュレーションと呼ばれるものを全自動化する、RadonPy というソフトウエアを開発しました。RadonPy は、ポリマーの組成や構造情報、また、実験、計算条件を入力すると、MD 計算の非常に煩雑なプロセスを全自動化して、いろいろな物性を網羅的に計算してくれます。2022 年に RadonPy の論文を発表しましたが、いろいろな反響がありました。例えば、Nature Publishing Group の online attention スコアで、同時期に出版された論文の、今、トップ、3%にランクインしているという状況です。今後、この分野のデファクトスタンダードのソフトウエアになっていくということを期待しています。

このような計算機の中の高分子物性の実験を全自動化することに成功しました。あともう一つは、私たちのような小さな研究グループでは、リソース的な制約により、大規模且つ包括的なデータベースを構築することは難しいという問題があります。そこで、単独でできないのであれば、いろいろな人を巻き込もうということで、2021 年に産学連携コンソーシアムを形成しました。今は統数研と 3 大学、24 企業、138 名の参画者がいますが、このような方々を巻き込んで、RadonPy のソフトウエアの開発とデータベースの共同開発を推進しています。

また、この事業は、今、富岳成果創出加速プログラムという国プロに採択されていまして、富岳の大量の計算資源を割り当てられて、富岳をフル稼働して、データのプロダクションを行っています。

最終的には、10 の 5 乗や 10 の 7 乗のポリマーが包含する、世界最大級の計算物性のデータベースを創出することを目指しています。

今、粛々とこのプロジェクトを進めておりまして、現時点で約 5 万ポリマーの計算が完了しています。その結果、このような複数の高分子物性の同時分布などが見えてきたり、物性間のトレードオフの関係が形成するパレートフロントが鮮明に見えてきました。実験だけでは、このような物性分布の全体像を観測するのは難しいですが、計算の世界では、高分子物性の世界地図のようなものを作成することがきます。

また、RadonPy データベースの主な用途としては、機械学習に活用するデータ資源を考えています。例えば、転移学習やマルチフィディリティ学習のデータリソースへの活用が感がられます。例えば、RadonPy で、5 万ポリマーぐらいの、このような 17 種類の物性がデータを生成しましたが、これを転移学習におけるソースタスクのデータとして活用して、限られた少数の実験データでファインチューニングしたり、転移を行ったりして、実際のリアルワールドの問題を解いていていきます。

これはいわゆる Sim2Real や Sinthetic-to-Real という設定の転移学習になります。当センターでは、このような技術を使って、これまでに様々な材料研究を展開してきました。

あともう一つ重要なのは、計算機実験のデータベースは不完全であるということです。MD 計算は、計算条件やパラメータに依存するので、それらをどのように適切に設定するのか。適切な条件を設定したとしても、やはり現実系との間に、このようなバイアスやばらつきが出てきます。横軸は実験値で、縦軸は MD の計算値です。このように幾つかの物性については、どうしてもこのように不整合性が生じてしまいます。その間を機械学習で埋めることで、このようにバイアスやバリアンスをうまく低減していくことが重要になってきます。

次に、Sim2Real 転移のべき乗則についてお話しいたします。シミュレーションデータを、これからどんどん大きくしていくわけですが、シミュレーションデータを用いてプレトレーニングしたモデルを現実系にファインチューニングしたモデルの汎化性能は、このようなべき乗則に従うことが知られています。シミュレーションデータのサイズを増やしていくと、実験物性に対する予測性能が、このように改善していきます。このようにべき乗則のパターンを実験的に観測しながら、どのようにデータを作ればいいのか、そのようなデータベース設計の指針を得るのに活用していきます。Sim2Real 転移のべき乗則は福水先生らによって発見されたものですが、そのアイデアを我々のプロジェクトに技術的に移転したということになります。

また、このように高分子の合成や物性測定が計算機の中で全自動化されると、それと機械学習を接続して、コンピュータの中でバーチャルな物質探索を行うことができます。例えば、機械学習を使って、あるターゲット物性を対象に、生成 AI などを使って、仮想分子を大量に生成していきます。これを全部、実験で物性を測定するのは難しいので、実験計画等を使って、うまく絞り込んであげます。その絞り込んだ仮想物質の物性を、第一原理計算や RadonPy の MD 計算を使って測定します。ただ、これは計算なので、機械学習を使って実験データにキャリブレーションしてあげます。さらに、新しく生成されたデータを使って機械学習の予測性能を改善します。このループを全自動化するシステムを開発しています。ちなみに、このシステムには、私たちのセンターの共同研究者、センターの教員、スタッフの人たちの様々な研究成果が含まれています。

このようなシステムができると、コンピュータの中で、いろいろな物質を自由自在に設計できるようになります。

例えば、これは比熱と屈折率を対象にした高分子の設計を表した事例です。データが全くない状況でも適当にターゲット物性の範囲を設定すれば、ここに存在する高分子をコンピュータの中で自動生成します。次に実際の応用例を紹介します。光学材料の高分子のアッベ数と屈折率を向上させることが目標となりますが、これらの物性の間にはトレードオフの関係が存在するため、このあたりにパレートフロントがあることが経験的に分かっています。これを突破するような分子を、どんどんコンピュータの中で生成して、それを合成に回

していくという、そのようなスキームが完成しています。

あとは、RadonPy の計算機実験の多機能化・高度化をどんどん進めていけば、もっといろいろなことができるようになっていきます。ただ、なかなか私たちのような 1 研究グループでこれを実現するのは大変なので、自分たち単独でできないのであれば、コンソーシアムの力を借りて、みんなでオープンソースの開発をやっていこうということになります。最終的には、ここに書いてある全ての高分子物性の全自動計算を RadonPy に実装します。この緑のところが、今実装しているところで、次のリリースで、この青の部分、最終的には、このグレーの部分を含めた全部の物性をコンプリートします。コンソーシアムには、産学の多彩な高度専門人材が参画しています。そのような方々が、今、RadonPy のアルゴリズム開発やソフトウエア開発を推進しています。このような活動を今後何年も積み重ねていきながら、世界標準のソフトウエアを完成させます。

こちらのスライドはセンターの大型競争的外部資金の獲得状況や産学連携の概要をまとめたものです。重要な点は、2015 年にイノベーションハブ構築支援事業がスタートして、そこから年々プロジェクトが増加していっているということと、産学連携につきましては、三菱ケミカル株式会社、JSR 株式会社との共同研究部門を設立して、研究所の中に、バーチャルなラボを形成して、共同研究を推進しているということです。また、企業との共同研究の件数も、年々増加していて、計算すると、約 6 年半で累計 75 件に達しています。これに合わせて RadonPy のコンソーシアムに 3 大学 24 企業、138 名が参画しています。

ここでお伝えした点が 2 つあります。まずは、センターの設立時からパートナーシップをずっと継続している企業がほとんどであるということが、センターの取り組みの成功の証の一つであるということです。もう一つは、なぜこのようにたくさんの人が当センターのパートナーシップに加わっていただけるかと点についての私なりの分析ですが、現在のマテリアルズインフォマティクスのプロジェクトは、基本的には、材料の研究拠点が、世界的にも、日本においても中心になっています。一方で、当センターはデータ科学の研究拠点でありながら、独自の材料データベースを持っています。そこに最先端の材料分野の研究者が合流していくという、非常にユニークな拠点が形成されています。これがもう 1 つの成功の理由であると思っています。

それから、当センターは産学連携に大きな特色があります。企業との共同研究では特定課題だけではなくて、オープンサイエンスでの協働プロジェクトをスピンアウトさせています。実際、RadonPy のデータベースのコンソーシアムはスピンアウト企画の典型例です。長年の間の多数の企業との協業を通じて信頼関係を構築してきたからこそ、たったの 2 年でこのような大きな取り組みを実現できたのだと思います。

それから、バーチャルラボという話がありましたが、今、現時点においても、それに資するような取り組み、活動をやっています。統計数理研究所のものづくりデータ科学研究センターには様々な計算技術や研究リソースがあります。

そこに共同研究部門などがあるわけですが、私たち、単独では、できることは限られてい

るので、外部の人たちを、1つのラボのような形で取り込んでいこうと考えています。昨今のサイエンスは、複合化が進んでいて、1研究グループでできることというのは、非常に限られてくるし、今後も、そのような傾向は、どんどん進んでいくでしょう。

私たちの考え方は非常にシンプルで、私たちの普段の研究活動を外に開放して、一緒にみんなで何か勉強したり、セミナーをしたり、そのような非常に緩い連携を創り出すことを心掛けています。そうすると、このような100名、200名規模の仮想的なラボを非常に簡単につくることができます。そうすると、単独のグループではできないような未踏領域を切り開くチャンスが生まれるという、そのような感じです。実際に、このような活動を6年間いろいろとやってきたことで、日常的な議論の中からスピンアウトさせた面白い取り組みができていると思っています。

最後に、今後の目標について説明いたします、短期的な目標につきましては、これまでやってきたことを粛々と続けていくだけです。RadonPyのオープンデータベース開発については、2年以内にコンソーシアムの連名で論文発表をして、データベースを公開します。産学連携によるオープンデータベースの共同開発という、新しいモデルをデモンストレーションします。また、特定の物質、革新的な特性を持った物質を、ターゲットにしたデータベース作成であったり、新材料創生をさらに加速していきたいです。

長期的な目標としましては、このような基本的には、データ駆動型研究というのは、本質的にはデータを持つものが勝つという資本のゲームです。そのような時代の流れを見ながら、私たちも、ビッグデータホルダーになろうということで、このような取り組みを始めたわけですが、そのような超巨大なデータと機械学習を組み合わせることで、データ駆動型研究に民主化を起こしていきたいと、次のステージでは考えています。

今はデータ駆動型研究の第1フェーズです。現時点で、データに基づき予測をしたり、意思決定をしたりしている人たちは、材料研究者の中のトップ1%にすぎません。ただ、この人たちだけでデータ駆動型研究をやっていても、あまり大きなイノベーションは起きないと予想しています。最終的には、材料研究の中心プレーヤーであるこの層にどのようにアプローチしていくかということが、重要になってきます。しかしながら、この層の人たちに、自らデータを取得し、Python でコードを書いて、データ解析をしろというのは無理です。ですから、そのような人たちが何らかの形で、ビッグデータに基づく予測や意思決定ができるような、そのようなことをどのようにしたら実現できるのかということを、このセンターの次のフェーズで考えていきたいです。以上となります。どうもありがとうございました。

【司会】どうもありがとうございました。ただ今の報告におけるコメント、質問等があったら、お願いします。伊藤先生、お願いします。

【伊藤顧問】伊藤です。大変素晴らしい成果が上がっているので、非常に感銘を受けたところですけれども、このような話は、材料科学なので、最後は、フィジカルの世界、実際の材料につながらなければならないと思います。そこに関しては、吉田先生のご説明にもありま

したように、統数研で行うよりも、むしろ、パートナーのほうにやっていただくという考え方だと思うのですけれども、この考え方は非常によろしいと思います。ただ、その時に、そのパートナーがやったことを、どのようにセンター側にうまくフィードバックをかけていただくか。というのは、新しい人たちに、この分野に参加してもらうためには、このようなやり方によって、どのような事例ができたのかという、事例がないと、なかなか参入してこないと感じています。そのような事例をセンターとして示すためにも、ぜひパートナーにやっていただいたもののうち、もちろん公開できないものはいいのですけれども、うまく戻せるものは、戻していただくような仕組みを、ぜひお考えいただければと思いました。それが1点です。

それから、あと2点。既にたくさんの事業と連携しているので、これ以上、連携しろというのは、やや酷かもしれないのですけれども、今、文部科学省のほうでは、マテリアルDXのようなプロジェクトも動いていて、研究 DX を進めましょうという結構大きな事業が幾つか動いています。

そこに直接参画していただかなくて結構なのですけれども、そことうまく連携を取っていただいて、ぜひこのような統数研の仕事が、他のマテリアルDXのプロジェクトでも使っていただけるように、ご検討いただければと思います。

最後の点は、質問なのですけれども、このような事業は、NIMSに私がいた時も非常に思ったのですが、これは継続が、なかなか大変で、例えば、今、富岳を使って高分子物性データベースをおつくりになられています。COI-NEXT は金沢大学拠点の事業だと思うのですけれども、COI-NEXTに採択されているので、10年ぐらいこのようなことを進めることができるとは思います。けれども、やはり自立的にデータがうまく集まってくるような仕組みを作る必要があると思います。この辺に関して、今の時点で、吉田先生は、何かお考えになっておられるでしょうか。

【吉田教授】どうもありがとうございます。いろいろとあるのですけれども、まず持続性というところに関してです。持続性は、非常に重要な問題で、データベース開発は富岳成果創出加速プログラムにより、今、支援されていますけれども、この事業が、あと3年で終了になります。いまは後継の同じようなプログラムが始まることを見越して、いろいろな活動を展開しています。それから、コンソーシアムは、今、無料でやっていますが、富岳成果創出加速プログラムが終了する 3 年後は、場合によっては、有料化への移行もあり得ると思います。

また、研究所として、このような活動を支援していただけるチャンスもあるのかなと思っています。それから、国全体のデータソースへの動きとの接続なのですけれども、まず私たちは、RadonPyのデータベースを国内の唯一の高分子物性のデータベース、これは世界的にも唯一ですが、この実験データベースであるPoLyInfoとうまく接続したいと考えています。

【伊藤顧問】これだけいい仕事が進んでいるので、ぜひ継続してやっていただく、少なくと

も 10 年スパンで維持できる仕組みを作っていただきたいです。NIMS の材料データベース MatNavi も、当初は 3 年間のプロジェクトで、データベースをつくるという事業でした。3 年間でつくって、その後、ずっとそれを維持するのでは、データベースとしては意味がないですよね。

【吉田教授】そうですね。

【伊藤顧問】そこは、吉田先生がよくご存じのとおりなので、そこをうまくやっていく。計算資源という意味で、今富岳が必要ならば、もちろん富岳を使っていけるといいと思いますし、少し古い話になりますけれども、SETI@home のような BOINC、ネット上で分散してやるボランタリコンピューティングのようなやり方もないことはないというので、いろいろなやり方があると思います。そこら辺をぜひご検討いただければと思いました。また、もう一つ大事なのは、研究所としてデータベースシステムを 10 年スパンで維持していただくということが、とても大事だと思います。ぜひよろしくお願いします。大変面白かったです。ありがとうございました。

【吉田教授】ありがとうございます。

【司会】ありがとうございます。他、質問やコメントはいかがでしょうか。

時間も迫っていますので、次のプログラムに行かせていただければ、ありがたいです。次は、所内のスーパーコンピュータについて上野統計科学技術センター長から説明させていただきたいと思います。上野先生、よろしくお願いします。

The Supercomputer System

第4回 統計数理研究所 NOE形成事業 顧問会議

統計数理研究所の スーパーコンピュータ

統計科学技術センター長
モデリング研究系 データ同化グループ
教授　上野　玄太

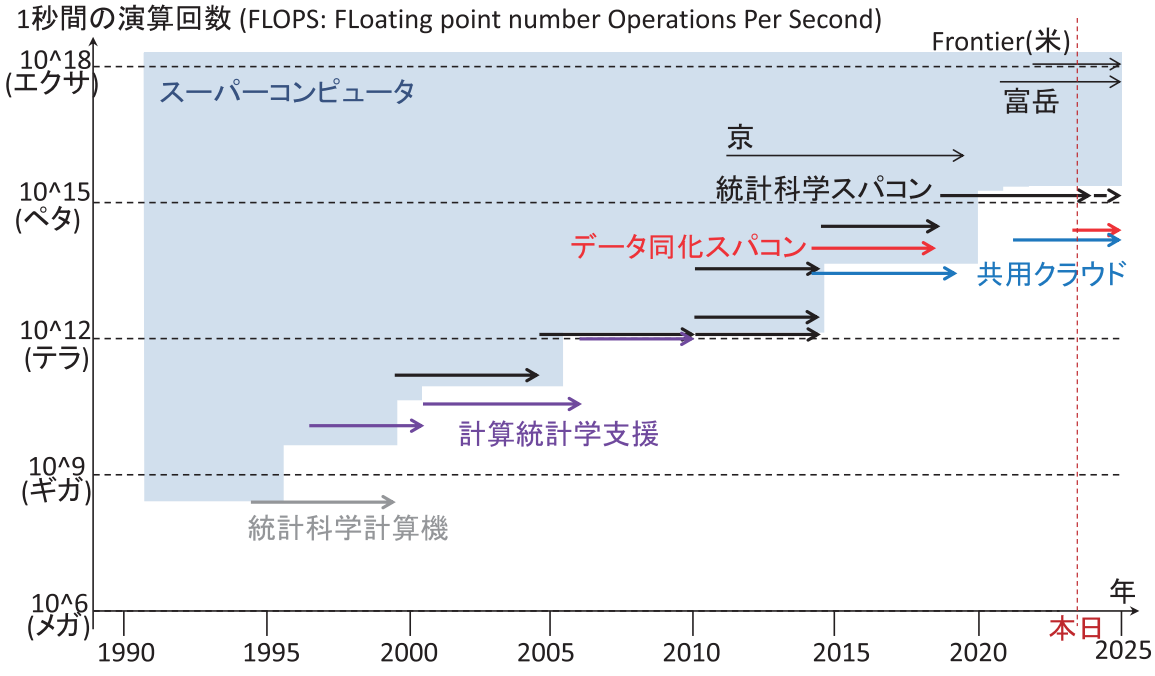
統計数理研究所の スーパーコンピュータについて

2023年5月16日
統計科学技術センター
上野玄太

大学共同利用機関法人 情報・システム研究機構
統計数理研究所

● スライド 1

スパコンの定義 (政府調達)



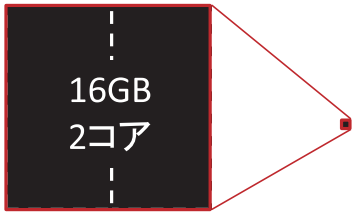
● スライド 2

上野のPC（2017年購入）

- 1台 = 1ノード
- 1台 = 1ノードあたり
  メモリ16GB、2コア

メモリ≒敷地面積、まな板の大きさ

コア数≒テニスコートが何面とれるか、板前の人数

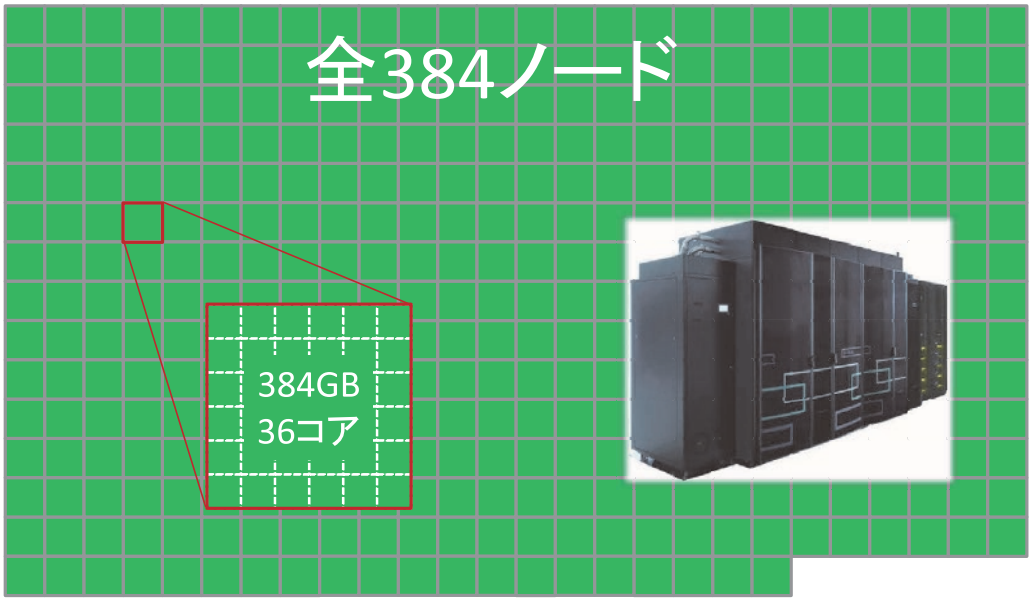


← 以降での16GBの大きさ

● スライド 3

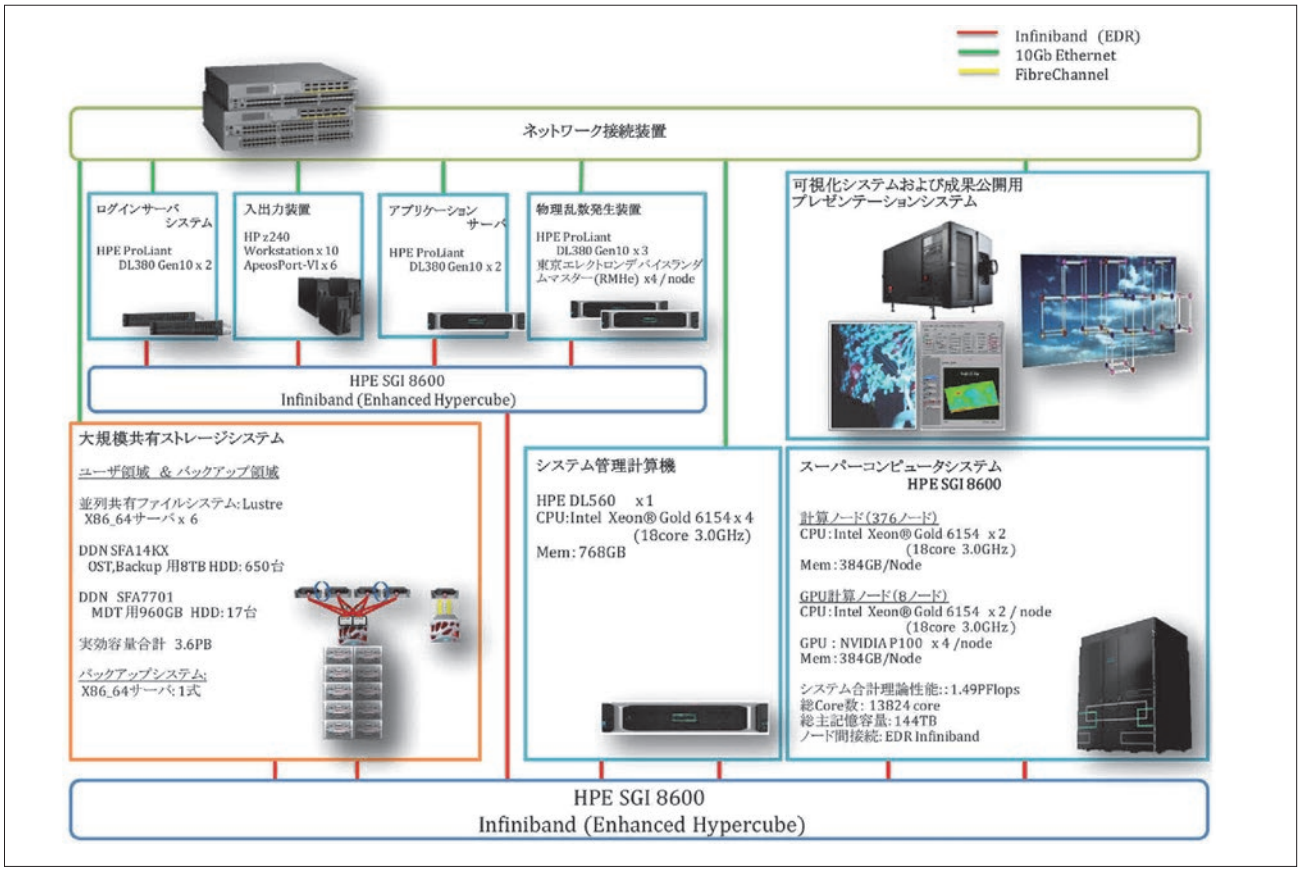
統計科学スーパーコンピュータシステム

2018年9月～

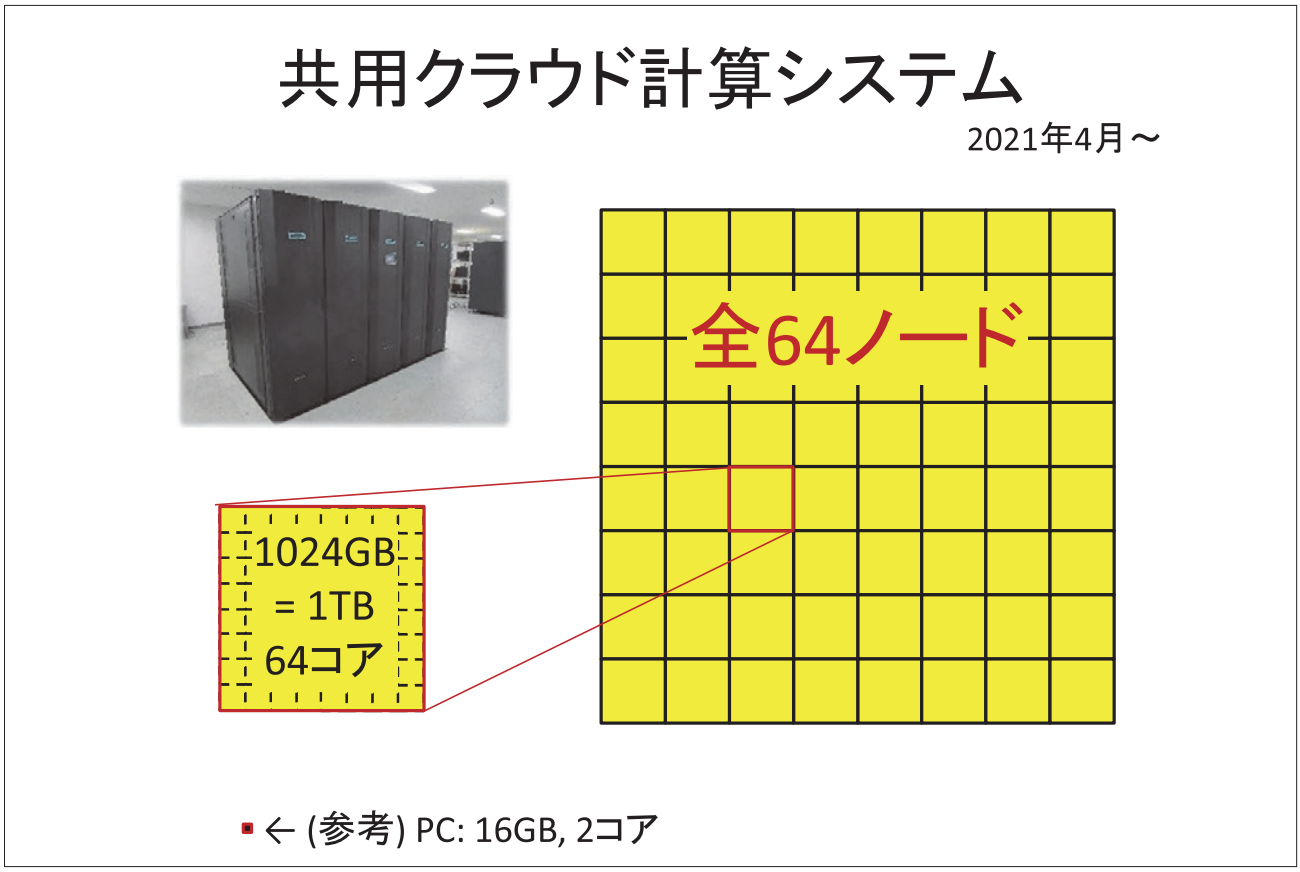


- ← (参考) PC: 16GB, 2コア

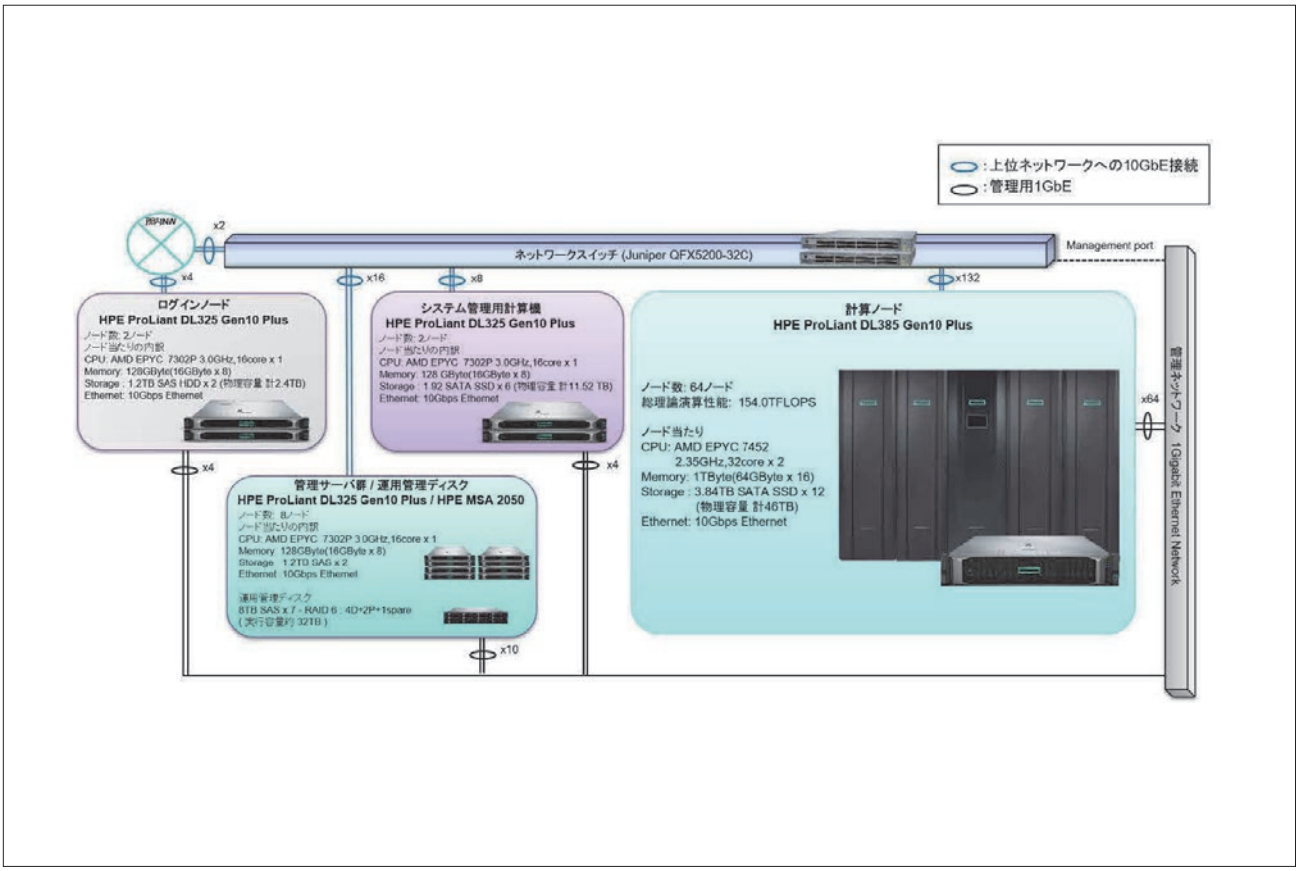
● スライド 4


Infiniband (EDR)
10Gb Ethernet
FibreChannel
ネットワーク接続装置
ログインサーバ システム
HPE ProLiant DL380 Gen10 x 2
入出力装置
HP z240 Workstation x 10
ApeosPort-VI x 6
アプリケーション サーバ
HPE ProLiant DL380 Gen10 x 2
物理乱数発生装置
HPE ProLiant DL380 Gen10 x 3
東京エレクトロンデバイスランダムマスター(RMHe) x4 / node
可視化システムおよび成果公開用プレゼンテーションシステム
HPE SGI 8600
Infiniband (Enhanced Hypercube)
大規模共有ストレージシステム
ユーザ領域 & バックアップ領域
並列共有ファイルシステム: Lustre
X86_64サーバ x 6
DDN SFA14KX
OST,Backup 用8TB HDD: 650台
DDN SFA7701
MDT用960GB HDD: 17台
実効容量合計 3.6PB
バックアップシステム:
X86_64サーバ: 1式
システム管理計算機
HPE DL560 x1
CPU:Intel Xeon® Gold 6154 x 4
(18core 3.0GHz)
Mem: 768GB
スーパーコンピュータシステム
HPE SGI 8600
計算ノード(376ノード)
CPU: Intel Xeon® Gold 6154 x 2
(18core 3.0GHz)
Mem: 384GB/Node
GPU計算ノード(8ノード)
CPU: Intel Xeon® Gold 6154 x 2 / node
(18core 3.0GHz)
GPU : NVIDIA P100 x 4 /node
Mem: 384GB/Node
システム合計理論性能:: 1.49PFlops
総Core数: 13824 core
総主記憶容量: 144TB
ノード間接続: EDR Infiniband
HPE SGI 8600
Infiniband (Enhanced Hypercube)

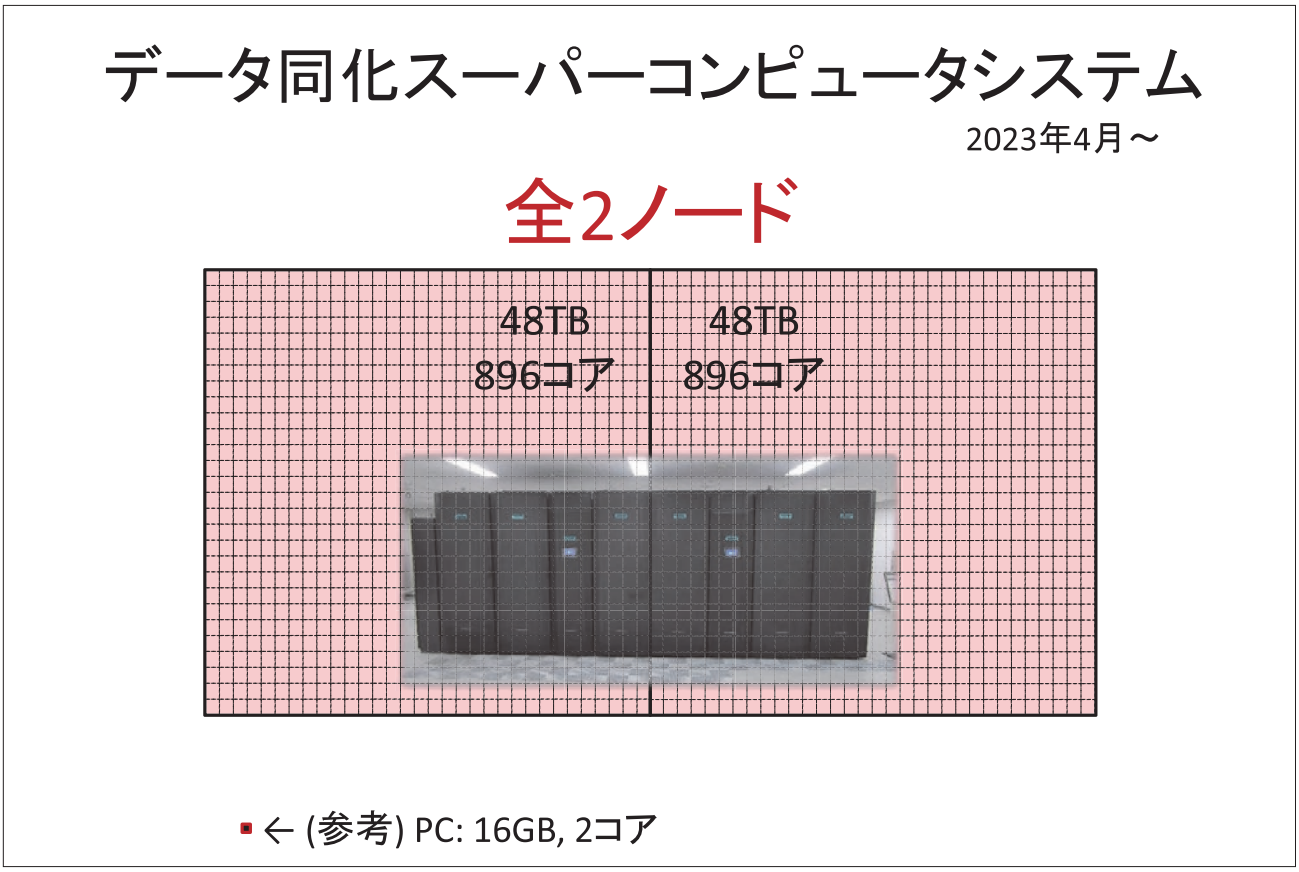
● スライド 5


共用クラウド計算システム
2021年4月～
全64ノード
1024GB
= 1TB
64コア
← (参考) PC: 16GB, 2コア

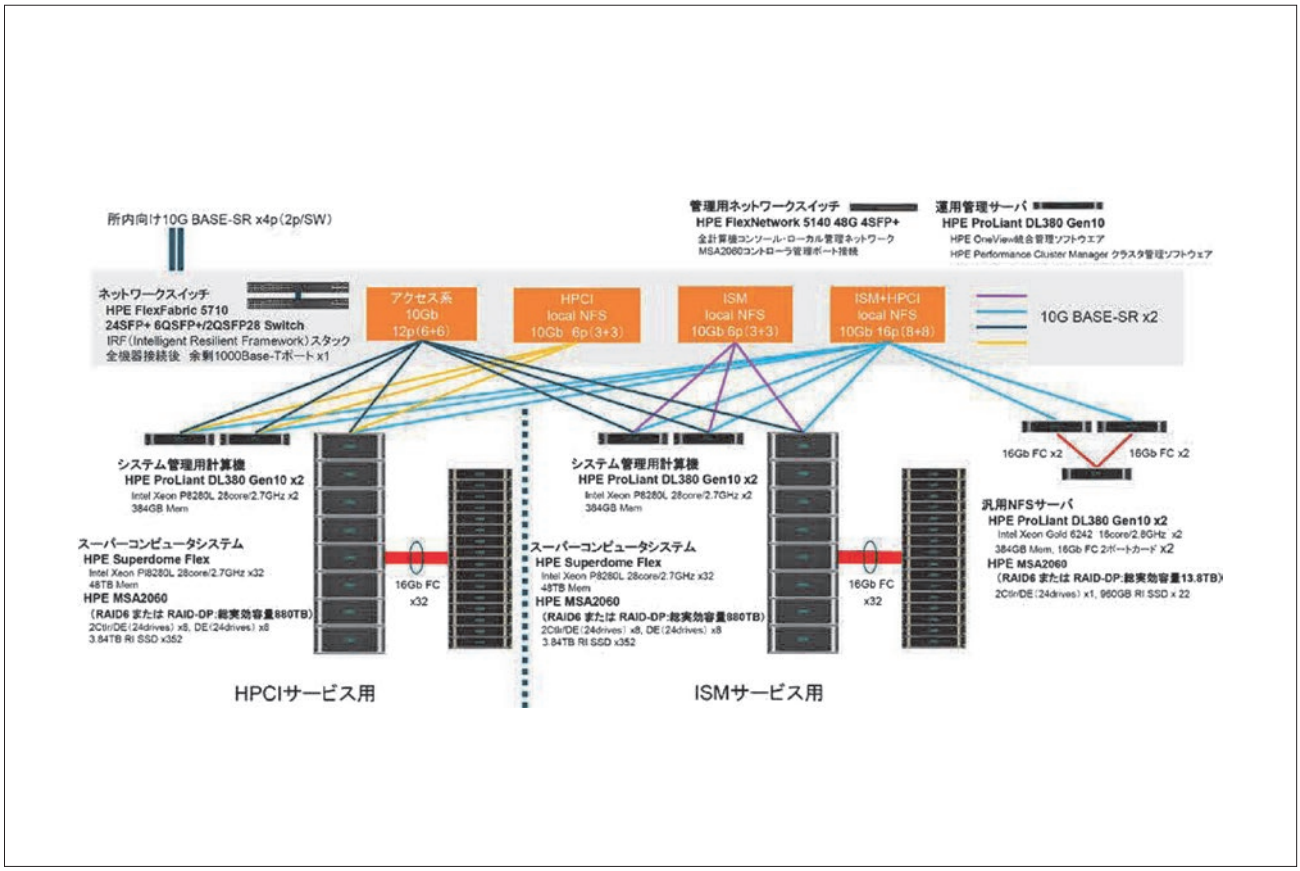
● スライド 6


:上位ネットワークへの10GbE接続
:管理用1GbE
x2
ネットワークスイッチ (Juniper QFX5200-32C)
Management port
x4
x16
x8
x132
ログインノード
HPE ProLiant DL325 Gen10 Plus
ノード数: 2ノード
ノード当たりの内訳
CPU: AMD EPYC 7302P 3.0GHz,16core x 1
Memory: 128GByte(16GByte x 8)
Storage : 1.2TB SAS HDD x 2 (物理容量 計2.4TB)
Ethernet: 10Gbps Ethernet
システム管理用計算機
HPE ProLiant DL325 Gen10 Plus
ノード数: 2ノード
ノード当たりの内訳
CPU: AMD EPYC 7302P 3.0GHz,16core x 1
Memory: 128 GByte(16GByte x 8)
Storage : 1.92 SATA SSD x 6 (物理容量 計11.52 TB)
Ethernet: 10Gbps Ethernet
計算ノード
HPE ProLiant DL385 Gen10 Plus
ノード数: 64ノード
総理論演算性能: 154.0TFLOPS
ノード当たり
CPU: AMD EPYC 7452
2.35GHz,32core x 2
Memory: 1TByte(64GByte x 16)
Storage : 3.84TB SATA SSD x 12
(物理容量 計46TB)
Ethernet: 10Gbps Ethernet
x64
管理ネットワーク 1Gigabit Ethernet Network
x4
x4
管理サーバ群 / 運用管理ディスク
HPE ProLiant DL325 Gen10 Plus / HPE MSA 2050
ノード数: 8ノード
ノード当たりの内訳
CPU: AMD EPYC 7302P 3.0GHz,16core x 1
Memory: 128GByte(16GByte x 8)
Storage : 1.2TB SAS x 2
Ethernet: 10Gbps Ethernet
運用管理ディスク
8TB SAS x 7 - RAID 6 : 4D+2P+1spare
(実行容量約 32TB)
x10

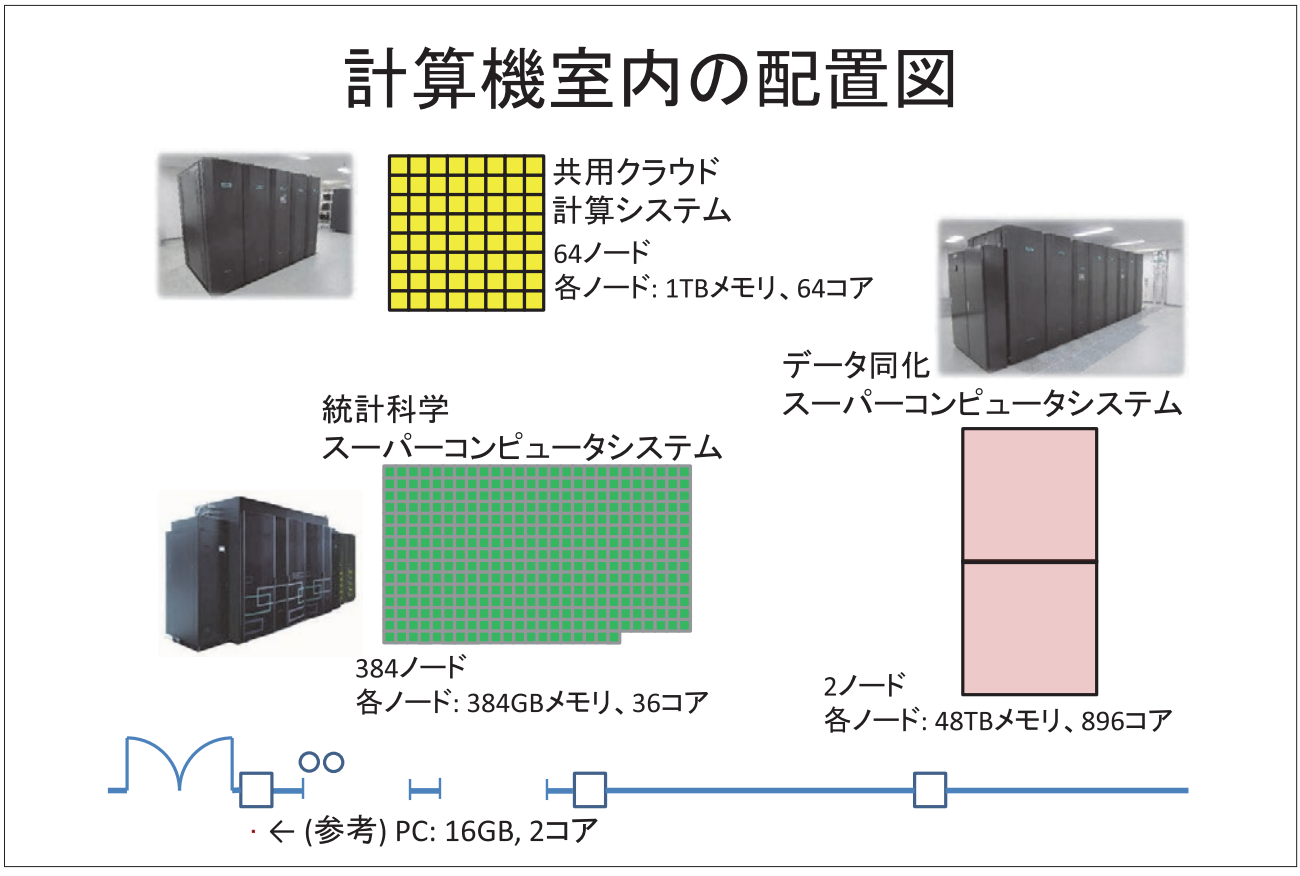
● スライド 7


データ同化スーパーコンピュータシステム
2023年4月～
全2ノード
48TB
896コア
48TB
896コア
← (参考) PC: 16GB, 2コア

● スライド 8



● スライド 9



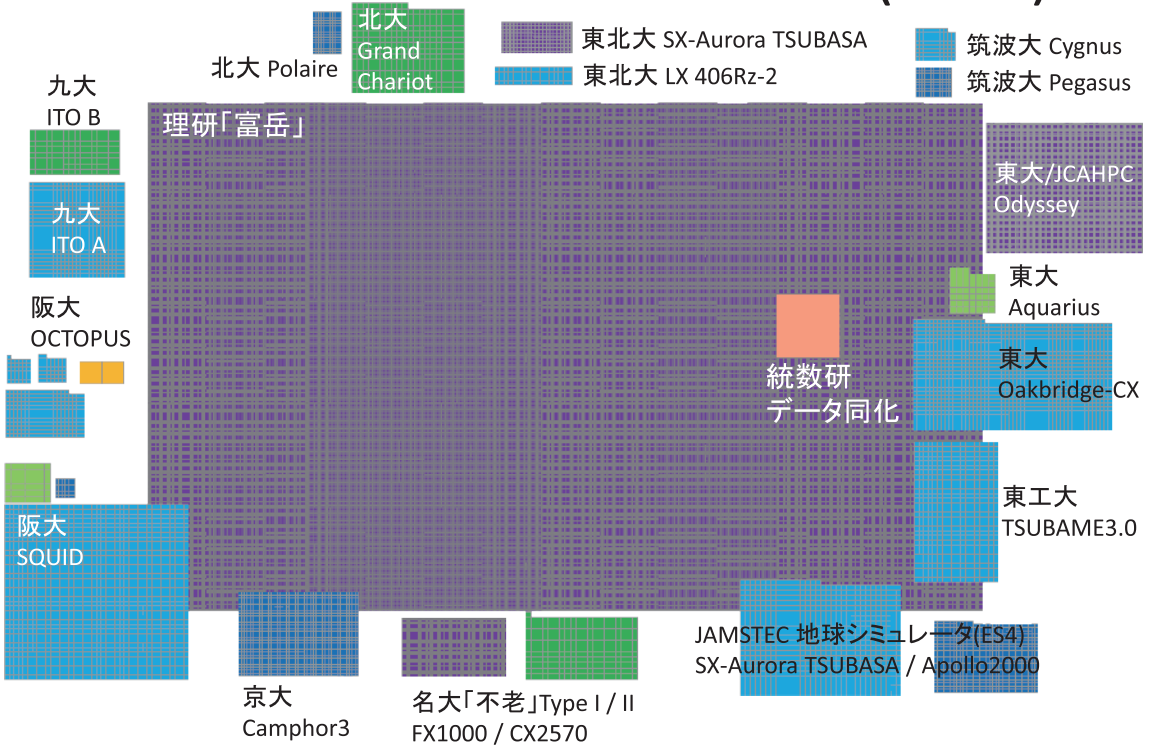
● スライド 10

データ同化スパコンの導入風景

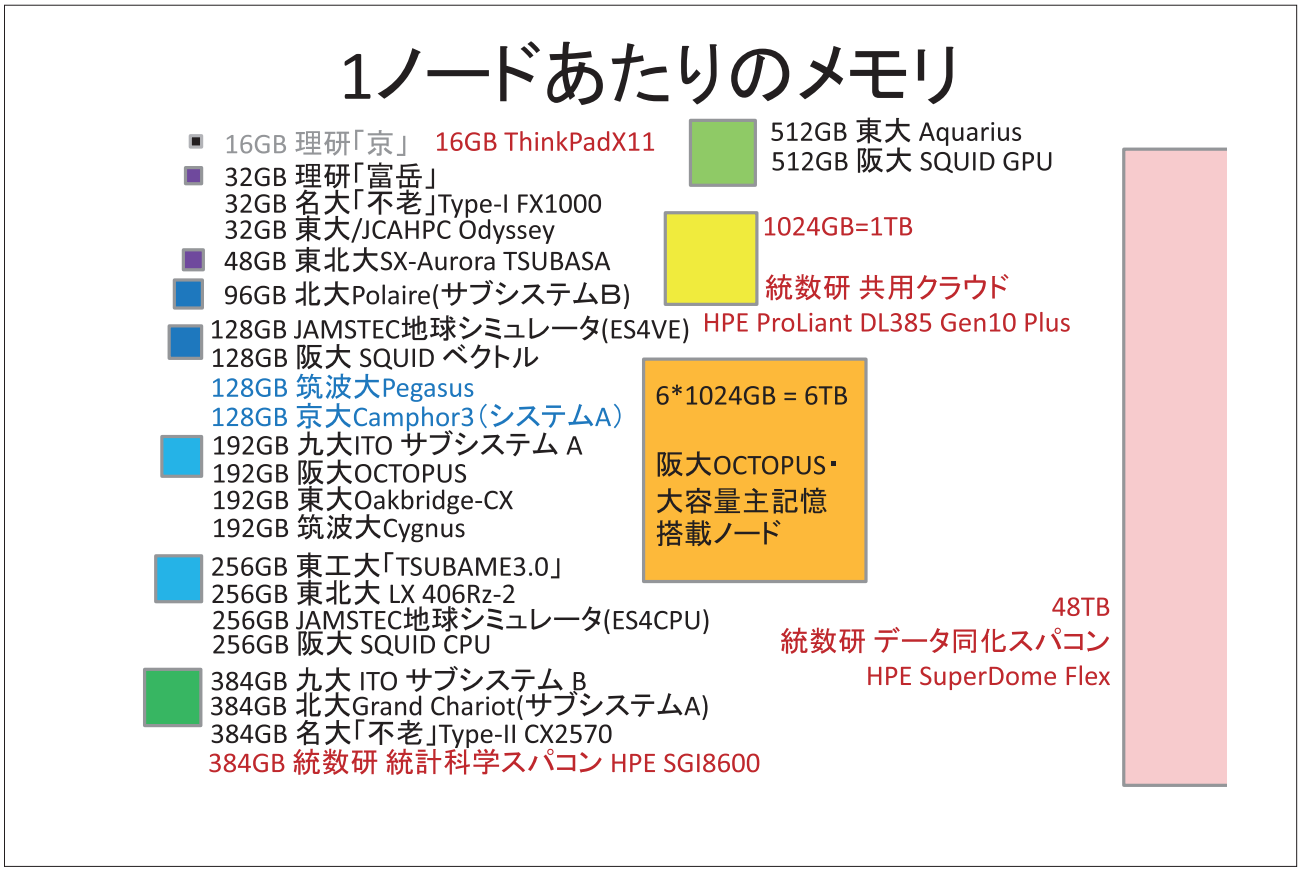
● スライド 11

HPCI: 革新的ハイパフォーマンス・コンピューティング・インフラ, 2012年9月28日～

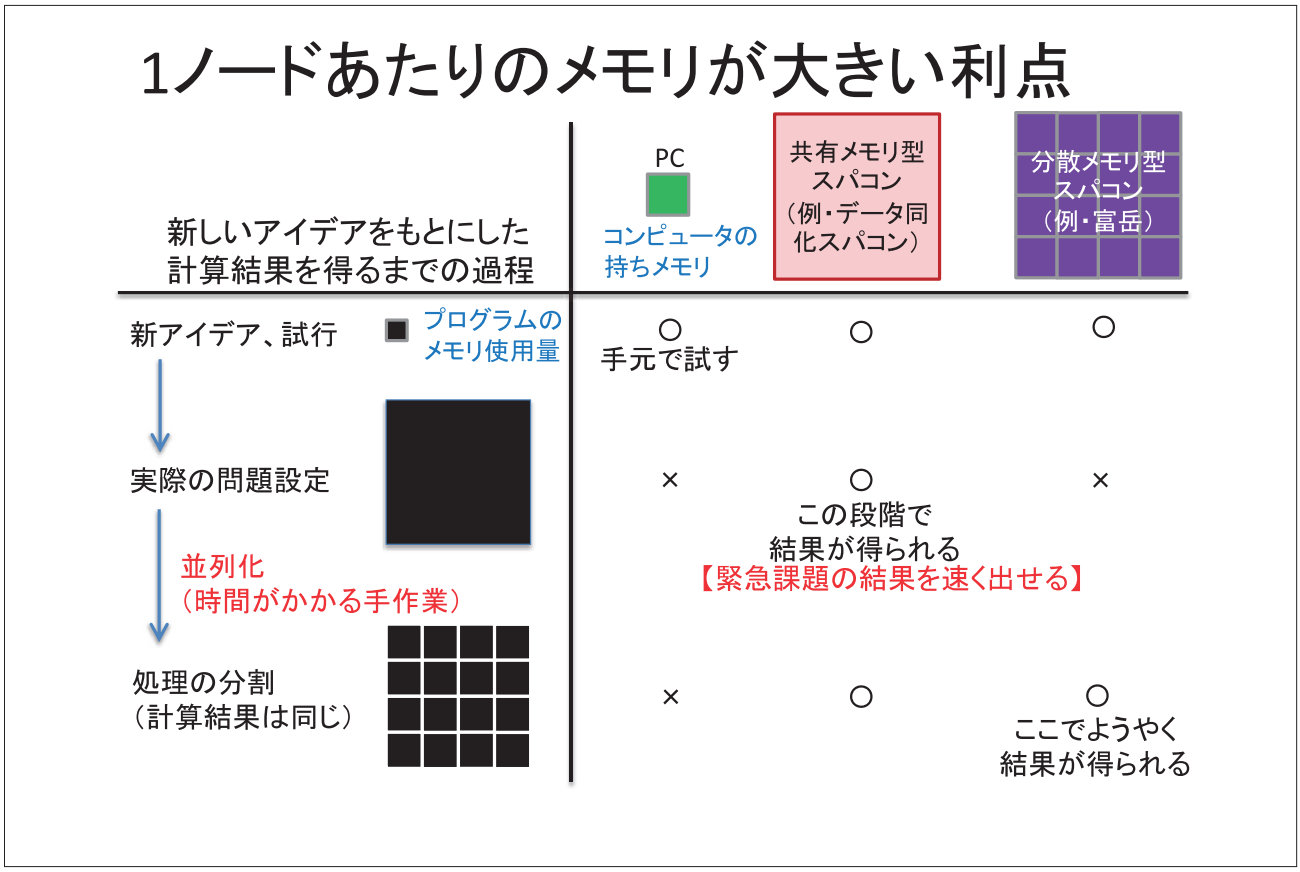
共用スーパーコンピュータ (HPCI)



● スライド 12



● スライド 13



● スライド 14

メモリが大きいことの利点（利用報告書から）

メモリが大きくないと計算できない

手元のプログラムの延長で、大規模のデータ解析が可能に

- 超高次元の超大量データの解析であり、統数研が保有するデータ同化スーパーコンピュータシステムのような世界最大規模の共有メモリ型計算機ではじめて、能率的な解析が可能になった。
- （R言語を用いている）従来の方法で解析に1カ月を要したデータを、0.5時間で解析することが可能になった。

既存の（他人が書いた・代々受け継がれている・長大な）プログラムを読解し、並列化処理を施すのは現実的に困難。既存のプログラムはそのまま動作するメモリが必要

メモリが大きくないと計算が遅い

並列化済みのプログラムでも、ノードあたりのメモリが大きくないと性能が出ない

- 京コンピュータではノードあたりのメモリ不足による問題が生じたため（16GB/node）、期待される性能を得ることができなかった。
- ノードあたりのメモリ容量が多い統計数理研究所所有のUV2000（512GB/node）(*) を使用することで大規模数値流体解析の実現を図る。（*本資料では64TB/nodeの意）

メモリが大きいと並列化もはかどる

実行環境が制限されないよう、メモリが小さい分散メモリ型スパコンでも動作するプログラムを作りたい。並列化の作業には、結果の一致が確認できるため、大きいメモリを持つ計算機が便利

● スライド 15

まとめ

- 統数研・統計科学技術センターでは、大型の計算機を3式運用している

1. 統計科学スーパーコンピュータ (2018-2023)
2. 共用クラウド計算システム (2021-2026)
3. データ同化スーパーコンピュータ (2023-2028)

- 使いやすさを重視し、メモリが大きい構成

● スライド 16

● 統計数理研究所のスーパーコンピュータについて

【上野教授】統計科学技術センターの上野です。オンサイトの NOE の会議では、通常、地下の見学というものがありましたので、それに代わるイベントとしての説明かと思います。

まず、スパコンの定義という話から始めたいと思います。この横軸が時間で、1990 年から 2020 年で、本日というところに縦線が入っていまして、縦軸は、1 秒間の演算回数、FLOPS になっています。ここに青く影を書きましたけれども、これを私は天上界と呼ぶことがあります。この各年で、青いエリアに入ったコンピュータはスーパーコンピュータであるということが、会計上の政府調達のルールで決められています。年々、どんどんバーが高くなっています。それが分かります。皆さんご存じのようなスーパーコンピュータ、京、富岳などは、この辺りに入ります。京は、2011 年あたりから 5 年間か、もうしばらくの運用をして、今、富岳がスタートしています。その上に、さらにアメリカの Frontier というところが、今、世界のスパコン性能ランキング「Top500」のトップにいます。

統数研のスパコンですけれども、このような形であります。この矢印は、何年間運用したかというのですけれども、1995 年から、たいてい 5 年ぐらいのスパンで運用していて、これを見ると、大体、スパコンに入るぎりぎりのところに矢印が引かれています。色が付いてくるのが、2015 年あたりからで、黒と赤と青がありまして、黒は統計科学スーパーコンピュータと呼ばれるもの、赤はデータ同化スーパーコンピュータで、青は、共用クラウド計算システムと呼んでいるものです。現在、本日というところですけれども、そこでは、今、黒が終わろうとしているところと、赤が始まるところで、青がちょうど中間地点のようなところになっています。ですから、現在は、3 台の計算機を運用しているところになります。

見学の時ですけれども、どのくらいの規模かというのを説明する時に、これは、今、Zoom で話している、私のノートパソコンなのですけれども、それとの比較で話を進めています。パソコンの台数を数える時に、通常は、1 台、2 台と数えますけれども、スーパーコンピュータの場合には、1 ノード、2 ノードと数えます。ですから、私の今話している Think Pad は、Think Pad、1 ノードといってもおかしくないものと考えてください。1 台、つまり 1 ノード当たりどのようなスペックですかという時に、いろいろ要素はあるのですけれども、私は、メモリとコア数のみ着目しています。CPU やインターコネクトなど、そのようなことは言わないで、メモリとコアだけを話しています。メモリは、どのようなものかというと、敷地面積に当たるものです。コア数というのは何かというと、テニスコートを何面取れるか、そのような例えとして説明をしています。他の説明の仕方で、メモリは、まな板の大きさであったり、コア数は板前の人数であったり、そのような説明もあります。私の Think Pad は 16 GB で 2 コアですから、テニスコートが 2 面取れます。だから、2 試合同時に実施することができることになります。以降での 16 GB の大きさです。これは、土地面積を、これからスパコンと比べるわけですけれども、この豆粒のサイズになっているということで書いています。これが私のノートパソコンの話です。

1 つ目のスパコンですけれども、統計科学スーパーコンピュータシステムです。2018 年

9 月から運用しているものです。これは緑で、画面は、ほとんどたくさん塗りつぶしていますが、この中には枠が 384 ありますから、384 台、384 ノードになります。1 ノードの中を見てみると、36 個テニスコートが取れます。だから、36 コアになっています。一つ一つのノードのサイズ、一軒一軒の土地の面積は、384 GB となっています。下のほうに、参考、PC、16 GB とありますから、この PC に比べて、金持ちの世帯が 384 世帯あります。それぞれの世帯の中にテニスコートが 36 個あります。そのようなものが固まったものが、統計科学スーパーコンピュータシステムになります。これは、詳しいので飛ばします。

2 つ目です。共用クラウド計算システムです。これは、全 64 ノードになりますから、64 台になります。

先ほどの緑から黄色になった点は、1 ノード当たりの土地面積が広くなったことです。このメモリは、1TB になります。1 TB で 64 コアですから、64 試合が同時にできます。これは、2 台目の計算機として運用しているものです。これを 2021 年の 4 月からです。これは、また詳しいので飛ばしまして、こちらは、AMD です。

3 つ目ですが、3 つ目は、この 4 月、先月から運用を始めたもので、データ同化スーパーコンピュータシステムになります。これは、全 2 ノードですから、2 世帯しかないのですが、それぞれの世帯が大金持ちです。私は大富豪と言っていますが、大富豪が 2 世帯あると、そのようなものになります。持っている土地は、48 TB で、これは、非常に大きいです。中に 896 コアですから、896 個のテニスコートが、それぞれの家にあると、そのようなものになります。

これは飛ばして、大体、この後に計算機室に入るのですけれども、計算機室の配置図ということで、これは上から見て、図を書いています。この黄色の共用クラウド計算システムは、緑の統計科学スーパーコンピュータシステムとして、薄い赤のデータ同化スーパーコンピュータシステムが、地下の計算機室に 3 つあります。これを順番にご覧いただくということになっています。

最近のことなので、写真がありますが、データ同化スパコンの導入風景というところをご覧いただいています。これは、まず左のほうですけれども、地下計算機室に部材が到着したという写真があります。青いシートが敷いてあって、そこに計算ノードがずらずらと並んでいるところです。そして、真ん中の下ですけれども、それから時間を置いて、アメリカから遅れてラックが到着しました。そして、そのラックは、地下計算機室に運ぶ必要があるのですけれども、クレーンで外側から下ろして、地下へ引き込むという手段で入れました。

このような形でスパコンの話をしたのですが、日本全国には、大学などが運営しているスーパーコンピュータが幾つもありますので、比較も交えながら、その紹介を少ししたいと思います。これは共用スーパーコンピュータ HPCI、革新的ハイパフォーマンス・コンピューティング・インフラです。2012 年からとありますが、これは文部科学省が中心になって進めている事業になります。始めた頃は、京コンピュータですけれども、現在は、富岳を中心としたものになっています。中央に日本の地図がありまして、その日本の地図に幾つか各大

学の計算機センターがいろいろと示してあり、見えると思います。この地図の背景に理研富岳と、大きな紫色のものがありますが、これが富岳を表しています。この大きな四角というのは、富岳が持っている全ノード、16 万ノードあるのですが、その 16 万ノードを全部まとめると、このようなサイズになるというものになります。

その周りに、上から、北大、東北大、筑波大、東大、東工大 TSUBAME で、右の下は、JAMSTEC の地球シミュレータです。その左側に、名大の不老、京大。左側は、阪大の SQUID、OCTOPUS、左の上には、九大の ITO システムの A、B というものがあります。統数研データ同化というのが赤くありますけれども、これは、今年度の 10 月から HPCI に提供をする予定にしているシステム、これで 1 ノードになります。この各種、色が緑や紫や水色や赤があるのですが、これは何を示しているかというと、1 ノード当たりのメモリの大きさです。ですから、各世帯が持っている土地の広さになるわけです。

そうなると、各世帯でどのくらい土地を持っているかというのを表にしたくなりましたので、作ってみたのが、この図です。これを見ますと、左の上に、16GB、理研、京、黒い四角がありますが、それは、私が持っている Think Pad と同じサイズの 1 台のメモリはそのくらいになります。それから、1 つ増やしますと、32 GB です。これが、理研の富岳です。富岳は、先ほど巨大でしたけれども、16 万ノードを集めると、あれだけ巨大なシステムになるということが分かります。それから、48、96 といろいろと進めていきまして、統計科学スーパーコンピュータは緑でしたから、それはこの色で言うと緑の 384 GB に当たっています。

それから右側に行きまして、統数研の共用クラウド計算システムは、メモリが 1 TB で、このくらいのサイズになっています。最後、3 つ目の最新のシステムとして紹介したのが、データ同化スーパーコンピュータシステムで、これは、48TB で 1 ノードのシステムになっているものです。

1 ノード当たりのメモリが多いのはなぜかという説明をしたいと思います。これで、新しいアイデアを基にした計算過程を入れます。新しい研究を始めました。新しいアイデアで、簡単なプログラムを作って、実験をしてみましょう。プログラムで、このくらいのメモリが必要とあります。このくらいのサイズでしたら、手持ちの PC であったり、あるいはスパコンでもちろん動くわけです。ですが、次に、実際の問題として、この黒い大きな四角を動かさなければなりません。そのような場合、この緑の PC のサイズでは、とても入らないわけです。この大きさのものをここに入れるというのは無理ですけれども、土地があれば、そこに入れる実験ができます。これでいいことは、この大きなものがそのまま入るので、この段階で、プログラムを実行することができて、緊急課題であっても、結構速く出すことができます。そのようなメリットがあります。

一方、この分散メモリ型です。これは、一軒一軒の土地面積が小さいものをたくさん集めて、場所をつくっている場合には、仕切りがありますから、この大きなものは、そのまま入れることができません。ですから、その場合には、このプログラムを並列化という作業をし

て、時間がかかる手作業ではあります。この黒くて大きいけれども、処理を分割して、プログラムに書き直します。指示行を入れたり、プログラムに手を加えたりすることで、ここで分散メモリ型のスパコンでも結果がでます。ここで得られるので、この 2 段目から 3 段目までのステップをスキップして、結果が出せるという意味で、メモリが大きいと便利であるということです。

メモリが大きいことの利点というのは、利用報告書などで、このように報告がありまして、赤い文字だけ読み上げますけれども、そもそもメモリが大きくないと、計算ができません。2 つ目は、メモリが大きくないと、計算が遅いです。3 つ目が、メモリが大きいと、将来的に並列化をする場合でも、その並列化の作業自体もはかどるということが利用報告書などで報告が上がっています。

まとめですが、統数研の統計科学技術センターでは、大型の計算機、3 式運用しています。1 つ目が、統計科学スーパーコンピュータシステム、2 つ目が、共有クラウド計算システム、3 つ目が、データ同化スーパーコンピュータシステムになります。方針としては、使いやすさ、主に、並列化の手もなるべくかけたくないという使いやすさを重視して、メモリが大きい構成としています。説明は以上となります。

【司会】ありがとうございました。NOE 形成事業そのものではないのですが、それを支援するスパコンについての説明でした。

少し時間が押していますので、次の統計エキスパート人材育成プロジェクトの解説の後に、まとめてご質問やご指摘を頂きたいと思います。次に、千野先生の発表をすぐにつなげていただければ、ありがたいです。よろしくお願いします。

Project for Training Experts in Statistical Sciences

第4回 統計数理研究所 NOE形成事業 顧問会議

統計エキスパート人材育成プロジェクト

大学統計教員育成センター長
特任教授　千野　雅人

統計エキスパート人材育成コンソーシアム
Consortium for training experts in statistical sciences

資料８

第４回NOE形成事業顧問会議

統計エキスパート人材育成プロジェクトの概要

2023年5月16日

大学統計教員育成センター長
千野　雅人

大学共同利用機関法人 情報・システム研究機構
統計数理研究所

● スライド 1

人材育成

大学統計教員育成センター
Center for Training Professors in Statistics

設置目的

- Society 5.0の実現を目指す我が国において，高度な統計学のスキルを有する人材の育成が急務
- 統計エキスパートの最上位に位置する大学統計教員を育成　⇒ 育成された大学統計教員が全国の大学等で中核となり更に統計エキスパートを育成　～ 人材育成の好循環システムの構築を目指す

センター構成

大学統計教員育成センター
- 統括部
- 研修部

2023年4月1日現在

センター長	1名
教授（兼）	3名
助教（兼）	1名
特任教授	13名
特任准教授	1名

千野センター長

✓ 各部に部長を置く
✓ 研修部に研修主幹及び教育システム開発主幹を置く
✓ 各部に，大学統計教員育成センターに所属する研究教育職員を配置する

シニア大学統計教員を特任教授として雇用
⇒ 研修生の育成や教育システムの開発に従事

「統計エキスパート人材育成プロジェクト」の中核機関におけるマネージメントの場として「TESS 運営会議」を定例開催
（Training Experts in Statistical Sciences）

大学統計教員育成センター看板設置式にて
（左から　椿所長，川崎副所長，千野センター長，山下副所長）

● スライド 2

プロジェクト創設の経緯 ～ 危機的な状況にある

科学技術イノベーション総合戦略（2015年6月 閣議決定）

○ 科学技術基本計画の中長期方針の下、各年度の重点項目を明確化
○ 成長戦略の一環として、毎年策定、閣議決定

○ 第5期科学技術基本計画の始動に向けた3つの政策分野
⇒ **大変革時代の挑戦**、地方創生、2020東京オリンピック・パラリンピック

重点的取組

・「**超スマート社会**」の実現に向けた共通基盤技術や**人材の強化**
（IoT、**ビッグデータ解析、数理科学、AI**、サイバーセキュリティ、センサ、ロボット、素材、ナノテクノロジー等）

「**重点的課題**」に、**強烈な問題意識**が示された

第1章 大変革時代における挑戦　2. 重点的課題

「我が国では欧米等と比較し、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く、・・・危機的な状況にある。」

● スライド 3

プロジェクト創設の経緯 ～ AI戦略 2019

AI戦略 2019（2019年6月 統合イノベーション戦略推進会議決定）

○ **3つの理念**　～「人間尊重」、「多様性」、「持続可能」
⇒ Society 5.0 を実現し、SDGs に貢献

○ **4つの戦略目標**　～ ①人材、②産業競争力、③技術体系、④国際
⇒ **AI時代に対応した人材を育成**、持続的に実現する仕組を構築

小中学校	高校	大学	社会人
基礎的学力・情報活用【100万人卒/年】	文理問わず数理・データ関連教育【100万人卒/年】	AI・数理・データサイエンス教育/エキスパート教育【50万人卒/年】	リカレント教育/待遇【多くの社会人に教育機会を提供】
	大学入試 応用基礎を重視する入試に採用する大学への重点支援	**応用基礎** 大学・高専生が自らの専門分野へのDS・AIの応用力を習得（25万人規模/年） ➢ AI×専門のダブルメジャーを可能とする環境 ➢ 専門教育レベルのコース認定の導入	**エキスパート** 年間2000人、トップ100人育成 ➢ PBL中心のAI実践スクール制度 ➢ 若手の海外挑戦機会の拡充

2017年以降、毎年、**データサイエンス系学部**が新設

○ ここで**顕在化した問題**が、**大学の統計教員の不足**

● スライド 4

プロジェクト創設の経緯 ～ 骨太方針・科技基本計画

経済財政運営と改革の基本方針 2020（骨太方針）
（2020年7月　閣議決定）

3. 人・イノベーションへの投資の強化
（1）課題設定・解決力や創造力のある**人材の育成**　②大学改革等
「 データサイエンス教育や統計学に関する専門教員の早期育成体制等を整備する。」

科学技術・イノベーション基本計画（2021年3月　閣議決定）

（1）ｻｲﾊﾞｰ空間とﾌｨｼﾞｶﾙ空間の融合による新たな価値の創出
(a) 現状認識
ﾃﾞｼﾞﾀﾙ社会において、**データが国の豊かさや国際競争力の基盤**
(b) あるべき姿とその実現に向けた方向性
このような社会を支えるのは、人材と社会ｲﾝﾌﾗ。**「数理・ﾃﾞｰﾀｻｲｴﾝｽ・AI」**に関する素養を備え、社会のあらゆる分野で活躍する**人材を大量に育成**
(c) 具体的な取組
② デジタル社会を担う人材育成
統計学の専門教員の早期育成体制を整備

● スライド 5

プロジェクトの今後の取組方針 ～ 統合戦略2022

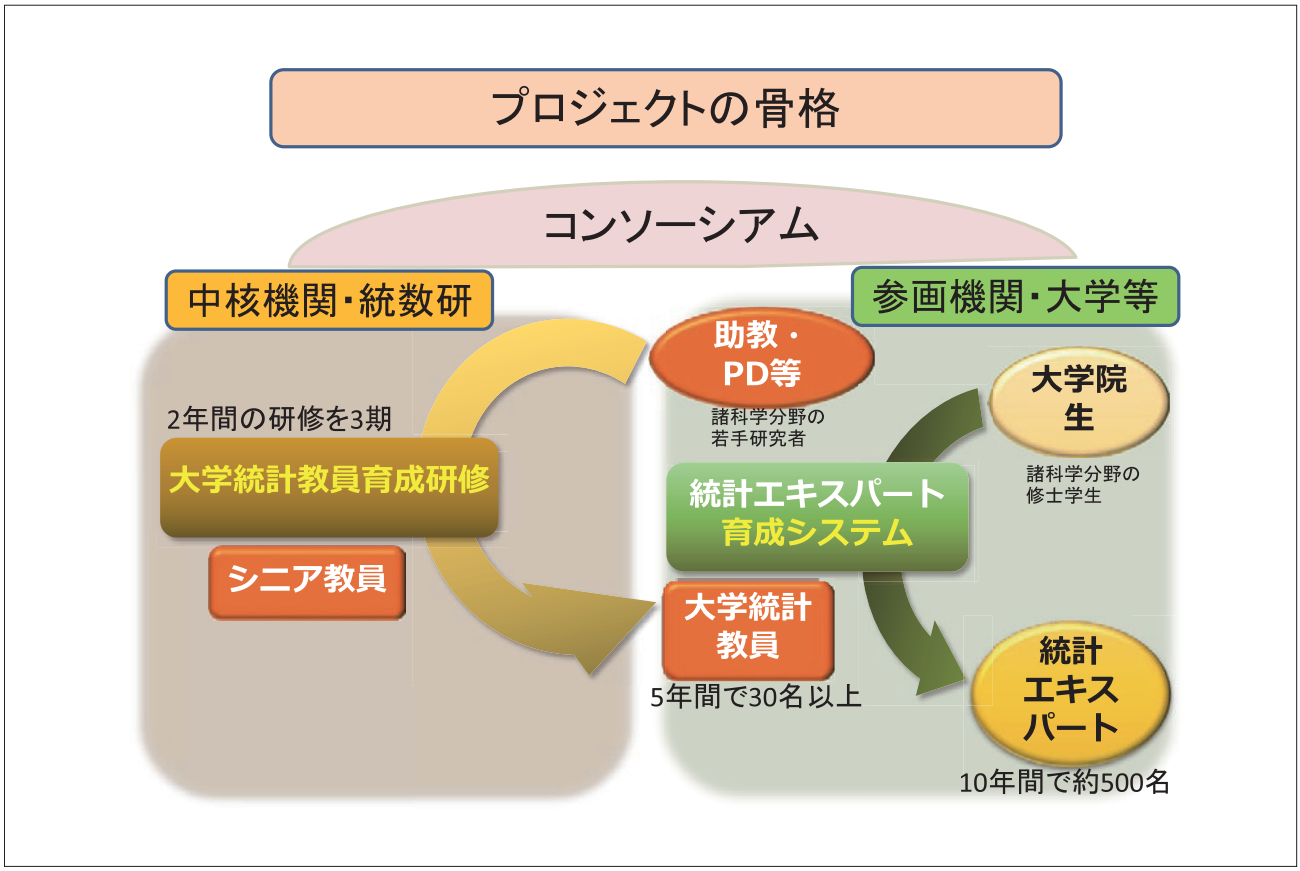
統合イノベーション戦略 2022（2022年6月　閣議決定）

○ 「第6期 科学技術・イノベーション基本計画」（2021年3月 閣議決定）の2年目の年次戦略
○ 「第6期基本計画」の初年度の進捗状況も踏まえ、政策を機動的に見直し、実行することを目指す2022年度の政府計画（閣議決定）

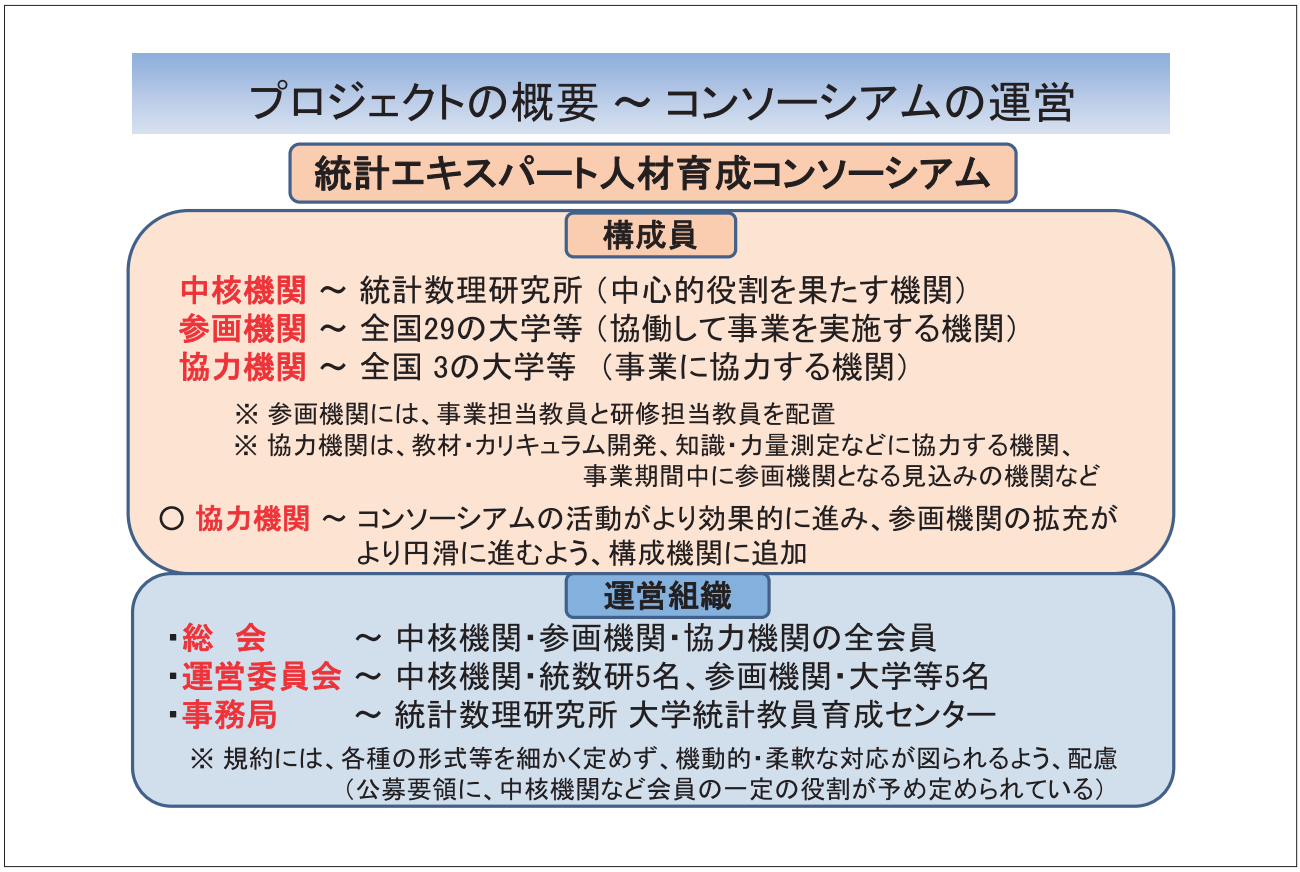
第2章　Society 5.0 の実現に向けた科学技術・イノベーション政策
1. 国民の安全と安心を確保する持続可能で強靭な社会への変革
（1）サイバー空間とフィジカル空間の融合による新たな価値の創出
⑤　デジタル社会を担う人材育成

基本計画・具体的な取組	実施状況・現状分析	今後の取組方針
○デジタル社会を担う人材が輩出・採用され、社会で活躍できるよう、・・・ 2021年度より、大学と政府や産業界等との対話を加速し、統計学の専門教員の早期育成体制整備、・・・	・統計学に関する専門教員の早期育成体制の整備に関して、統計エキスパート人材育成プロジェクトの公募を実施し、大学共同利用機関法人情報・システム研究機構　統計数理研究所を中核機関としたコンソーシアムを選定して事業を開始。	・「統計エキスパート人材育成プロジェクト」を継続的かつ着実に実施。【文】

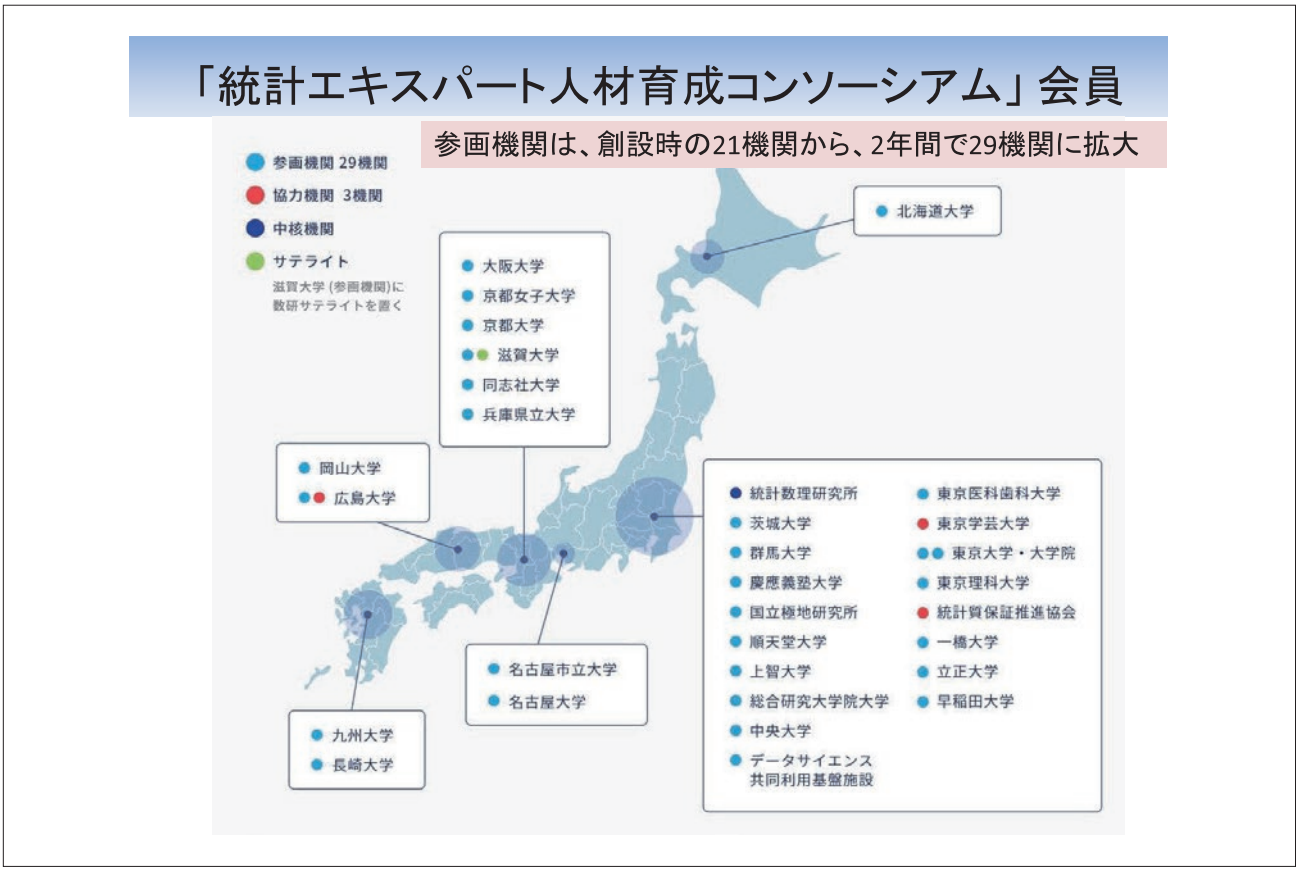
● スライド 6


プロジェクトの骨格
コンソーシアム
中核機関・統数研
参画機関・大学等
2年間の研修を3期
大学統計教員育成研修
シニア教員
助教・PD等
諸科学分野の若手研究者
大学院生
諸科学分野の修士学生
統計エキスパート育成システム
大学統計教員
5年間で30名以上
統計エキスパート
10年間で約500名

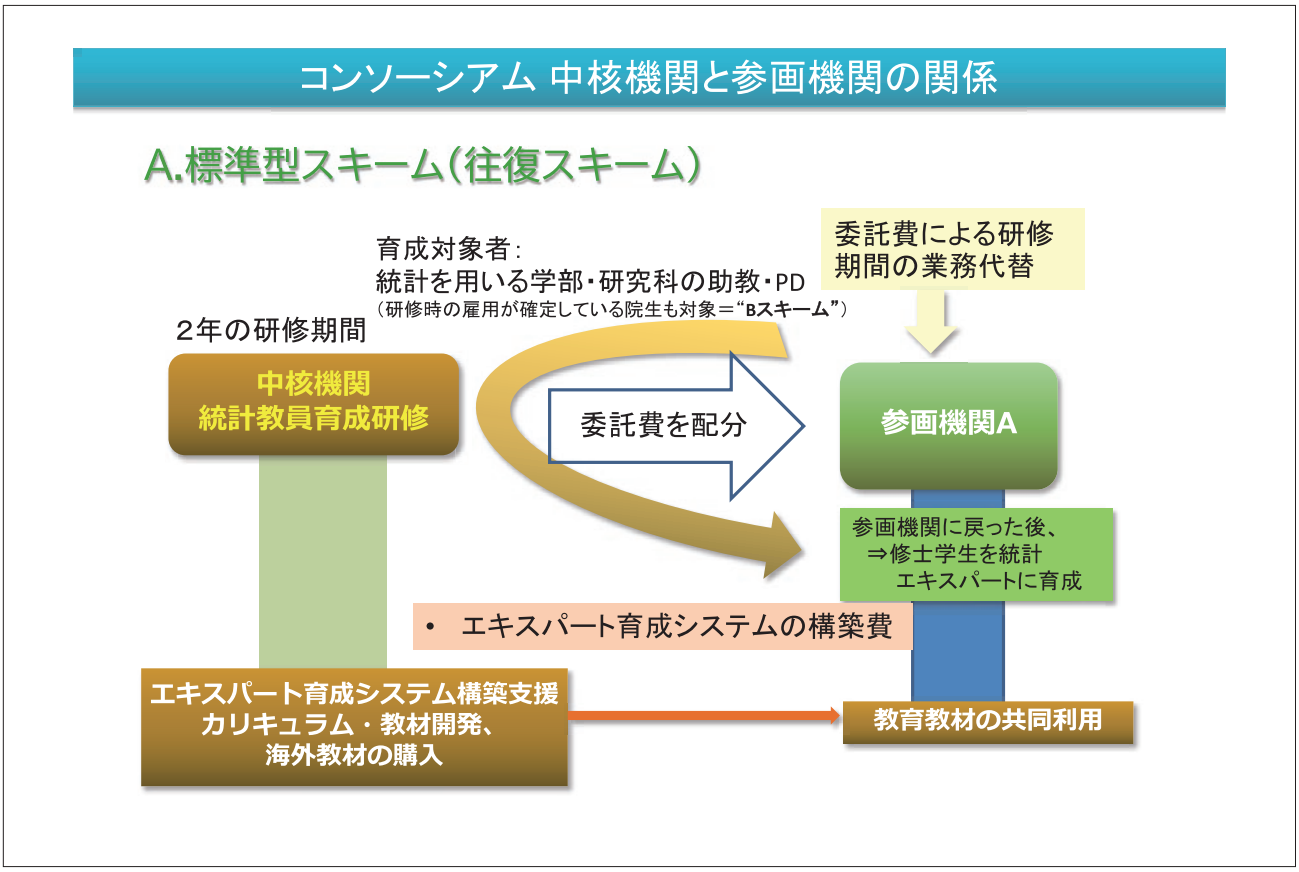
● スライド 7


プロジェクトの概要 ～ コンソーシアムの運営
統計エキスパート人材育成コンソーシアム
構成員
中核機関 ～ 統計数理研究所（中心的役割を果たす機関）
参画機関 ～ 全国29の大学等（協働して事業を実施する機関）
協力機関 ～ 全国 3の大学等 （事業に協力する機関）
※ 参画機関には、事業担当教員と研修担当教員を配置
※ 協力機関は、教材・カリキュラム開発、知識・力量測定などに協力する機関、事業期間中に参画機関となる見込みの機関など
○ 協力機関 ～ コンソーシアムの活動がより効果的に進み、参画機関の拡充がより円滑に進むよう、構成機関に追加
運営組織
・総 会 ～ 中核機関・参画機関・協力機関の全会員
・運営委員会 ～ 中核機関・統数研5名、参画機関・大学等5名
・事務局 ～ 統計数理研究所 大学統計教員育成センター
※ 規約には、各種の形式等を細かく定めず、機動的・柔軟な対応が図られるよう、配慮（公募要領に、中核機関など会員の一定の役割が予め定められている）

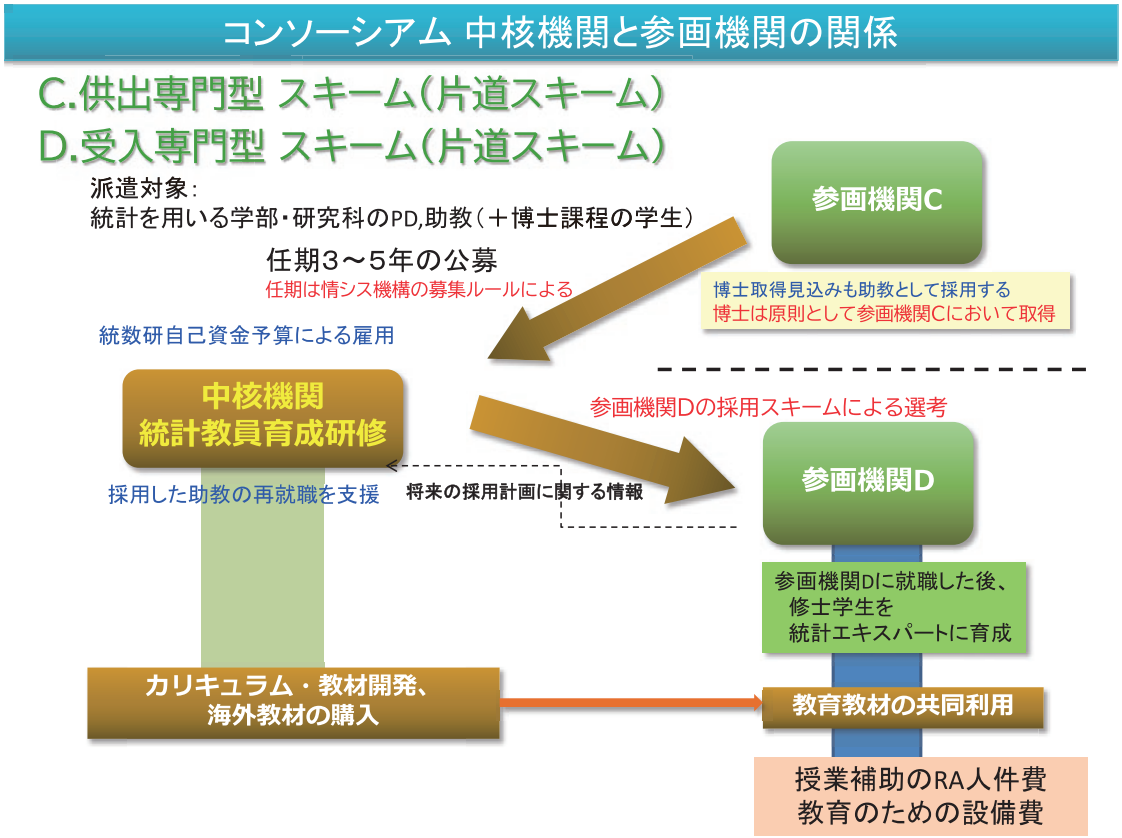
● スライド 8



● スライド 9



● スライド 10



● スライド 11

自己点検の結果概要

【2022年度事業実施計画】

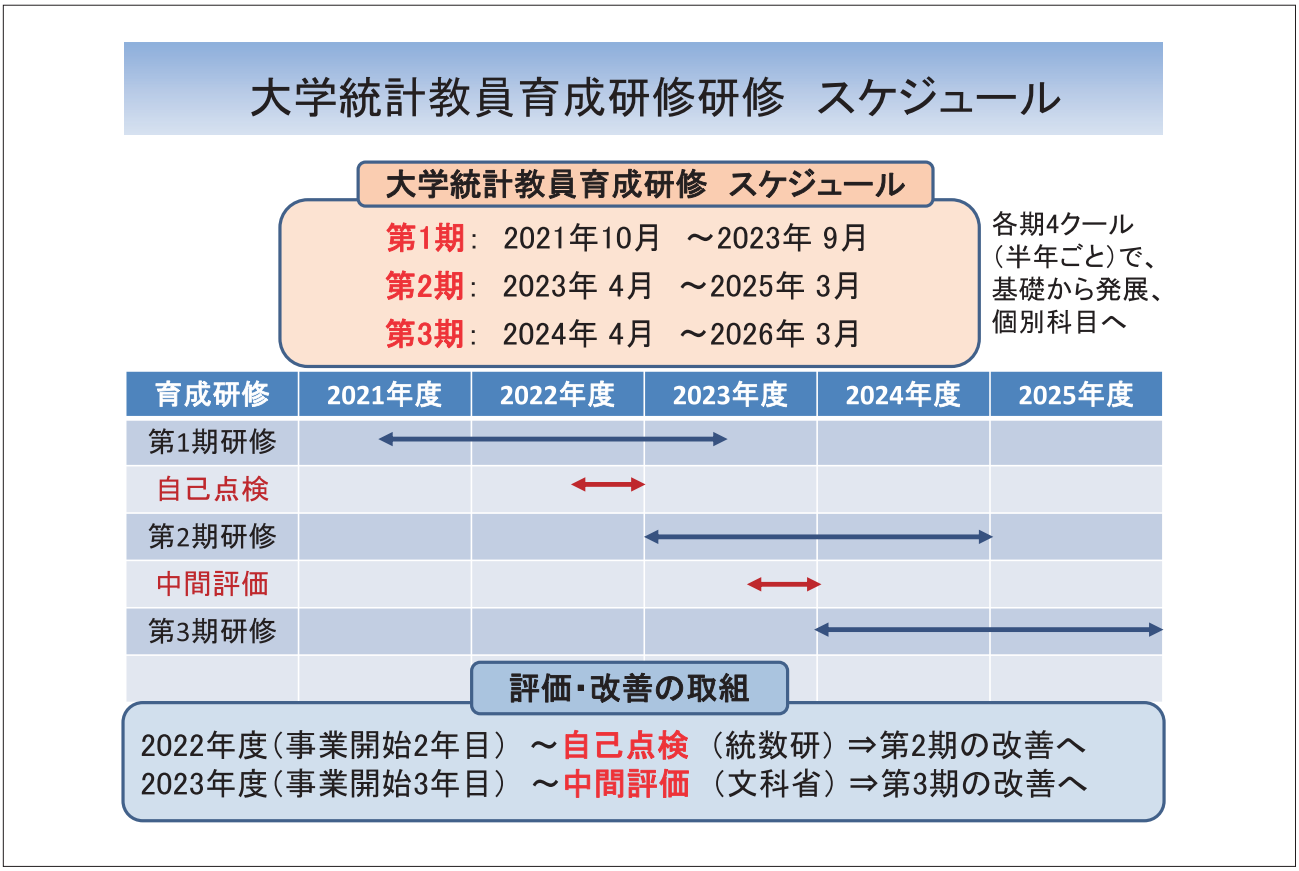
◇ 自己点検を実施（10月）し、第1期大学統計教員育成研修や初年度のコンソーシアム事業の課題の抽出・評価を行った上で、改善方策を検討

◇ この自己点検に当たって、評価アドバイザリーボードを開催し、その助言を活用

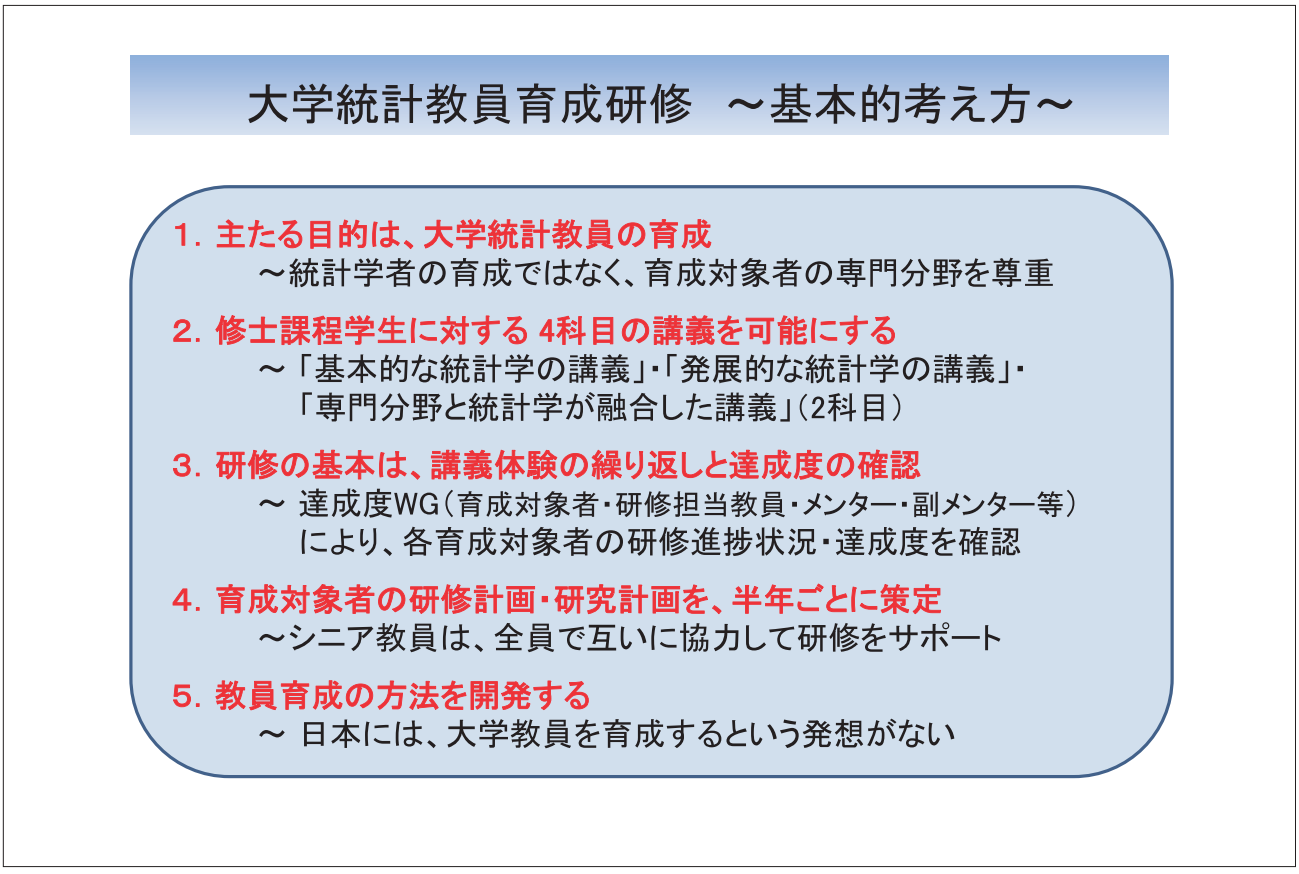
- **中核機関内のTESS運営会議での検討と、運営委員会での議論を踏まえ、アドバイザリーボードを設置し、その助言も活用して自己点検を実施**（10月～令和５年２月）
- 自己点検の一環として、すべての参画機関と育成対象者を対象に、アンケート調査を実施した結果、**「当初の期待を超える活動である」と回答した参画機関が25機関中13機関（52%）、第１期研修参加機関に限ると９機関中８機関（89%）、他の若手研究者に本研修への参加を「勧める」又は「どちらかと言えば勧める」と回答した育成対象者が12名中8名（67%）との評価**
- アンケート調査結果は、第１期大学統計教員育成研修や初年度のコンソーシアム事業の課題の抽出・評価や、今後の改善方策の検討に活用
- **自己点検結果については、「自己点検結果報告書」として取りまとめ、コンソーシアムホームページにおいて公表するとともに、コンソーシアムの2023年度事業計画の検討に活用**

※ アドバイザリーボード
○ 小木 しのぶ（㈱NTTデータ数理システム取締役）【日本計算機統計学会会長】
竹内 光悦（実践女子大学教授）【日本統計学会理事】
山本 義郎（東海大学教授）【日本分類学会評議員】
「○」印：座長

● スライド 12



● スライド 13



● スライド 14

大学統計教員育成研修 ～ 研修科目・カリキュラム

○ 半年間のクールごとに、「**基礎科目**」を全員が受講、「**個別科目**」をできるだけ多く選択して受講。 現状は、ほぼすべてオンラインによる遠隔研修。

基礎科目 ～必須科目

(基本1)「統計ベースライン特習」(4か月)
- 全体研修：全員参加の1時間模擬講義体験（全8回、各3時間）
- グループ研修：メンターごとの研修～統計学の知識向上（月数回、各2時間程度）

(基本2)「統計教育力育成演習」(2か月)
- 全体研修：全員参加の1コマ講義体験（全4～6回、各3時間）
- グループ研修：メンターごとの研修～シラバス・教材作成（月数回、各2時間程度）

(基本3)「統計研究力強化演習」
- 所属参画機関でのFD活動 ～学部生・修士学生に対する講義など

個別科目 ～選択科目

【短期研修】(2～3か月程度)

(個別1) 先端的データサイエンス特論
- 因果推論、時系列解析、多変量解析 などのトピックス
- 各トピックスごとに10回程度・各90分

(個別2) 先端的データ分析演習
- Python・Rの基礎学習、同発展学習 などのトピックス
- 各トピックスごとに5～6回程度・各90分

【長期研修】(1年以上の期間が必要)

(個別3) コンサルテーション演習
- 官庁・企業等からの依頼に応じたコンサルテーション

(個別4) 教材開発演習
- 所属参画機関用の教材等を作成

(個別5) 共同研究演習
- 他分野の研修生等との交流を通じて行う共同研究

● スライド 15

大学統計教員育成研修 ～ 育成対象者とシニア教員

第1期研修 育成対象者（12名）とメンター（6名）の配置

参画機関	学　位	メンター	副メンター
東京医科歯科大学	博士（工学）	岩崎	折笠
兵庫県立大学	博士（情報学）	中西	岩崎
同志社大学	博士（文化情報学）	中西	岩崎
群馬大学	博士（保健学）	岩崎	折笠
岡山大学	博士（経済学）	中西	岩崎
東京理科大学	博士（薬学）	水田	中西
東京理科大学	博士（情報科学）	国友	岩崎
滋賀大学	博士（工学）	神保	中西
滋賀大学	博士（工学）	神保	中西
長崎大学	博士（理学）	折笠	岩崎
国立極地研究所	博士（情報）	水田	岩崎
統計数理研究所（C方式）	博士（経済学）	国友	中西

メンター（シニア教員）
岩崎　学、折笠秀樹、国友直人、神保雅一、中西寛子、水田正弘

● スライド 16

大学統計教員育成研修 ～ 育成対象者とシニア教員

第2期研修 育成予定者（13名）とメンター（8名）の配置

参画機関	学　位	メンター	副メンター
北海道大学	博士（保健科学）	三輪	水田
早稲田大学	博士（理学）	水田	三輪
順天堂大学	博士（医学）	岩崎	折笠
名古屋市立大学	博士（医学）	折笠	岩崎
一橋大学	修士（学際情報学）	清水	水田
中央大学	博士（数理科学）	清水	水田
一橋大学	博士（工学）	国友	岩崎
北海道大学	博士（経済学）	国友	岩崎
滋賀大学	博士（情報学）	神保	中西
京都大学	博士（情報学）	中西	神保
九州大学	博士（理学）	三輪	水田
東京理科大学	博士（薬学）	折笠	岩崎
同志社大学	博士（文学）	中西	神保

メンター（シニア教員）～岩崎　学、折笠秀樹、国友直人、神保雅一、中西寛子、水田正弘、清水邦夫、三輪哲久

● スライド 17

２０２２年度 活動概要

■ **2022年度の活動では、「第２期研修の企画・準備」及び「自己点検の実施」を追加**

2021年度

① コンソーシアムの運営
② 第1期教育システムの開発
③ 第１期研修の実施
④ 必要な環境整備

2022年度

① コンソーシアムの運営
② 第1期教育システムの開発
③ 第１期研修の実施
④ 第2期研修の企画・準備
⑤ 自己点検の実施
⑥ 必要な環境整備

■ **2022年度の主なスケジュール**

5月１日		第１期大学統計教員育成研修・**第２クール開始**
5月中旬	第２回 コンソーシアム**総会**	
5月下旬～	2022年度参画機関等との業務委託契約事務	統計検定2級の受験チケット配布（予定）
7月下旬	コンソーシアム運営委員会	8月　第２期研修・C方式育成対象者応募締切
8月下旬	第２回**コンソーシアムワークショップ**	9月　統計関連学会等での成果発表
9月上旬		第２期研修・A方式育成対象者推薦締切
10月1日～	10～2月　**自己点検の実施**	第1期研修・**第３クール開始**
10～11月		第２期研修・A方式・コンソーシアム選考委員会
12月中旬		第２期研修・C方式・統計数理研究所による選考
1月中旬	コンソーシアム運営委員会	

● スライド 18

第1期大学統計教員育成研修の概要　～2021年度

■ **2021年度の主なスケジュール**

21年9月	第１期生　説明会（17日　統計数理研究所）
10月	第１期生　開講式（４日　統計数理研究所）
10月～	第１期生　大学統計教員育成研修･第１クール開始
10～11月	第1回　全体研修（模擬講義）
12月～1月	第2回　全体研修（模擬講義）
３月～4月	第1回　全体研修（公開模擬講義）
～4月	第１期生　大学統計教員育成研修･第１クール修了

全体研修　３人が発表（各自60分）×４日
互いの模擬講義を見ることで教育方法を学ぶ
Zoomを利用，全て録画，参加できないときは後で録画をみる

グループ研修　全体研修のない週等に行うメンターとの研修
統計に関する知識を確かなものとする
各自の研究を進める

● スライド 19

第1期大学統計教員育成研修の概要　～2022年度以降

■ **2022年度以降の主なスケジュール**

５月～	第１期生　大学統計教員育成研修･第２クール開始	
5月～6月	先端的データサイエンス特論（統計的因果推論、統計的時系列解析）	コンサルテーション演習
6月～7月	第3回　全体研修（模擬講義）	
８月	統数研公開講座受講（多変量解析）	教材開発演習
9月	統計関連学会　全員発表（6日　成蹊大学）	共同研究演習
9月～10月	先端的データ分析演習（R環境）	
9月～10月	第2回　全体研修（公開模擬講義）	
10月～	第１期生　大学統計教員育成研修･第３クール開始	
10月	先端的データサイエンス特論（統計学実践とその数理）	
11月	中間報告会（4日　滋賀大学）	
12月	第4回　全体研修（模擬講義）	
23年１月	先端的データサイエンス特論（医療統計と研究公正）	
３月	第3回　全体研修（公開模擬講義）	
３月	中間報告会（24日　長崎大学）	
４月～	第１期生　大学統計教員育成研修･第4クール開始	
5月～6月	先端的データサイエンス特論（連続講義ｼﾘｰｽﾞ、計算ベイズ）	
9月	第4回　全体研修（公開模擬講義）	

● スライド 20

● 統計エキスパート人材育成プロジェクトについて

【千野特任教授】大学統計教員育成センター長の千野です。よろしくお願いします。それでは、統計エキスパート人材育成プロジェクトについて説明します。このプロジェクトは文科省の公募事業で、統数研が受託することができました。

そのプロジェクトを実施するために、昨年 1 月、統数研に新たな研究施設として、大学統計教員育成センターを設置しました。特任教員が 10 名以上いる組織になります。プロジェクトの経緯なのですが、2015 年 6 月の閣議決定での重点的課題というところに強烈な問題意識が示されました。それは何かというと、統計科学を専攻する人材が極めて少なく、危機的な状況にあるということです。これが政府の閣議決定です。

その後、2019 年、AI 戦略 2019 です。ここでデータサイエンス AI の人材育成ということが掲げられて、この前後、2017 年の滋賀大学データサイエンス学部以降、毎年、データサイエンス系学部、あるいは学科を新設する大学が現れまして、特に最近、加速化しています。

ここで顕在化した問題が、大学の統計教員、統計をきちんと教えられる教員の不足ということでした。そのようなことを受けまして、2020 年の骨太方針、それから、科学技術基本計画、いずれも閣議決定ですが、ここにおいて、統計学の専門教員の早期育成体制を整備するということが記載されました。これを受けて、文科省で予算要求をして、2021 年公募事業となったということです。

このプロジェクト開始後も、2022 年、昨年の最近の閣議決定の中でも、政府としての記述で、今後の取り組み方針として、このプロジェクトを継続的かつ着実に実施しますと、【文】と書いてありますので文科省が書いた文言ですが、このように位置付けられています。

プロジェクトの内容ですが、これが骨格です。中核機関が統数研、参画機関が全国の大学等です。ここでコンソーシアムをつくりまして、この中で人材を育成します。どのように行うかというと、中核機関、統数研が 2 年間の大学統計教員を育成する研修を 3 期にわたって実施します。コンソーシアムの参画機関の大学等は、ここにいろいろな分野の若手研究者を参加させます。2 年間たつと、大学統計教員としての能力を持つことになります。この方と共に、所属の大学では、統計エキスパートを育成する仕組みを作っていただいて、所属の大学院生、これも諸科学分野の修士学生なのですが、そのような方々をいろいろな分野で活躍できる統計エキスパートに育成するということです。統数研が教員を育成し、育成された各大学の教員が統計エキスパートを育成するという 2 段階の仕組みになっていまして、このサイクルがうまく回りだせば、人材育成の好循環システムが構築されるだろうということを目指しています。

コンソーシアムですが、中核機関は、統数研、それから、参画機関は、現在、29 の大学になっています。北は北海道大学から、南は九州、長崎大学まで、このような大学が参加しています。なお、2 年前の創設時には、これが 21 機関でしたので、2 年間で 29 機関に拡大したということで、われわれは、このプロジェクトに対するニーズの強さを感じています。

少し細かい話ですが、このスキームとしまして、研修期間の研修生の業務を代替する教員を雇用できるような委託費を参画機関には配分しています。それから、片道だけ、研修生を送りだすだけ、育成された教員を受け入れるだけという大学の参加も認めています。

昨年度、自己点検を実施しました。これに当たっては、第三者的なアドバイザリーボードの助言をいただいて、また、全ての参画機関、それから、育成対象者にアンケート調査を実施しました。そのアンケート調査の結果、いずれも高評価を頂きました。ちなみに、参画機関で高く評価していただいたのは、25 機関中 13 機関、52%です。それほど多くないようなのですが、まだ第 1 期が終わっただけですので、研修に参加していない参画機関のほうが多いです。そのようなところは、評価保留ということにしていましたので、研修に参加した機関に限ると、9 機関中、8 機関が、当初の期待を超える活動であると高く評価しています。この自己点検結果報告書は、紙の報告書としてもまとめたのですが、ホームページで公開しています。

これが、研修のスケジュールなのですが、3 期ありまして、第 1 期は、21 年の 10 月開始で、今年の 9 月に終了になります。第 2 期が、今年の 4 月に始まったところで 2 年間、第 3 期は、来年 4 月に開始して 2 年間ということになっています。この初期の段階に自己点検を行って、5 年間の真ん中のところで文科省の中間評価があります。今年度は、中間評価の年になっています。

研修の考え方なのですが、われわれが育成するのは統計学者ではなくて、いろいろな専門分野の若手研究者を、統計も教えられる教員に育成するということで、専門分野を尊重しようという考え方で研修を実施しています。具体的には、4 科目の講義を可能にするということで、基本的な統計学、発展的な統計学、専門分野と融合した講義 2 科目という考え方になっています。

これは、具体的なカリキュラムですが、必須科目と選択科目となっていまして、必須科目のほうは、統計学の基礎知識の履修、それから、教育力の育成、研究力の強化という形になっています。実際の研修ですが、模擬講義という形で行っていまして、研修生が教師役になって、1 コマ講義をします。他の研修生が学生役になって、いろいろと質問するという講義体験の繰り返しという中で、統計の基礎知識、教育力を付けるようにしています。選択科目は、ご覧のようなものになっています。

これが、今年 9 月に終了する 1 期生なのですが、12 名です。第 1 期は、メンター6 名体制で実施しました。専門分野ですが、このように、情報学、保健学、薬学等いろいろな分野にわたっています。第 2 期は、今年 4 月開始ですが、研修生 13 名で、第 1 期、第 2 期、重なる時期もありますので、メンターを 2 名増加させまして、8 名体制で研修を行っています。専門分野はやはりいろいろな分野、医学、それから、経済学や、初めて文学という方もいらっしゃいますが、このような幅広い分野になっています。

あとは活動概要、実際、昨年度どのような活動をしたかということですが、年 1 回、コンソーシアムの総会を行っています。それから、8 月にワークショップを行って、10 月から 2

月、自己点検を実施しました。研修のほうは、第 1 期研修が実施中で、併せて第 2 期研修の選考を行いました。このような形で進めています。ご参考までに、21 年度、これは初年度のスケジュール、それから、22 年度以降のスケジュールという形になっています。私からの説明は以上です。

【司会】千野先生、ありがとうございました。先ほどのスーパーコンピュータに関するものと、このエキスパート人材育成プログラムに関する発表で、顧問の先生方、何かコメント、ご質問がありましたら、よろしくお願いします。

【蒲地顧問】よろしいですか。

【司会】蒲地先生、よろしくお願いします。

【蒲地顧問】上野さんの発表を聞いて、見学会の代わりということで、申し訳ないのですが、データ同化用のスパコンということで、非常に大きなメモリを持った汎用（はんよう）機というより、専用機のようなのを導入されて、非常にうらやましい限りです。データ同化研究をやっているところは、例えば私が以前働いていた気象研究所でも本当は、そのようなものがあったら、非常に楽だろうなという気持ちがあって、うらやましいなと思いながら聞いていました。

質問は、全く違う質問で申し訳ないのですけれども、省かれたところで、少し気が付いたのは、GPU タイプも導入されていますよね。機械学習などそのようなグループがよく使っているのは、どちらかというと、CPU タイプよりも GPU タイプのコンピュータのほうを使っています。ソフトウエアを使ったりはしているのですけれども、そのような時の CPU から GPU への全く全部書き換えるというのは、とても大変な作業で、予算も人も要るだろうと思います。何か GPU タイプを使用する時の機械学習グループが使用するような時に、何かサポートなど、そのようなものは、スパコンのセンターというか、そのような関係者としてはされているのでしょうか。

【上野教授】実は GPU は、そこまで積極的に入れていません。

【蒲地顧問】そうですか。

【上野教授】はい。統計科学スーパーコンピュータには、GPU ノードが少数あるのですが。

【蒲地顧問】ありましたね。

【上野教授】以降の共有クラウドや、データ同化スパコンでは GPU は入れていません。

【蒲地顧問】入れていないのですか。

【上野教授】はい。これをユーザーとして従来のデータ同化は、元の力学モデルを基準とするというところがあるので、まずフォートランが動かないと、ということがあるので。

【蒲地顧問】フォートランですね。

【上野教授】そこの現在のユーザーからすると、スパコンで GPU を使う人は、あまりいないという状況があります。ですから GPU は、これまでも発表があったセンターのほうで特別な何か使用というか、使い方で準備して、グループで使っているという現状になっていま

別な何か使用というか、使い方で準備して、グループで使っているという現状になっています。

【蒲地顧問】そうですか。分かりました。ありがとうございます。

【司会】ありがとうございました。他、いかがでしょうか。よろしいでしょうか。最後に、またまとめて質問いただくつもりでいますので、次に進ませていただきます。

Survey Science NOE

第4回 統計数理研究所 NOE形成事業 顧問会議

調査科学NOE

活動紹介・事業計画

データ科学研究系 調査科学グループ
准教授　前田　忠彦

資料５

2023年5月16日　統計数理研究所NOE形成事業顧問会議

調査科学NOE活動紹介

―「日本人の国民性調査」を中心に―

調査科学NOE
データ科学研究系・調査科学グループ
グループリーダー　前田　忠彦

● スライド 1

本日の話題提供の概要

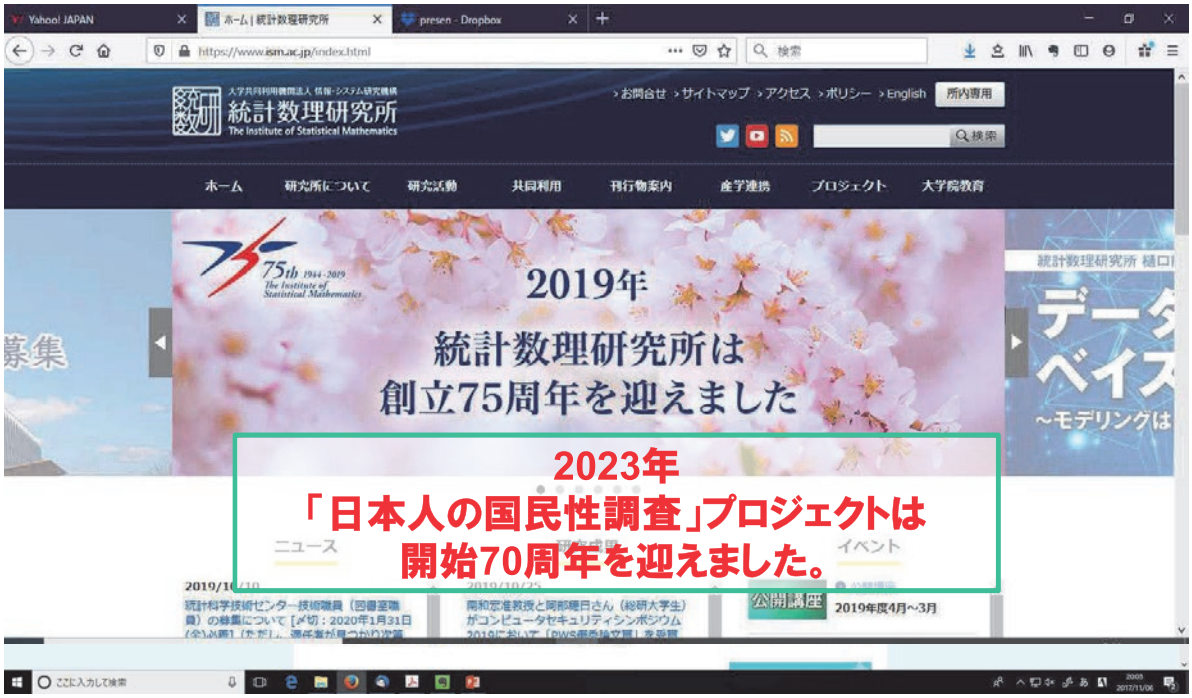
主たる活動である「日本人の国民性調査」に関わる話題を中心に調査科学NOEのこれまでの活動と，今後の活動計画（短期的・中長期的課題）を紹介します。

統計数理研究所の社会調査グループでは1953年以来，5年毎に「日本人の国民性調査」という社会調査を継続実施してきました。

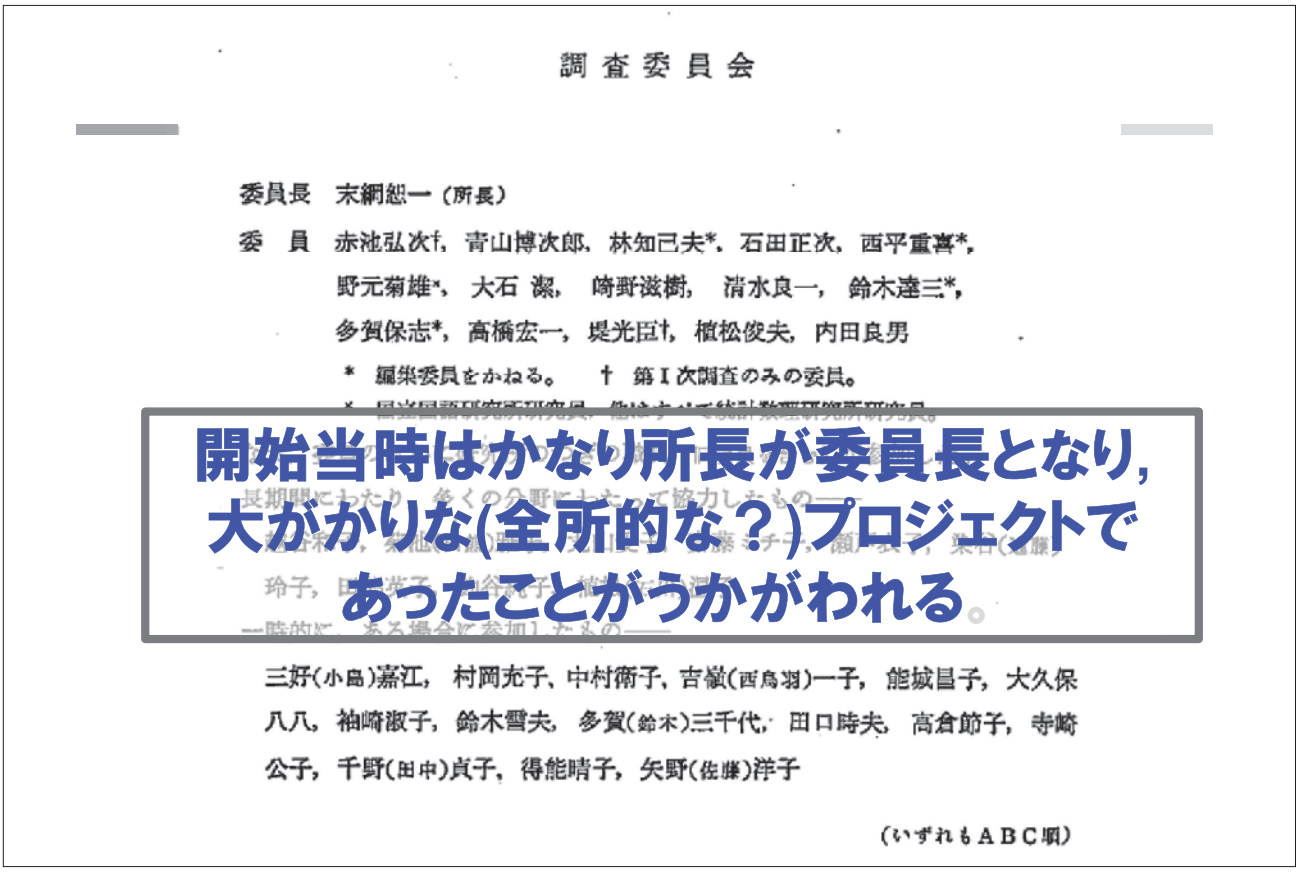
この伝統的なプロジェクトも岐路を迎えているかも知れません。具体的な調査結果や関連研究を少し紹介しつつ【伝統を守りながらも，新しい姿で（社会）調査の分野のNOE形成に果たしていくべき役割】について考えています。

● スライド 2

研究所のHP（http://www.ism.ac.jp）4年前



● スライド 3



● スライド 4

初期の調査委員会

委員長　末綱恕一（所長）

委　員　赤池弘次†，青山博次郎，林知己夫*，石田正次，西平重喜*，
野元菊雄*，大石 潔，崎野滋樹，清水良一，鈴木達三*，
多賀保志*，高橋宏一，堤光臣†，檜松俊夫，内田良男

*　編集委員をかねる。　†　第Ⅰ次調査のみの委員。

×　国立国語研究所研究員，他はすべて統計数理研究所研究員。

- 今から思えば，少し意外な名前も含まれている
- 全所的なプロジェクトであったばかりでなく，国立国語研究所との共同作業の跡も見える
  - 1948年「日本人の読み書き能力調査」
  - 1950年「鶴岡市における言語調査」

● スライド 5

調査科学NOE/調査科学グループ

- 統計数理研究所の70年以上にわたる社会調査研究の成果を基盤として，そのさらなる発展と社会的貢献の促進のために活動するグループ
  - 社会調査法に関する理論や応用を研究し，NOE（ネットワークオブエクセレンス：国内外の同様の研究機関や組織との間の連携を通じて，研究の発展を目指すための仕組み）活動行う
- いくつかのプロジェクトを展開
  - 日本人の国民性調査
  - 意識の国際比較調査
  - 社会調査情報集積プロジェクト
  - 連携研修調査プロジェクト
  - 社会調査情報活用プロジェクト
- なぜ統計学の研究所で社会調査の研究・NOE活動を行う？
  - 社会調査は，統計学に支えられた重要な応用分野の一つであるから

● スライド 6

調査科学NOE　令和5年度事業計画

調査科学グループ人員配置

2023年4月1日現在

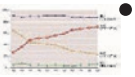
- 所内
  - 准教授　前田忠彦(*1, グループリーダー)
  - 朴　堯星
  - 特任助教　芝井清久 (*1)
  - 特任研究員　田中康裕(*1)
- 客員教員　9 名

※DS共同利用基盤施設社会データ構造化センターと密接に連携して活動
(*1)DS社会データ構造化センターが本務

研究会・シンポジウム等開催計画

（計画なし）
- 令和5年度は「日本人の国民性　第15次全国調査」への注力のため，大規模研究集会会・シンポジウム等の開催は予定しない。
  - 客員教員を中心とした，国民性調査データの活用法に関する内部研究会等を実施予定
  - オンラインチュートリアルセミナーの実施可能性検討中

主要プロジェクト紹介



● 日本人の国民性調査プロジェクト（PL：前田忠彦）
1953年より5年に一度継続実施している「日本人の国民性調査」(直近は平成30年度の第14次全国調査)を通じて，日本人の国民性の解明，調査法研究，データ解析法の開発，という3つの目標に資する研究を行う。今年度は秋に第15次全国調査を予定する。

● 国際連携調査プロジェクト（PL：朴　堯星）:
主に東アジア地域を念頭に，MOU締結先の機関等と連携し，既存の国際比較調査データの解析を通じた研究交流を進めると共に，国際比較調査の企画・推進に協力する。今年度は上記第15次全国調査に合わせ，協定先成均館（ソンギュングァン）大学主催の韓国総合社会調査KGSSと一部項目を共通化して実施し，国際比較研究を行う。

● 連携研修調査プロジェクト（PL：加藤直子(客員)）
大学等他機関と連携し，統計数理研究所が実施する社会調査において，学生・研究者等に実際の社会調査のプロセスを実地に体験する機会（連係研修調査）を提供し，社会調査における人材育成に資する。令和2年度以降は共同研究型調査の実施のための関居作りや，その調査設計支援の側面を強化している。

● 社会調査情報活用プロジェクト（PL：稲垣佑典(客員)）
統計数理研究所が保有するさまざまな社会調査データを活用して，新たな統計分析の手法の開発、調査方法論の研究や統計教育の素材とする等の活動を推進することを通じて統計科学の発展に寄与することを目的とする。更に，公的ミクロ統計と社会調査データを連結した分析など，調査データの新たな活用方法を検討する。今年度はR4年度実施のモード間比較調査のデータを活用した統計エキスパート人材養成プロジェクト用の教材開発を目指す。

● 意識の国際比較調査プロジェクト（PL：芝井清久）
→DS社会データ構造化センターへ機能移転・縮小中
アジア・太平洋価値観国際比較調査など，「日本人の国民性」調査をCLA(文化の連鎖的比較)やCULMAN(文化多様体解析)という国際比較の枠組みに拡大して、各国の国民性を多次元的に総合的に解析する。
ＤＳ施設社会データ構造化センターは，本プロジェクトからの資産を継承し，既存国際比較調査データの公開と共同利用に資する活動を推進する。

※調査(共同調査含む)の実施については,資金獲得が前提となる活動が多く、資金獲得を目指した外部連携の強化を調査科学NOEの中心的活動目標とする。

● スライド 7

別紙

調査科学NOE　令和5年度事業計画

調査科学グループ客員教員一覧

2023年4月1日現在

【継続】
・客員教授
吉川　徹（大阪大学）
佐藤 嘉倫（京都先端科学大学）
松本　渉（関西大学）
今田 高俊(東京工業大学名誉教授)
真鍋　一史(関西学院大学名誉教授)
・客員准教授
藤田 泰昌（長崎大学）
尾碕 幸謙（筑波大学）

【新規】
・客員准教授教授
稲垣 佑典（成城大学）
加藤 直子（桜美林大学）

調査科学グループ年度NOE活動に係る目標

- 「日本人の国民性調査」プロジェクト，
  - ✓ 第15次全国調査(秋実施)の準備・実査・基礎分析までを進める
  - ✓ 開発中の疑似音声面接型調査モードによる実験調査を並行実施
- 国際連携調査プロジェクト
  - ✓ 第15次国民性調査と韓国総合社会調査の連携(項目提供と国際比較研究)
  - ✓ R4年度に引き続きネットワーク形成の方針策定を検討。
- 「意識の国際比較調査プロジェクト」プロジェクト
  - ✓ R4年度実施ウェブ国際比較調査結果の公表
- 「連携研修調査プロジェクト」
  - ✓ DS施設との役割分担を踏まえて，位置づけを微修正する(継続検討)。
  - ✓ 機構内予算などの取得を通じた「共同研究型調査」の強化を図る。
- 「社会調査情報活用」プロジェクト
  - ✓ 社会調査データ(特に前年度実施のモード間比較調査)を通じた統計教育教材の開発の検討

その他

【協定機関との交流】
＜国内＞　特になし
＜海外＞
- 協定先成均館(ソンギュングァン)大学調査科学研究センターとの研究交流(2回程度の先方への訪問を予定):
  - KGSS韓国総合社会調査への(6月実施予定)との連携と共同分析

【広報計画】
- 日本人の国民性調査関係のウェブサイトの再整備。

【その他】
- 令和5年度は,「日本人の国民性　第15次全国調査」(5年ぶりの実施)を中心に活動計画を立案し,主に2つの関連研究を行う。→ ※1参照
  - 関連研究1:音声疑似対面型調査の実験調査
  - 関連研究2:KGSSへの項目提案と共同分析
- 選挙人名簿登録者数・投票区性格調査(マスコミ4社との協同調査:例年実施案件)

※1　資金源として下記外部資金も活用
科研費基盤研究(A)　R5～R8年度
日本人の国民性の統計的研究:継承と発展
（代表　前田忠彦）
科研費基盤研究(C)　R3～R6
Web会議アプリとAIアシスタントを用いたインタラクティブ調査方法の開発(代表　朴 堯星)

● スライド 8

「日本人の国民性調査」とは

- 統計数理研究所が1953（昭和28）年以来5年に1度実施している「継続社会調査」 ↓入口ページ

https://www.ism.ac.jp/ism_info_j/kokuminsei.html

● スライド 9

第14次全国調査の主な結果
https://www.ism.ac.jp/survey/index_ks14.html

1. 伝統的意見の退潮
2. 緩やかに進む「ドライ志向」
3. 持ち直し安定する経済面の評価
4. 伸び悩む環境保護に対する意識
5. 安定した「日本人の性格」に関する自己評価
6. 社会生活上の不安や原子力施設の事故に対する不安の変化
7. 新規項目に関する結果

● スライド 10

「日本人の国民性調査」の3つの目的

なぜ統計数理研究所は，かくも長きにわたって，この研究を続けてきたか？

＝下記の3つの目的を達成するため

1. 国民性の解明：調査結果をとおして，日本人のものの見方や考え方とその変化を明らかにしていくこと
2. 調査手法の研究開発：実際の調査を行いながら，これからの社会変化にも対応できる新たな統計調査手法を研究していくこと。
3. 統計手法の研究開発：調査のデータを素材として，より優れた統計的方法を研究していくこと。

● スライド 11

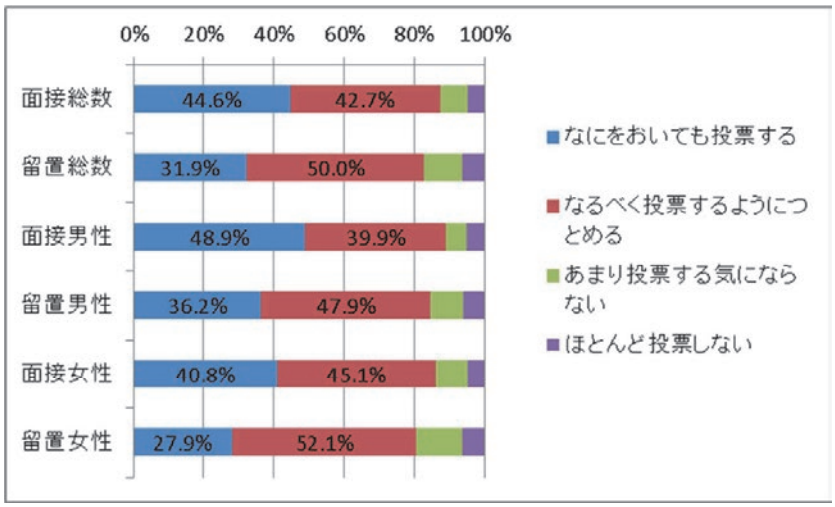
継続社会調査で何が分かるか？

- 「日本人の国民性調査」；本体調査は，5年に1度の実施
  - 継続調査として：調査方法（サンプリングの方法や調査実施の方法：実際には面接法）と調査項目などをなるべく「同一に」して，調査を繰り返す（同じといっても限界はあるが）
  - 古くからの質問項目の他に，新しい国民の意見の動向を知るために，新規の項目を導入することに。
  - 1973年の第5次調査からは，調査票をK型とM型の2種類用いている：M型は主に新規に導入した項目で構成される
- なぜ同じ調査方法・同じ調査項目で「継続」してきたのか？
  - 「日本人の意見の変化」を知るために
  - 調査の結果は，調査方法によっても異なり，質問の表現のちょっとした変更でも変わってしまうことがある
    →同じ方法・項目で継続しないと，変化がわからない。
- 2023年第15次全国調査：基本は従来と同じ手法で今秋実施準備中

● スライド 12

調査方法によって結果が変わってしまう例

“#8.6　選挙への関心”
あなたは衆議院の総選挙があるとき、ふつうはどうしますか？



留置法よりも面接法によるほうが，投票意向がかなり高めに回答される
（統計数理研究所による2012年度中の実施調査から）

● スライド 13

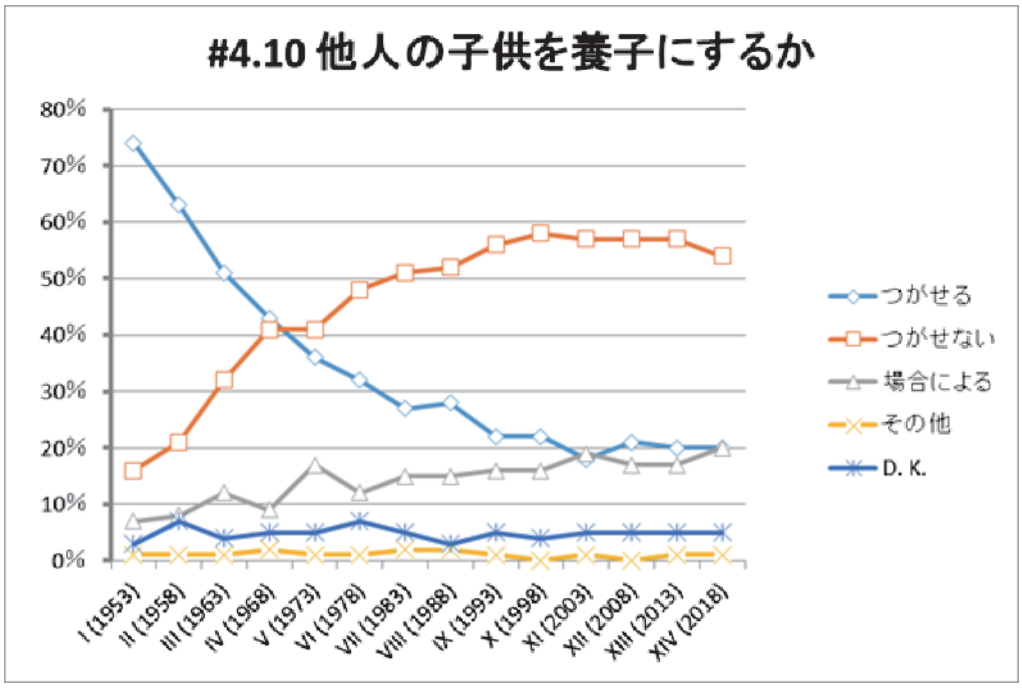
長期にわたる日本人の意見の変化の例（1）
#4.10　他人の子供を養子にするか？

質問文と選択肢
#4.10 他人の子供を養子にするか
　子供がないときは、たとえ血のつながりがない他人の子供でも養子にもらって家をつがせた方がよいと思いますか、それともつがせる必要はないと思いますか？

1　つがせた方がよい
2　つがせないでもよい、意味がない
3　場合による
4　その他［記入］　　5 D.K.

● スライド 14

長期にわたる日本人の意見の変化の例（1）
#4.10　他人の子供を養子にするか？



● スライド 15

長期にわたる日本人の意見の変化の例（2）
#5.6　めんどうをみる課長

質問文と選択肢

#5.6 めんどうをみる課長

[リスト]ある会社につぎのような2人の課長がいます。もしあなたが使われるとしたら、どちらの課長につかわれる方がよいと思いますか、どちらか一つあげて下さい？

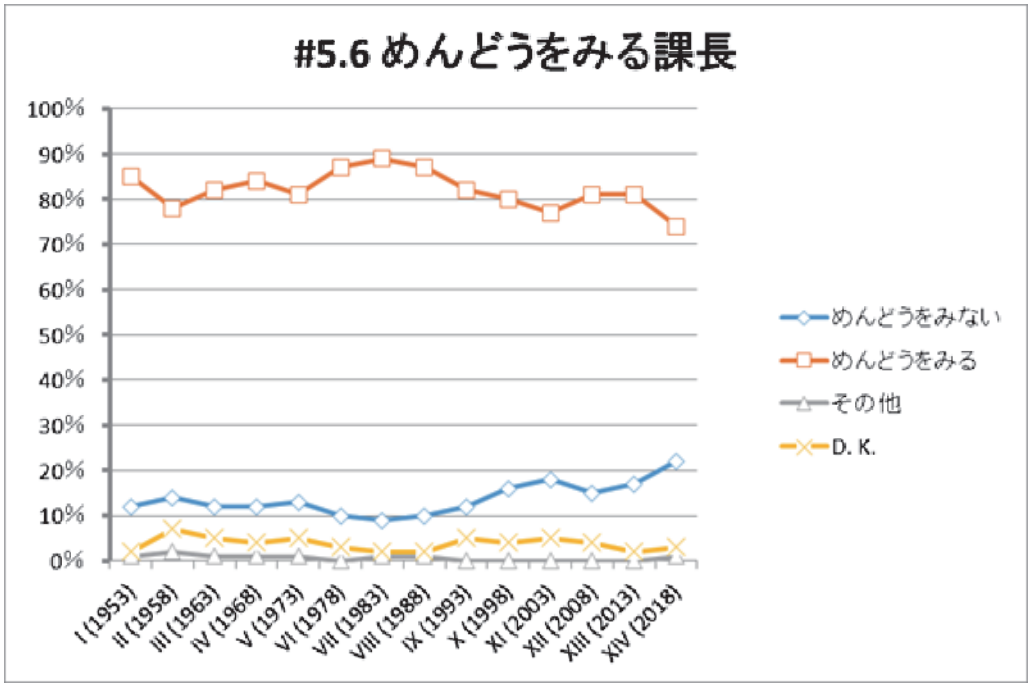
1　規則をまげてまで、無理な仕事をさせることはありませんが、仕事以外のことでは人のめんどうを見ません

2　時には規則をまげて、無理な仕事をさせることもありますが、仕事のこと以外でも人のめんどうをよく見ます

3　その他[記入]　　4 D.K.

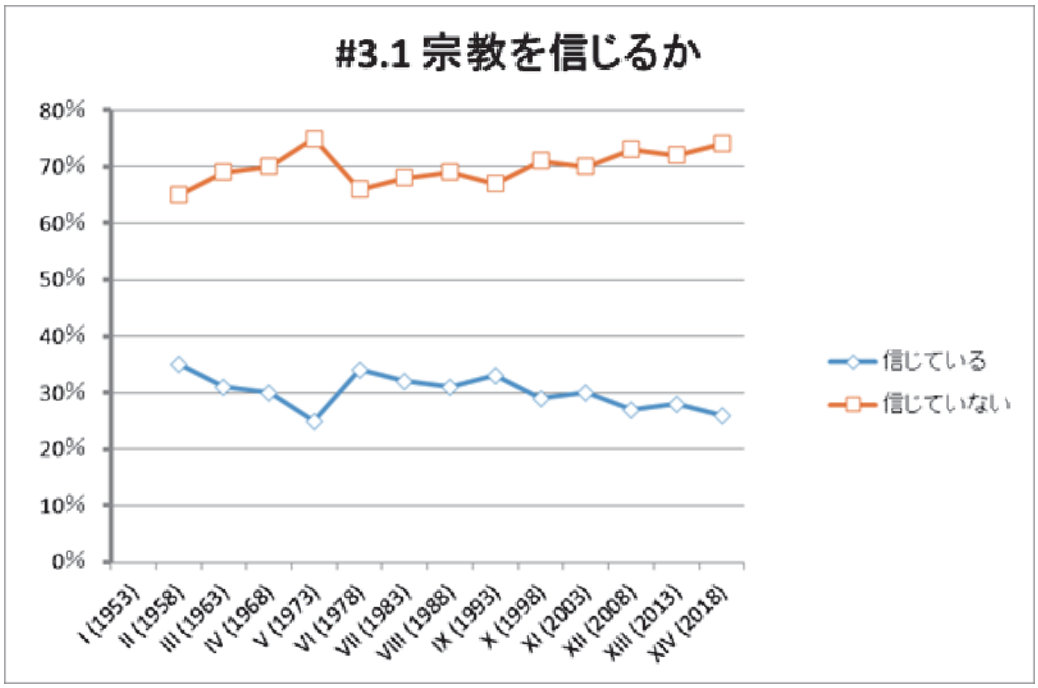
● スライド 16

長期にわたる日本人の意見の変化の例（2）
#5.6　めんどうをみる課長



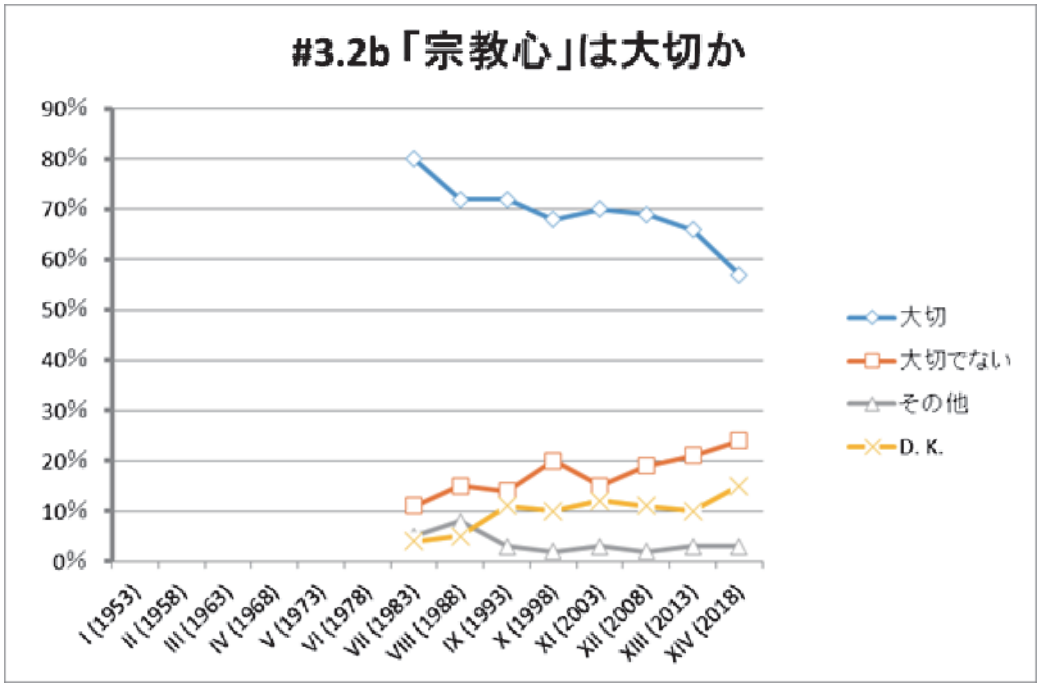
● スライド 17

長期にわたる日本人の意見の変化の例（3）
日本人の宗教心：宗教を信じるか？



● スライド 18

日本人の宗教心：宗教心は大切か？



● スライド 19

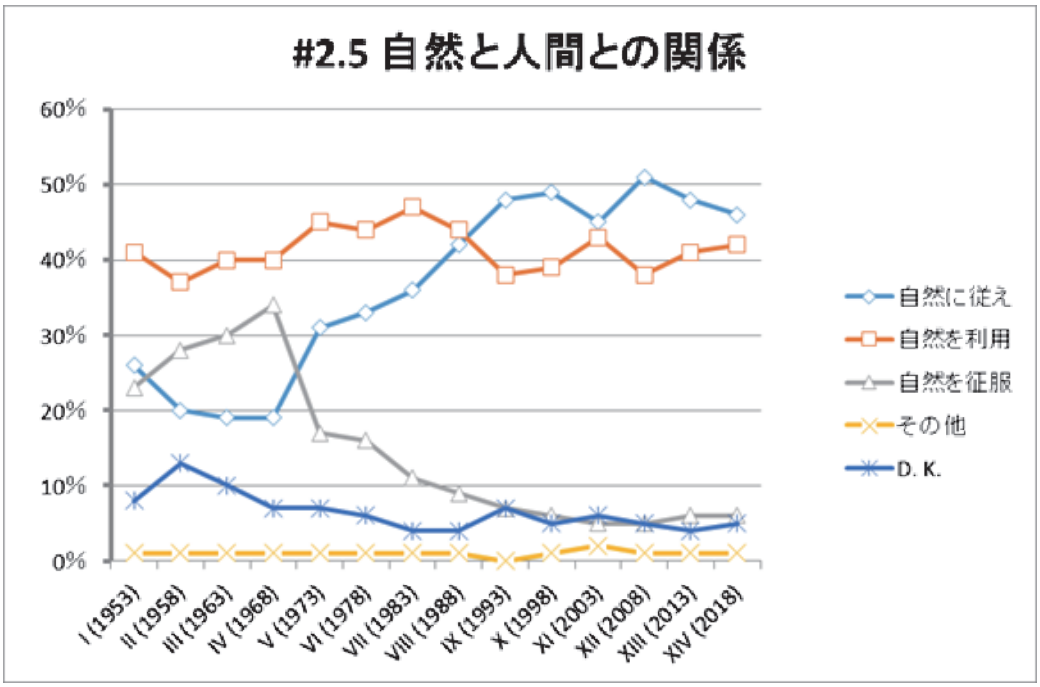
自然をめぐる意識　#2.5 自然と人間との関係

[リスト]自然と人間との関係について、つぎのような意見があります。あなたがこのうち真実に近い（ほんとうのことに近い）と思うものを、ひとつだけえらんで下さい？

1. 人間が幸福になるためには、自然に従わなければならない
2. 人間が幸福になるためには、自然を利用しなければならない
3. 人間が幸福になるためには、自然を征服してゆかなければならない
4. その他[記入]　　5 D.K.

● スライド 20

自然をめぐる意識　#2.5 自然と人間との関係



● スライド 21

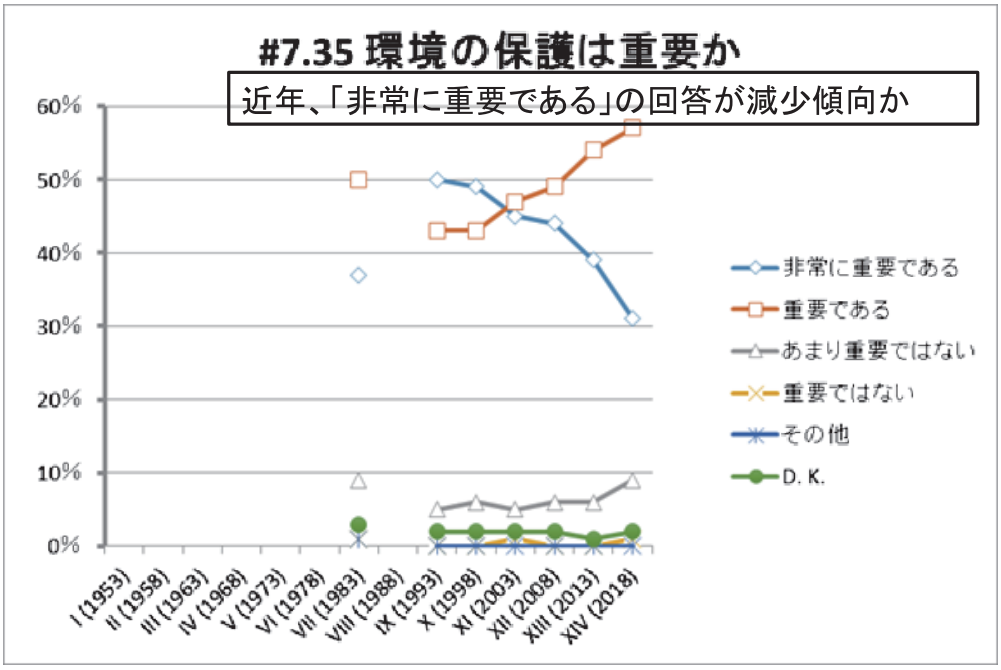
環境をめぐる意識　#7.35 環境の保護は重要か

〔リスト〕環境の保護は、あなたにとってどのくらい重要な問題ですか？

1. 非常に重要である
2. 重要である
3. あまり重要ではない
4. 重要ではない
5. その他［記入］　　6 D.K.

● スライド 22

環境をめぐる意識　#7.35 環境の保護は重要か



● スライド 23

政治に関する項目　#8.6 選挙への関心

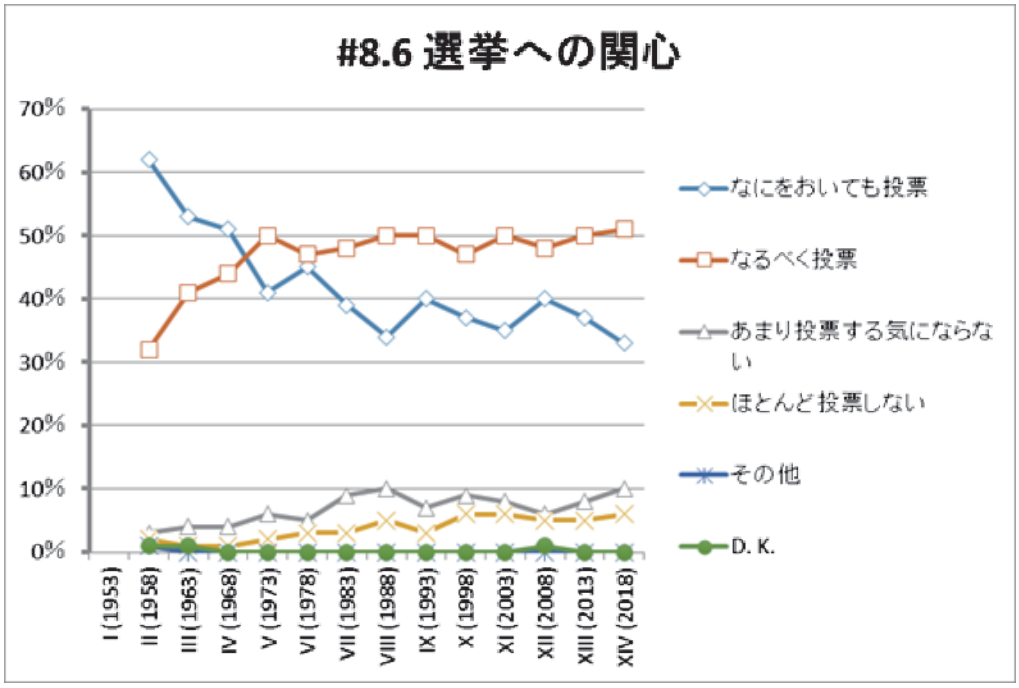
#8.6 選挙への関心(1958年より調査)

[リスト]あなたは衆議院の総選挙があるとき、ふつうはどうしますか？

1　なにをおいても投票する
2　なるべく投票するようにつとめる
3　あまり投票する気にならない
4　ほとんど投票しない
5　その他[記入]　　6 D.K.

● スライド 24

政治に関する項目　#8.6 選挙への関心



● スライド 25

意見の変化はなぜ起こるのか？

こうした意見の変化はなぜおこるのか？

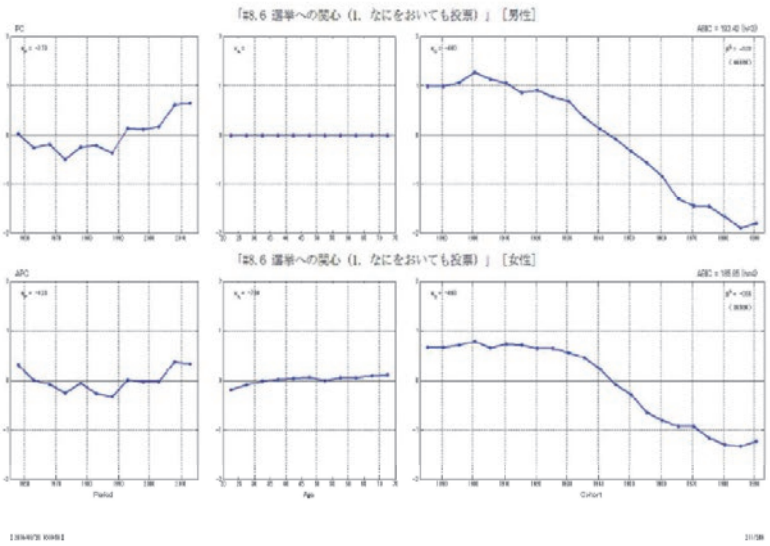
- 時代の変化と共に全ての人が意見を変えていく（時勢の効果：変化は比較的大きく）
- 新しい意見を持った世代が調査対象に参入してくる（出生コウホートの効果（または世代効果）：変化は緩やか）
- 加齢によって（個人の中でも）変化が生じていく場合（加齢効果：変化はあまり生じない，社会の高齢化に伴いわずかずつ）

⇒長期にわたる継続調査では，こうした要因の効果が見えてくる

● スライド 26

“なにをおいても投票”に対するコウホート分析

コウホート効果が優勢



OB 中村 隆名誉教授提供

● スライド 27

国民性調査関連の近年の大型科研費取得

- アジア・太平洋価値観国際比較--文化多様体の統計科学的解析
  2010-14年度　基盤研究(S)　代表 吉野諒三
  　アジア・太平洋地域11カ国・地域の国際比較調査
- 日本人の価値意識の変容に関する統計的研究
  2012-15年度　基盤研究(A)　代表 中村　隆
  　第13次全国調査(2013年)の側方支援，モード間比較，パネル調査
- 計量的日本人研究の新展開
  2018-21年度　基盤研究(A)　代表 前田 忠彦
  　第14次全国調査(2018年)の基盤
- 日本人の国民性の統計的研究：継承と発展
  2023-26年度　基盤研究(A)　代表 前田 忠彦
  　第15次全国調査(2023年)の基盤，伝統の継承と次代を見据えた発展

● スライド 28

国民性調査の課題と第15次調査での対応

【手法面】

- 回収率の長期的な低下傾向
  →調査不能バイアスの研究，回収率維持の方策
- 面接調査法の危機（長期的な問題とコロナ禍後の調査と）
  →他の調査モードによる実験を含ませる（Push to Web，疑似音声型面接）

【内容面】

- 他の（標準的な）社会調査と比べた時の変数の“偏り”
  →属性変数(Demographic Variables)の充実を図る
- 「日本人」の研究で，日本人の国民性が分かるか？
  →近隣の（ある程度文化を共有する）国との協力による国際比較の導入：手始めに韓国一般社会調査（KGSS）との項目共有
- 調査項目（内容面）での疲弊
  →第14次調査からの新規項目の継続と発展

● スライド 29

調査科学NOE近年の研究テーマ例(1/2)（国民性調査プロジェクトを除く調査方法論を中心に）

- 階層意識全国調査の時系列データの収集と標本抽出WEB調査法の確立 [ネットワーク型]
  科研費基盤研究(A) 2019- 2023年度 代表 吉川 徹 分担 前田忠彦 他.
- 調査票調査におけるセンシティブな質問への回答に対する調査員の影響の計量的研究 [ネットワーク型]
  科研費基盤(C) 2020-23年度 代表 小林 大祐 分担 前田忠彦
- 郵送調査法の回収率向上方策 [基盤研究型]
  e.g. “Improving mail survey response rates in Japan: empirical tests for envelopes, request letters, questionnaires, and schedules”, (Yoosung Park, Takahiro Tsuchiya (2020), International Journal of Social Research Methodology 24(1).
- Web会議アプリとAIアシスタントを用いたインタラクティブ調査方法の開発 [基盤研究型]
  科研費基盤(C) 2021-24年度 代表 朴 堯星，分担 前田 忠彦，持橋 大地

● スライド 30

調査科学NOE近年の研究テーマ例(2/2)
(その他の共同研究，科研分担者等)

- 心理実験を取り入れた先進的Web調査におけるサティスファイスの実態解明
  科研費基盤研究(C) 2022-24年度　代表稲垣 佑典(成城大学)
- 「調査協力態度尺度」構築による各種社会調査データの質の評価
  科研費基盤研究(B) 2022-25年度 代表 吉野諒三(同志社大学)
- 社会的危機状況下における人びとの意識の変容とその階層差に関する社会学的解明
  科研費基盤研究(B) 2021-23年度　代表 数土 直紀(一橋大学)
- 難解な感染症関連用語の言い換えや説明の案出と理解促進効果の検証
  科研費基盤研究(C) 2021-23年度　代表 横山詔一(国立国語研究所)
- 基礎教育を保障する社会の基盤となる日本語リテラシー調査の開発に向けた学際的研究
  科研費基盤研究(A) 2019年-21年度　代表 野山 広(国立国語研究所)
- 「不適切回答アラートシステムの開発」
  筑波大学ビジネスサイエンス系との共同研究 2022年8月 - 2023年10月

@自己評価：いろいろな課題に協力しているが
個別の共同研究を超えた「ネットワークハブ」的な役割を果たせているか？

● スライド 31

調査科学NOEの（長期的）課題と対応策(?)

- 人員の減少 (承継職員 2012年 6名→2023年 2名)
  →若手人材のリクルート
- ネットワーク形成型活動の強化
  →（ハブ的組織との連携 e.g. 大学共同利用機関法人の他機構間，同機構内DS施設の活動の共有，共同研究の戦略的活用）
- 新しいデータサイエンスの潮流下，“ビッグデータ”時代下での社会調査の役割？
  →伝統の継承と「発想の転換」の双方が必要
  e.g. 高質少量データと低質大量データの統合的活用
  まだなお「少量データ」を質高く取得するための技法の鍛錬は必要
  –1988年第8次調査までは，「自前」で調査を実施していた：この体制は望めないにしても
- 国際性の担保(調査科学に関わる”ガラパゴス化”の回避)
  →構成メンバーの努力による，戦略的国際連携
- 発信の不足
  →成果の書籍化など

● スライド 32

● 調査科学 NOE 活動紹介・事業計画について

【司会】続きまして、3 つ目の NOE ですが、調査科学 NOE の活動紹介・事業計画について、これは所内にセンターがありませんので、所内の肩書では、データ科学研究系調査科学グループの前田准教授より説明させていただきます。前田先生、お願いします。

【前田准教授】調査科学グループの前田と申します。今日は、よろしくお願いします。本日の話題ですけれども、日本人の国民性調査という話題を中心に、調査科学 NOE のこれまでの活動と、今後の計画についてご紹介します。私どものグループでは、1953 年以来、5 年ごとに日本人の国民性調査という社会調査を継続実施してきました。この伝統プロジェクトも岐路を迎えていると考えています。具体的な調査結果や関連研究を少し紹介しながら、伝統を守りながらも、新しい姿で社会調査の分野の NOE 形成に果たしていくべき役割について考えているところです。アドバイスを頂けますと幸いです。

4 年前の研究所のホームページには、トップに創立 75 周年を迎えましたというページがありました。さしずめ、日本人の国民性調査プロジェクトに関して言いますと、1953 年に始まりましたので、開始 70 周年を迎えました。古希ということになります。少しプロジェクトもくたびれているかもしれません。

こちらのスライドで初期の調査委員会は、かなり大がかりなものであったことが分かります。所長が委員長となっていまして、大がかりな人数でプロジェクトを進めて、全所的なプロジェクトにあったという感じがあります。今、少しそこを拡大してみますと、初期の調査委員会には、今から思えば、少し意外な名前も含まれています。赤線を引きましたお名前、林知己夫、西平重喜、鈴木達三あたりは調査科学の分野では、ビッグネーム、特に林知己夫が統数研の所長もされましたし、林の数量化理論で有名な名前で、顧問の先生方にもお名前をご存じの方が多いのではないかと思います。緑の線を引いた赤池弘次先生、それから清水良一先生は、共に統数研の所長をなさった方ですが、調査の分野と少し離れている先生にもかかわらず、当初の委員会には、名前を連ねていたということがあります。また、青線で下線を引きました野元菊雄先生、これは国語学の分野では本当にビッグネームで、国立国語研究所の所長もなされた方です。国語研究所の共同研究の跡も見えるということがありまして、1948 年には、統計数理研究所のメンバーが参加して、日本人の読み書き能力調査というとても大規模な調査もありました。言葉に関する調査としては、1950 年、鶴岡市における言語調査なども有名なプロジェクトです。

さて、現在につながる調査科学関係の NOE 活動としては、統数研が NOE 活動を始めた比較的早期に、調査科学 NOE が設置されました。70 年以上にわたる社会調査研究の成果を基盤として、その発展や、社会的貢献促進のために活動をしています。本 NOE は幾つかのプロジェクトを展開しています。そのプロジェクトを今日全て紹介することはできませんが、日本人の国民性調査を中心に紹介させていただきます。

今となっては、なぜ統計学の研究所で社会調査の研究なのかという、若い人などはそのよ

うに思ってしまうかもしれませんけれども、ご承知のとおり、社会調査は、統計学に支えられた重要な応用分野の 1 つであり、標本調査理論だけにとどまらず、調査方法論、survey methodology と呼ばれる分野も統計学が支えているといって過言ではありません。

私ども、調査科学 NOE は、統数研の中には、センターを持っていません。調査科学グループというのが担当しています。この 3 月で、人数が少し減ってしまいまして、所内では、いわゆる承継教員は、私、グループリーダーを務めている前田と、それから朴准教授がいまして、特任助教、特任研究員が 1 人おり、9 人ほどの客員の先生にご協力をいただいているという状態になっています。図の右側に主要なプロジェクトが紹介してあります。赤字で今年度やろうとしているということが書き込まれているという状態ですけれども、ご興味があれば、内容を少しお読みになっていただければと思います。後ほど部分的に説明します。客員の先生は、こちらにもお名前がありますけれども、佐藤嘉倫先生にも来ていただいているということです。幾つかの目標を掲げて、今年度の活動を予定しているところです。

さて、今日の話題の中心である日本人の国民性調査ですけれども、1953 年以来、5 年に一度実施している継続社会調査でして、研究所のホームページにも、入り口ページから中身をご覧いただけるようなページを用意しています。興味があれば、ぜひご覧いただければと思います。

直近の調査は、2018 年に行われました第 14 次全国調査ということになりなります。そちらに紹介したページに 7 つのポイントで主な結果を紹介しています。タイトルだけ読み上げますと、伝統的意見の退潮、緩やかに進むドライ志向、持ち直し安定する経済面の評価、伸び悩む環境保護に対する意識、安定した日本人の性格に関する自己評価、社会生活上の不安や原子力施設の事故に対する不安の変化等々でした。

さて、なぜ統計数理研究所が、かくも長きにわたってこの研究を続けてきたかという点については、いつもこの 3 つの目的を説明させていただいています。1 つ目は、もちろん、これは、社会調査を実施するモチベーションそのものですが、日本人の国民性を解明するということです。調査結果を通して日本人のものの見方や考え方と、その変化を明らかにすることが 1 つの目的です。2 つ目の目的は、調査手法の研究開発です。実際の調査を行いながら、これからの社会変化にも対応できる新たな統計調査手法を研究していくことが必要です。3 つ目は、統計手法の研究開発で、調査のデータを素材として、より優れた統計的な方法を研究していくこと、これがモチベーションになっています。

継続社会調査で何が分かるかということですけれども、先ほどの目的 1 に関連したことを知るというのが大きな目的になっていまして、5 年に一度繰り返してきていますけれども、基本的には、同じ調査方法、同じ調査項目で継続してきたということがあります。同じ調査項目とはいうものの、実は項目の入れ替わりは非常に激しくて、完全に続けて調査できている項目というのは、それほど多くありません。しかし、同じような方法で、同じ調査項目で継続することによって、日本人の意見の変化を知ることができるということです。逆に、同じ調査方法ということを守ってきた理由は、調査の結果が、調査方法によって非常に異なっ

てしまうということがあるからです。同じ方法、項目で継続しないと、変化が分からないということがあります。1953 年に開始し、5 年に一度のプロジェクトですので、2023 年は、ちょうど古希の実施年でして、第 15 次全国調査を予定しています。

今、申し上げました調査方法によって結果が変わってしまう劇的な例を、2012 年の私どもの研究から紹介します。衆議院議員の総選挙がある時、普通はどうしますかという質問に、何をおいても、いつでも投票しますという回答は、面接法と呼ばれる方法で調査をすると、45%弱の回答なのですが、留置法と呼ばれて、自分で記入する方法で回答していただくと、32%ということで、調査の結果が 13 ポイントも違ってしまいます。そのような違いが生じてしまいまして、調査手法を変えるだけで、調査の結果というのは非常に大きく変わってしまうことを示す劇的な例です。そうすると、逆に言うと、調査手法を変えてしまうと、調査の継続性が失われるということですから、私どもは、面接調査という方法で調査をずっと続けてきています。

長期にわたる日本人の意見の変化の例を少しだけご紹介します。ここの調査結果のご紹介は、時間調整用のバッファとして用意しているところです。全てはご紹介せず、ほんの一部だけご紹介して時間を節約したいと思います。

典型的な例は、項目整理番号 4-10 で、他人の子どもを養子にするかという質問項目です。「子どもがない時は、たとえ血のつながりがない他人の子どもでも、養子にもらって、家を継がせたほうがよいと思いますか、それとも、継がせる必要はないと思いますか？」という意見について、継がせたほうがいい、継がせなくてもいい、意味がない、場合によるという意見がどのように変化してきたかというのを示したのがこちらのグラフです。

1953 年の開始当初は、継がせるという意見が 70%超で、大多数だったのですが、急激に減りまして、1968 年には、継がせないというほうの意見が多数派と、意見が逆転するような形になりました。その後も、ずっと意見は減ってきています。最近は動きが少なくなってしまっていますけれども、ともかくこれは日本の社会において家制度というものの重要性が薄れてきた証しということで、このことによって、過去 65 年にわたる日本人の価値観の変化というのを知ることができる例です。

他に、これは、林知己夫先生が書かれた本には、よく紹介されている人情課長という例なのですが、あまり意見の変化がない項目の例のような形で紹介されることが多いです。実は、これも少し最近は揺らいできているというのが、私の見立てです。

日本人の宗教心というのは、実は、国際的にもしばしば研究対象、興味の対象になる話題なのですが、これまでの調査回ではおおよそ、3 割くらいの人が信じていて、7 割ぐらいの人が信じていないという傾向が続いています。それでも日本人の宗教的な心は、少し最近やはり弱まってきているのかもしれないということを示唆する結果があるといったことがあります。

次に、自然と人間を巡る意識も少し変化が見られるかもしれません。これは少しゆゆしき話題ですけれども、環境の保護は重要かという質問に関して、非常に重要であるという回答

が、どうも最近少し減ってきているようで、環境の保護に関する日本人の意識はどうなっていくのかということが少し懸念されます。

政治に関する意識として、選挙への関心を問うた先ほどの質問なのですが、これは、実は、何をおいても投票するというのは、長期低落傾向です。このような変化がなぜ起きているのかということについての分析を、統計的な立場から行うことができます。これは、先ほど目的 3 で申し上げました、統計的な手法の研究の一番典型的な例だと私自身は考えておりますが、コウホート分析というのが可能です。意見の変化がなぜ起こるのかということについて、時勢の効果、世代の効果、出生コウホートの効果、そして、人が加齢することによる加齢の効果という 3 つの成分に分解することができます。このモデル化部分に赤池先生流のベイズ型のアプローチを導入するという手法開発を、OB の中村隆が行いまして、様々な項目に応用しております。この質問で「何をおいても投票する」という回答が減ってきたのは、主にコウホート効果が優勢であるということが分かっております。コウホートが入れ替わって、新しい世代が、別に投票しないという、何をおいても投票する、を選ばない傾向があることによって、意見が変化していったということが分かっています。

このように、統計的な立場の研究も、この調査データを用いて常に行われているということですが、このプロジェクトは、近年は、どのような形で大型の研究費を取得してきているかということを、少し NOE の活動としてスライドに紹介させていただきました。他の NOE に比べると、随分古いものから引用させていただいていますが、2010 から 14 年、これは国際比較の研究を行うための基盤 S というのを OB の吉野諒三が取得しています。12 年から 15 年には、中村隆代表で基盤 A、それから、18 年から 21 年度の計量的日本人研究の新展開というのを私が代表で取得していまして、これが、前回の第 14 次全国調査の基盤となる研究費となりました。そして今年度から 4 年度の、国民性調査の継承と発展という課題で、私が代表の基盤 A を取得していまして、これで 15 次調査、この科研費だけでまかなえる研究の規模ではないのですが、これを基盤に、今年度の秋の調査を予定しているということです。

国民性調査は、長く続いている調査で、非常に大きな課題を抱えています。手法面では、回収率が長期的に非常に低下していまして、それに対する対応としては、調査のバイアスを研究するといった統計学的な研究も必要になっています。もちろん回収率を維持することそのものが、とても重要になっています。それよりも今とても大きな問題は、面接調査法の危機を迎えているということです。長期的にも面接調査法は、なかなか難しくなっていますが、コロナ禍直後ということもあり、今年度の調査から、少しこの部分では、心配になっていまして、対応としては、他の調査モードによる実験を含ませて行うということを考えています。

内容面の課題は、他の標準的社会調査と比べた時に、項目がやや異なっています。統計学者が考えてきた社会調査ですので、いわゆる社会学者や政治学者などが考えてきた標準的な社会調査の内容と比べると、少し項目が個性的かなというところがあります。しかし、少

なくとも、ここに関してしては、属性変数について充実を図ることによって、他の分野の研究者にも、いろいろと使っていただける調査データセットにしていこうということを考えています。また、日本人の研究だけで日本人の国民性が分かるのかという問題に関しては、近隣の、ある程度文化を共有する国との協力による国際比較を導入することを考えます。今、手始めに、韓国の一般社会調査（KGSS）との項目共有を予定しています。

最後の点は、とても大きいことなのですが、調査項目が、内容面で疲弊しているということがあります。先ほど動きが鈍くなったという項目の例を少し紹介しましたが、動きが少ない、最近では意見の変化が少ない項目も増えてきていまして、従来項目の継続とともに新しい項目の導入による発展が必要となっています。

国民性調査プロジェクトを離れますと、調査科学 NOE では、いろいろな研究テーマに参画させていただいていて、例えば、科研費ですと、科研基盤 A、大阪大学の吉川徹代表や、金沢大学の小林大祐代表などといった課題に参加させていただいています。上の 2 つのプロジェクトネットワーク型と、一応、名前を仮に付けさせていただいています。他方例えば、下の 2 つは基盤研究型と仮称しています。これは先ほど少し紹介した朴准教授の研究ですけれども、郵送調査法の回収率向上方策などについての発表や、あるいは朴さん自身が代表になっている基盤研究で、新しい調査方法に関する研究を行っているといったことがあります。

あるいは、次のスライドではその他の調査プロジェクトに、いろいろと社会調査のプロジェクトにお声掛けいただいていまして、私の参加している幾つかの科研費のプロジェクトをリストにさせていただきました。最後のものは、朴さんが参加していることなのですが、筑波大学の先生との共同研究も行われています。自己評価として、いろいろな課題に協力していますけれども、NOE という意味で、Network Of Excellence の意味で、個別の共同研究を超えたネットワークハブ的な役割を果たしているかというと、少しそこの部分は、十分ではないのかなと考えているところです。

私自身が調査科学 NOE の長期的課題の対応策ということを考えているところは、1 つは、人員が減少しているということでして、2012 年、約 10 年前に承継職員が 6 人の体制でいました。特任についても、多少いたものが、現在では、承継職員 2 人で支えているという状態でして、若手人材のリクルートが急務です。また、先ほど少し足りないと申し上げたネットワーク形成型の活動を強化する必要がありまして、大学共同利用機関法人の他機構間との協働や、さまざまな大学の先生の共同研究を戦略的に活用していくことが必要であろうと考えます。また、新しいデータサイエンスの潮流の中、それこそビッグデータの時代です。社会調査は、50 年前では、ビッグデータのほうだったと思うのですけれども、今ではスモールデータになっているかと思います。伝統を継承しながら、発想の転換というものが必要になってくるかなと思いますが、私自身のスタンスは、まだ、なお少量データを質高く取得するための技法の鍛錬が必要と考えています。

また、国際性担保も重要な課題です。実は、所内の研究者を含めてやや内弁慶になってい

たところがあり、日本の調査科学全般にガラパゴス化が進んでいるかなと考えています。これは、戦略的国際連携を通じて、構成メンバーが努力して、国際性を担保していく必要があるかなと思います。最後に、少し発信が最近不足していますが、人的リソースが足りないこともありまして、なかなか強く成果発信ができていない点について、特に書籍化などを通じて、少しこの研究をさらに広く知っていただく活動を強化したいと考えています。私からの説明は以上です。ご清聴ありがとうございました。

【司会】前田先生、ありがとうございました。ただ今の調査科学 NOE のご発表に対して、コメント、ご質問いただければ、幸いです。いかがでしょうか。伊藤先生、お願いします。
【伊藤顧問】NISTEP（科学技術・学術政策研究所）がありますよね。NISTEP も同じような調査をやっていて、以前、私、そこの外部評価をやった時に、NISTEP では、人社と、それから、自然科学を融合して、全体を考えますという話がありました。NISTEP での技術予測動向調査、これは 5 年置きなのですけれども、かなり科学技術的なものばかりで、人社に関してどのように考えるかということに関しては、これからの課題だという議論もありました。このように NISTEP も社会活動の調査を進めようとしているのですが、何か連携するようなこととは、お考えになっているでしょうか。
【前田准教授】個人的に研究のプロジェクトに参加した経験があるようなメンバーは、過去にもいると思うのですが、現在は特別の協力関係というのはまだ構想できていません。でも、この課題に書きましたように、他の機構等との連携で伝統的な調査法の彫琢というか、技術を磨くということも必要になりますし、あるいは共創の中での新しいテーマの発掘といった活動も、当然重要になってくるかなと考えていまして、参考にさせていただきたいと思います。
【伊藤顧問】よろしくお願いします。
【司会】ありがとうございます。では、佐藤先生、お願いできますでしょうか。
【佐藤顧問】佐藤です。よろしくお願いします。まずは、70 年にわたって継続されているということには、本当に頭が下がる思いで、大変な努力をされたことだなと思って、敬意を表します。ちょうど出ている最後のスライドで、やはり国民性調査も、いろいろ曲がり角に来ている、それをどのように乗り越えていくかということがまとまっていると思います。特にビッグデータとの兼ね合いをどうするかなのですが、私は、今、日本学術会議の分科会でビッグデータの活用に関する報告を書いて、提出したところです。けれども、社会調査データとビッグデータは、必ずしも対立するものではなくて、どれほどビッグデータが使われるようになっても、やはり社会調査データというものは必要です。

特に意識を問う質問というのは、ビッグデータでは、まずは取れません。ビッグデータというのは基本的に、やはり行動レベルのデータなので、Twitter でも、本当の意見を表明しているかどうかは分かりません。意見を表明するという行動は分かりますけれども、先ほど前田先生が出されたように、少し聞き方が変わっただけで、もう分布が変わるようなセンシ

ティブな問題というのは、やはり社会調査データで、質の高い少量データをきちんと取っておくというのは、重要なことなので、今後も継続していただきたいと思います。

もう一つは、ビッグデータと社会調査データをリンクさせるという試みが、今、行われつつあります。要するに、社会調査データを取った時に、その対象者に、「Twitter のアカウントを教えてください。それで、いろいろトレースするけれども、いいですか」といって、同意を取った人にだけ、そのような Twitter データや、またモバイルデータなどを集めます。そうすると、国民性調査で回答をした人が、どのような行動をしているのか、どのようなツイートをしているのかというところで、また、多元的な分析ができるのではないかなと思っています。

最後、国際性の担保で、これは、これで非常に重要なことなのですが、少し気になるのは、成均館大学の調査研究センターと連携しながら、KGSS の項目の共有化という話がありましたが、一方で、大阪商業大学の KGSS の研究拠点のほうでも、EASS（East Asian Social Survey）をやっていて、そこでも KGSS と連携しているわけです。すると、大阪商大の EASS とのうまい連携というか、一歩間違うと、対立構造が起きそうな気がするので、そこら辺をどのようにお考えなのか、教えていただければと思います。以上です。

【前田准教授】最初の 2 点のコメントは、大変ありがとうございます。私も、ほとんど同じ意見を持っていまして、ここの赤字で書いたようなことについて開拓の余地があると考えて、書かせていただきました。3 番目の点ですが、大商大と何か特別にコンフリクトの関係があるとは、必ずしも考えていません。うまく国民性調査の中で生かしてきたものを、成均館大学のプロジェクトのほうで採用していただいて、という方向を考えています。また、調査テーマのモジュール化とローテーションといった EASS の動向についても、私は、この 2 月に合同カンファレンスという催しに参加させていただいて、動向としては承知していまして、うまくすみ分けができるかなと考えているところです。

【佐藤顧問】分かりました。どうもありがとうございます。

総合討論および顧問
お一人ずつからの総括コメント

•

謝辞および閉会

● 総合討論および顧問お一人ずつからの総括コメント

【司会】他、いかがでしょうか。スケジュールによりますと、ここで休憩を入れた後、総合討論ということになっていますが、時間も押し迫っていますし、このまま総合討論に移らせていただきたいと思います。ご了解いただければ、幸いです。

総合討論の仕方なのですが、ここまで統数研の NOE 形成事業と、それからスパコンと統計エキスパート人材育成プロジェクトの説明をさせていただきました。それで、ここからは各 NOE、もしくは事業の活動評価に関すること、また、全体の印象を含め、今回の発表になかったことまで含めて、統数研の印象等、そのようなものについて顧問の先生方から順番にご発言いただくということでお願いしたいと思っています。先生によっては、もうかなりご指摘を既に頂いたので、さらにというのは、何か心苦しい点がありますので、そこら辺は、無理のない範囲で、何かコメントいただければという程度で考えていただければ、幸いです。また、これは、ぜひ言いたいという点がありましたら、そのようなところについては、ディープに語っていただければ、われわれにとっては、力になりますので、よろしくお願いします。

伊藤先生には、もう既にいろいろとご意見を頂いているのですけれども、さらに総合的なご意見がありましたら、よろしくお願いします。

【伊藤顧問】伊藤です。今日は、いろいろ興味深いお話をお聞かせいただき、誠にありがとうございました。1 点だけ、少し屋上屋的な意見になってしまうかもしれません。最初の所長のお話にもありましたけれども、これから URA や、あるいは支援員など、いわゆる研究者でない方々も増やしていきたい、それによって研究力を高めるという話がありました。吉田先生のものづくりセンターでも、そうなのですが、例えば、NIMS の場合は、研究者と同じぐらいの人数のエンジニアがいます。もちろん統数研は、実際に材料を扱っているわけではないので、それほどの人数は要らないと思います。しかし、例えば、データベースを維持するなど、そのようなことに関しても、技術者、エンジニアという方々も、どうしても必要になってきますので、ぜひご検討されて、今後の計画に反映していただければと思っています。私から以上です。

【司会】ありがとうございます。山形先生、まだ大丈夫ですか。

【山形顧問】順番を変えてしまって、どうもすみません。私のほうから少し、中間のところが、少しだけ抜けてしまったのですが、いろいろとお話を聞かせていただいて、普段知らなかった素晴らしい業績についても知ることができて、本当にありがとうございました。私は、元々、国の研究機関にいまして、今、このようなアカデミックな統計数理研究所の研究業績が長い統計数理研究所の研究蓄積に基づいて、世界の最先端を切り開いているというのは、大変心強く思いました。これは、このままアカデミックな研究というのは、ぜひ進めていただきたいと思います。しかし、先ほど申しましたように、世の中の動きは、かなり速くて、今、社会人、大学院生のいる大学にいるのですが、やはりスマホの時代がさらに超えて、ChatGPT になってきて、世の中の動きが非常に速いものですから、このアカデミックの世

界と、実際の社会の動きのずれというのが、やはり少し気になるかなということです。

それで、多分、データサイエンスというのも、もう古くて、もう少し先に行って、Evidence-Based Decision Making というのですか、データを活用して、どのようにその科学的な知見なり、エビデンスに基づいてよりよい意思決定をするかというのが、サービスとして確立できるかというのが非常に求められているところです。

ですから、そのようなところを、データを集めるところから最後のアドバイスをするところまで、一気通貫というとマージャンのようであまりよくないのですけれども、できるようになると、素晴らしいなというのは、勝手に私の思っていることです。先ほど ROIS と組み合わせるという話もあったので、ROIS の話を聞いていると、データベースの技術ばかり研究しているように聞こえてしまって、これも外部からの勝手な思いなのですけれども、そのすごいデータベースで一体何をするのかというのが、よく分からないなというのがあります。実際の現実を考えると、あまりにも寒いということがあるので、日本全体としてのパフォーマンスを上げるために、せっかくの素晴らしい知識や技術が、どのようにしたら総合的に活用できるのかなというのも考えていただけるとありがたいなというのが最後のコメントです。以上です。

【司会】ありがとうございました。それでは、上田先生からコメントを頂ければありがたいです。

【上田顧問】先ほどもコメントをさせていただきましたけれども、やはりバーチャルラボというものに対する期待感は、個人的には、非常に大きいです。やはりご承知のように、日本は、研究開発者人口というのが、欧米、あるいは中国などと比べて圧倒的にスケールダウンしているので、やはり組織間で連携して、その知恵を融合していくということも大事です。ですから、そのコアになる統数研のバーチャルラボというのは、非常に重要なポジションになると思いますので、そこを期待しています。ぜひよろしくお願いします。以上です。

【司会】ありがとうございます。それでは、小宮山先生、よろしければお願いします。

【小宮山顧問】時間がないので、簡単に申し上げますけれども、まず 1 つは、統計エキスパート人材育成の話です。これは、多分公表されているので、言っていいのだと思うのですけれども、文科省側の推進委員もさせていただいているので、大応援団の団員です。これからもよろしくお願いします。

2 番目ですけれども、今までこのネットワークに製薬業界としてあまり関わりを持ってこられなかったなという、とても反省があって、先ほどのコンサルティング事業をきっかけにして、ネットワークに関わりを持たなければならないなと今日は本当に強く思ったところで、関わりを持ち始めなければ、ネットワークも何もないと思いますので、頑張りたいと思います。

あとは、この分野、医療の分野ですと、先ほど山形先生が Evidence-Based Decision Making という話もされましたけれども、医療や医薬の分野でも、このような話は出てきています。そのエビデンスを説明するために、丁寧なモデル化をしていき、それによって説明

に役立てるという流れが世界的な流れになってきていますので、これからもよろしくお願いしますというか、ご協力を賜ることもあるかもしれません。以上です。

【司会】どうもありがとうございます。こちらこそよろしくお願いします。

続きまして、佐藤先生、コメントをお願いしたいですが、いかがですか。よろしくお願いします。

【佐藤顧問】授業の関係で申し訳ないのですけれども、最後の 3 つしか拝聴することができませんでした。私は、統数研というのは、日本における統計の研究教育を下支えする重要な基盤になっていると思います。今日、3 つ伺ったのは、それぞれいろいろなところで、それをなされているなと思って、非常にありがたいと思っています。特に、最後の前田先生のは、私もよく知っているので、先ほどコメントをしましたが、統計エキスパート人材育成プロジェクトの概要というご説明を伺って、非常に、これは統数研らしいなと思いました。このようなことによって、日本の中で統計を教えられる教員をつくっていって、それを受けた若い人たちが、いろいろな分野で活躍できるようになっていったら、うれしいなと思っています。

ただ少し懸念事項は、このようなものは、一般市民の人たちまでどのように広がっていくか、また、今、私が非常に懸念しているのは、政治家です。例えば、コロナ対策というのは、本当にエビデンスに基づいた政策立案しているのかどうか、統計的ないろいろな知見をどのように取り込んでいるのかというものが、さっぱり分からないところがあります。私、今の統数研のアウトリーチ活動も非常に評価していまして、私自身も、幾つかアウトリーチの会合やシンポジウムにオンラインで参加していますけれども、そのように世間一般というのは、少し変ですけれども、日本全体的に、やはり統計の重要性を認識して、統計に基づいていろいろな政策を立てていきます。それを市民がやはり評価していくという雰囲気というか、環境が出来上がったらいいなと思っています。ですから、非常に私は統数研の研究教育、そして、アウトリーチ活動に期待しています。前田先生がおっしゃるように、人材が少なくなってくるという問題はあるとは思いますけれども、これからも大きな期待を持っていますので、よろしくお願いします。

【司会】ありがとうございます。副島先生、お願いできますか。

【副島顧問】今日、少しずついろいろなところで出てきたと思うのですけれども、私も研究所の経営というものに当たっていて気を付けているのは、このマーケティングの視線というのが、やはり大事だなと。研究の成果物を出していく時に、誰がその研究成果物のユーザーになるのだろうか、カスタマーになるのだろうかということを、要は、それが実務で使われる、ビジネスで使われるというところも意識して、研究内容を進めていく必要があります。そして、論文を出して終わりではなくて、どのように売っていくか、どんな人に知ってもらうかというところでマーケティングをもう少し頑張りましょうと声かけしています。

例えば、昔は、内部機関誌やジャーナルに投稿すれば、あとは学会報告ぐらいでした。情報の総供給が少ないので、需要してもらえたわけです。しかし、社会が情報過多となるなか

方法の変革が必要になりました。待っていては需要してもらえないのです。それで最近ではウェブマーケティングの強化に乗り出しています。

それから、誰がユーザーかを識別する課題です。3 つあると考えています。まず一つは、プロであり、それから、その応用を考える企業の人もプロだとすると、その対極に、先ほども少し出てきた普通の人というのがあり、そして、その中間に中継者になってくれるメディアや、あるいは専門家ではないのだけれども、ユーザーという意味で、例えば、意思決定に関わる企業幹部や政治家などが想定されます。そのような 3 つぐらいのレイヤーの人が大きく分けているのだろうなと。そのレイヤーごとに多分マーケティングは、恐らく変えていかなければならないということです。

論文を直接読めるのは、専門家だけです。なのでサマリーがあるわけですが、これも専門家向けと中階層向けでは書き方が異なります。例えば、エグゼクティブサマリーとノンテクニカルサマリーは異なります。前者は企業の意思決定者のために存在しているので、応用のところに重きを置いたような研究のような紹介の仕方になります。一般読者は、言葉の使い方から全く変えていく必要があります。分量という次元も大事です。複数のバリュエーションがあるはずです。要は、カスタマー向けの次元が 3 つあって、これに加えて、どのような媒体でどういう言葉で何に焦点をおいて伝えていくのかというところにもレイヤーがあります。もちろん、マトリックス全部やらなければならない必要性は、全然ありません。一般向けのところに論文などを展開する必要はないわけなので、分かりやすく語って、関心を持ってもらえればいいわけなので、そのような使い分けというのが一つあるのだろうなと思います。

もうひとつ別のお話をさせてください。結果の広報として研究をいかにオープンにしていくのかという話とは別に、データや手法のオープン化も重要です。先ほどもアンケートのお話がありましたけれども、あのようなアンケートをオープン化して、使えるようにしてくる、あるいは統計の大学の教員を育てるための研修プログラムというのがあります。これらがウェブ上に置かれていて、動画も含めてフリーに使えるようになっているところが大事なのだろうなと思います。私も会社の内部研修プログラムに携わっているのですけれども、一部の大学などは、学びのマテリアルのオープン化が進んでいて、自分で一から全然作らなくていいというところで、大変重宝させてもらっています。

最近は、ネット上にいろいろな大学のプログラムのようなものがあります。例えば、有名なものですと、Udemy です。Academy ではなくて Udemy、あるいは Coursera（コーセラ）などです。Coursera には、アメリカの大学の優れた授業が、そのまま大量に掲載提供されていて、非常に安く、場合によっては無料で使うことができます。ですから、その辺も研究成果のオープン化というところが、このアウトリーチにつながっていくのだろうなと考えています。以上です。

【司会】ありがとうございます。副島先生がせっかく顔を出されているので、少し私のほうから宣伝させていただきます。今月 26 日、統計数理研究所公開講演会というものがありま

す。その中で、副島先生から金融の現状およびその DX 絡み、フィンテック絡みのような講演をしていただくことになっています。ご興味のある方は、ぜひご参加していただくようにお願いします。副島先生、よろしくお願いします。

【副島顧問】ありがとうございます。すみません。では、宣伝絡みということで、私も 1 つ、少しだけ。金融研究所は昨年度に設立 40 周年を迎えました。日本銀行設立 100 周年を記念して 1982 年に設立され 40 年が経過したわけです。周年行事として Web コンテンツを作成しました。金融研究所のサイトの左上のトップのところに 40 周年記念というタグがあります。そこをクリックしていただくと、これまでの研究の振り返りや、それから、主要な研究者の方へのインタビューや対談などが載っています。最近デジタルマネーがとてもたくさん出てきているのですけれども、そもそもお金とは何か、現代のマネーシステムはどう構築されてきたのか、そのようなところからお話を始めた対談などもありますので、ご関心がある方は、どうぞということで、すみません。宣伝させていただきました。

【司会】順番が前後してしまいまして、すみません。蒲地先生に最後にスピーチいただきたいですけれども、いかがでしょうか。

【蒲地顧問】すみません。時間もないので、全体の印象として、前回の顧問会議から 3 年半ぐらいたって、非常に進展が見られているというのがよく分かりました。機械学習による異文化融合、それをネットワークとしてのハブとして進んでいるし、また、人材育成ということを生んでいるというのは、非常に大事だと思いました。日々、AI 関連の技術革新は進んでいるし、われわれの会社なり、研究機関なりでも、人材が足らないということを非常に感じていたので、特に専門的な分野を大事にしながら、AI もやっていくという、そのようなダブルメジャーでやっていくという例を非常によく示していただいて、ありがとうございました。単に統計解析だけではなくて、社会の問題に対応すべく、戦略的に取り組んでいらっしゃるというのが、よく分かりました。今日は、どうもありがとうございました。以上です。

● 謝辞および閉会

【司会】ありがとうございます。いろいろなご意見、顧問の先生方、ありがとうございました。それでは、最後に椿所長から一言お願いしたいと思います。

【椿所長】本日は、長時間にわたりまして、私どものNOE活動のみならず、統数研の第4期中期計画といいますか、これから進めなければならないということに関して、さまざまなご助言を頂いたことに大変心から感謝申し上げます。

われわれの研究所は、基本的に国立といいますか、そのような大学共同利用機関として、ずっと続けてきたわけですけれども、時代が非常に速く動いていきます。研究テーマ、研究活動も、大変速く動いていく中で、どのような新たな視点といいますか、考え方を投入していなければならないかということに関して、顧問の先生方から、大所高所によった大変貴重なアドバイスを頂けたと考えます。

今後、われわれとしては、私どもの活動の中で、先生方から頂いたコメントなどをどのように反映していくかということを、今日の議事録を作ると同時に、どのようにわれわれの活動に取り入れるか、あるいは、そのような視点を入れるかについて、執行部としても、あるいは、NOE形成事業の幹部も考えていかなければならないと承知しています。今回、NOEの顧問の先生方については、任期の最終年度を迎えているということではありますけれども、今後、ますますいろいろな意味で、所の活動に対して助言を頂く、忌憚（きたん）なく批判を頂くと共に、ある意味で甘えかもしれませんけれども、応援団としてわれわれの所の活動を支えていただくことを心から希望する次第です。

本日は本当に長時間にわたり、大変な活発な議論を頂いたことを、改めて感謝申し上げます。どうもありがとうございました。

【司会】本日、会議の冒頭で話しましたとおり、会議の内容については後日、報告集を作成し、発行させていただきます。顧問の先生方には、議事録に掲載する発言内容などのご確認のお願いを事務局からしますので、ご対応いただければ幸いです。本日は、私の司会が不手際で、大幅に時間を超過させてしまいまして、大変申し訳ありませんでした。先生方の意見、きちんと経営に反映できるように心掛けていきたいと思います。

以上をもちまして、第4回統計数理研究所NOE形成事業顧問会議を終了させていただきます。本日は、長時間、誠にありがとうございました。失礼します。

【一同】ありがとうございました。失礼します。

第4回 統計数理研究所NOE形成事業顧問会議 報告集

発　行　日　令和6年3月

編集・発行　情報・システム研究機構
統計数理研究所 NOE形成事業運営委員会
〒190-8562 東京都立川市緑町 10−3
Tel : 050 - 5533 - 8500（代表）
https://noe.ism.ac.jp/

統計数理研究所

NOE形成事業

Network Of Excellence