

情報・システム研究機構における データサイエンス推進の取組みについて

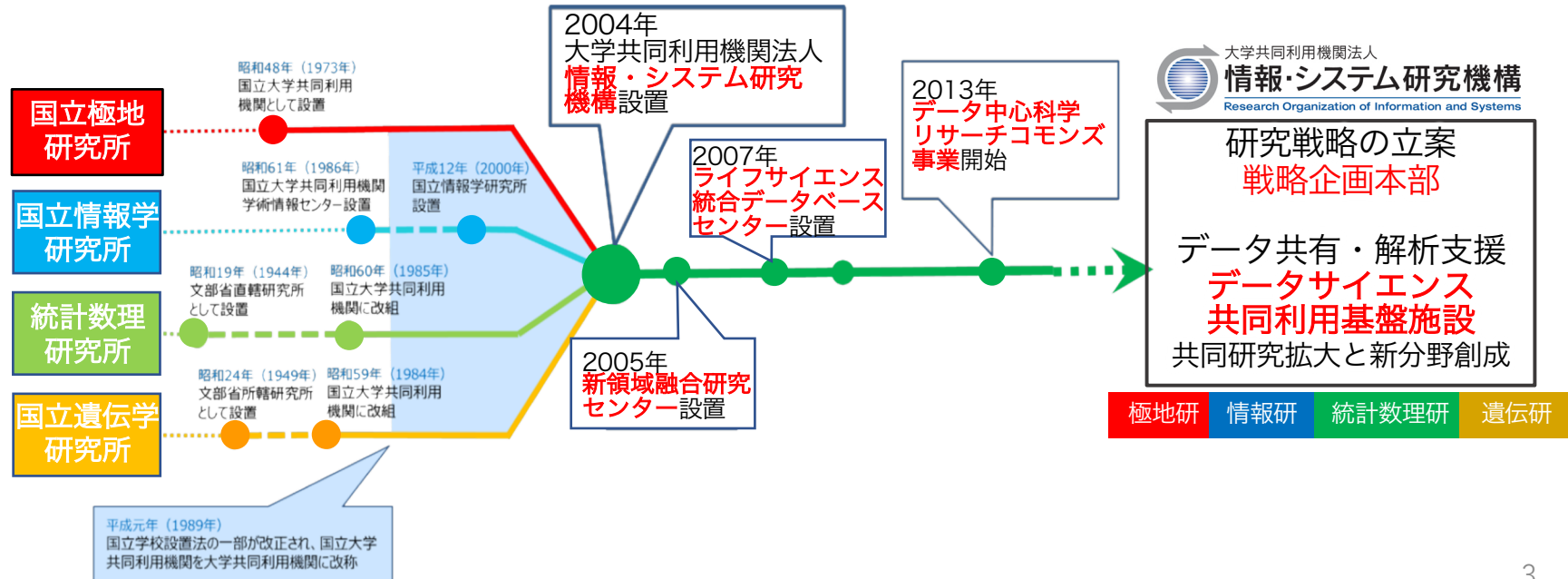
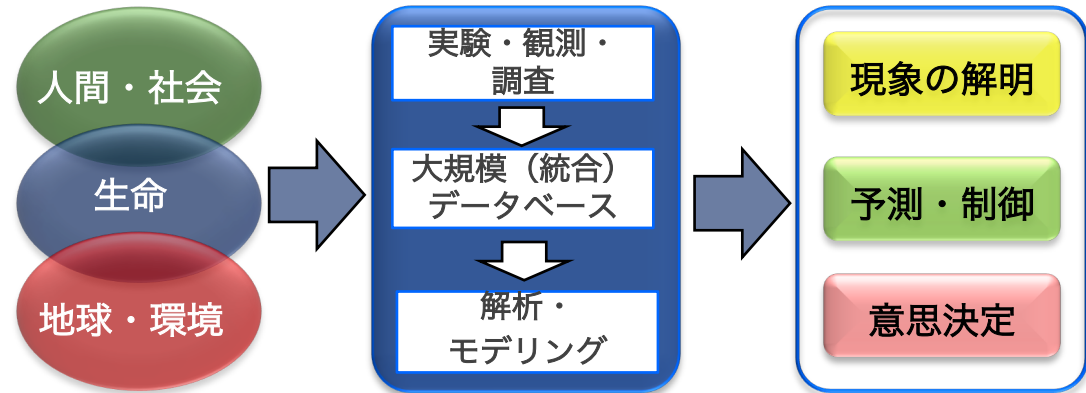
大学共同利用機関法人 情報・システム研究機構
機構長 藤井 良一

アウトライン

- 1) 情報・システム研究機構のミッション
- 2) データサイエンス時代の課題と当機構の役割
- 3) データサイエンス共同利用基盤施設を始めとする
データサイエンス推進の施策
- 4) データサイエンス人材の育成

機構の成立ちとミッションとしてのデータサイエンス推進

生命、地球、環境、社会などの複雑な問題を、物質とエネルギーの観点に替って**情報とシステム**という立場から捉えるための、方法の研究、研究基盤の整備および融合研究による新分野の開拓を行なう。



情報・システム研究機構の概要

法人の基本的な目標

国立極地研究所、国立情報学研究所、統計数理研究所、国立遺伝学研究所の4研究所が全国の大学等の研究者コミュニティと連携して、世界水準の総合研究を推進するとともに、機構長のリーダーシップのもと、21世紀社会の重要かつ複雑な現象に関する問題を情報とシステムという視点から捉えなおすことによって、その解決を目指す。

アクションプラン

1. 世界トップレベルの研究の推進

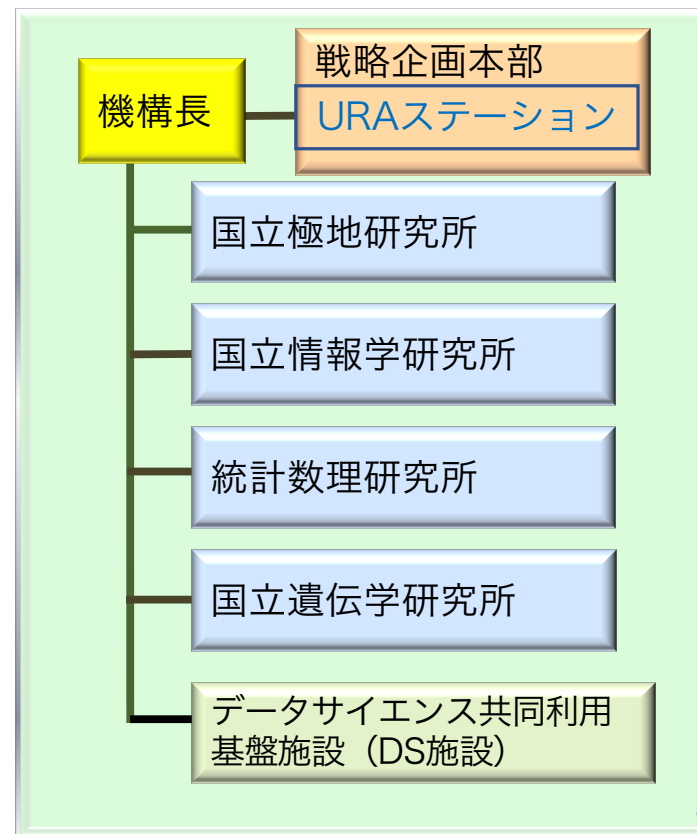
- ◆新たな研究パラダイムを構築し、新分野を開拓
- ◆世界水準の先進的な研究を推進

2. 共同利用・共同研究の推進

- ◆研究コミュニティの要請に応え、大学等の研究力アップに貢献
- ◆学術情報基盤(SINET, DDBJ等)を全国の大学等に提供

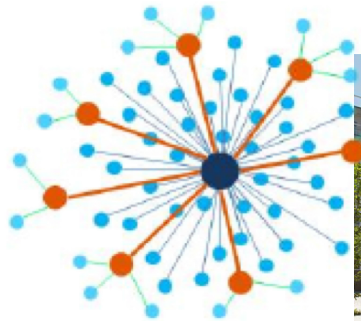
3. データサイエンスを振興

- ◆データサイエンス共同利用基盤施設(DS施設)の活動を強化し、データ駆動型科学を発展させて、社会イノベーションを推進
- ◆データ共有支援、データ解析支援を進め、同時にデータサイエンティストを育成



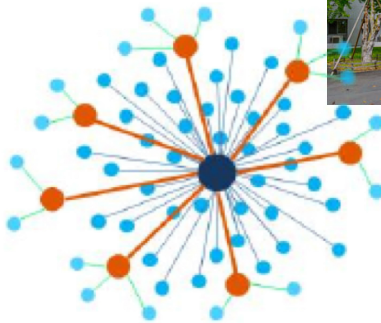
4 研究所 : 学術コミュニティの中核的拠点

学理の追求と大学共同利用、人材育成で
社会や大学等へ広く貢献する



各研究所（機関）は
各学術コミュニティ
のCOEとして機能

社会と連携し
基礎的基盤的な
成果を社会還元



大学等と法人単
位で連携し
最先端研究と
科学の発展支援



学術と社会のデータサイエンスへの要請

ビッグデータには膨大な知識や価値が埋もれている

しかし、現在の方法・技術では有効活用は不十分

- ・多くは構造化されていない
- ・価値密度が低い
- ・不均一（形式、精度、観測頻度、非定常性）、スパース

ビッグデータを効果的・効率的に集約し

革新的な科学的方法により知識発見や価値創造を行うことが重要

 データ中心の新たな方法論：第4の科学が必要

大規模データ処理
データ可視化

データ解析

大量の散在するデータをリアルタイム処理する技術
膨大な高次元データや計算結果を人間が把握可能にする技術

ビッグデータからの深い知識獲得のための方法

対応を迫られるビッグデータ例 EISCAT_3D

欧州非干渉散乱レーダー（北欧での超高層大気観測）

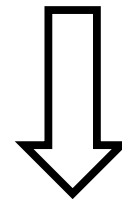
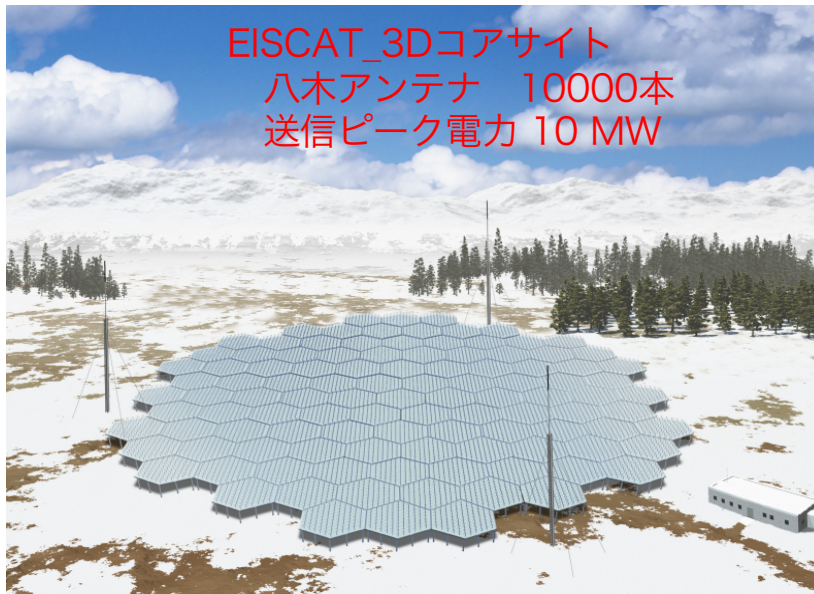
1997年から日本は極地研究所を代表機関として参加

2020年過ぎに新レーダー EISCAT_3D 観測開始

データの最速の出力は 1.6 Tbit/秒 (生データ)

リアルタイム処理後：

54 Gbit/秒 3.2 Tbit/分 4.7 Pbit/日 1.7 Ebit/年



数年後には対応が迫られる
Challengingな課題

分野融合・新分野創成のための機構の施策

●中期的対応（法人第3期）

データサイエンス（第4の科学）の推進
 データサイエンス共同利用基盤施設の

データ共有支援

データ解析支援

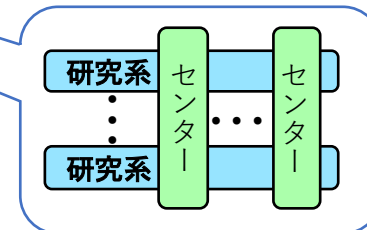
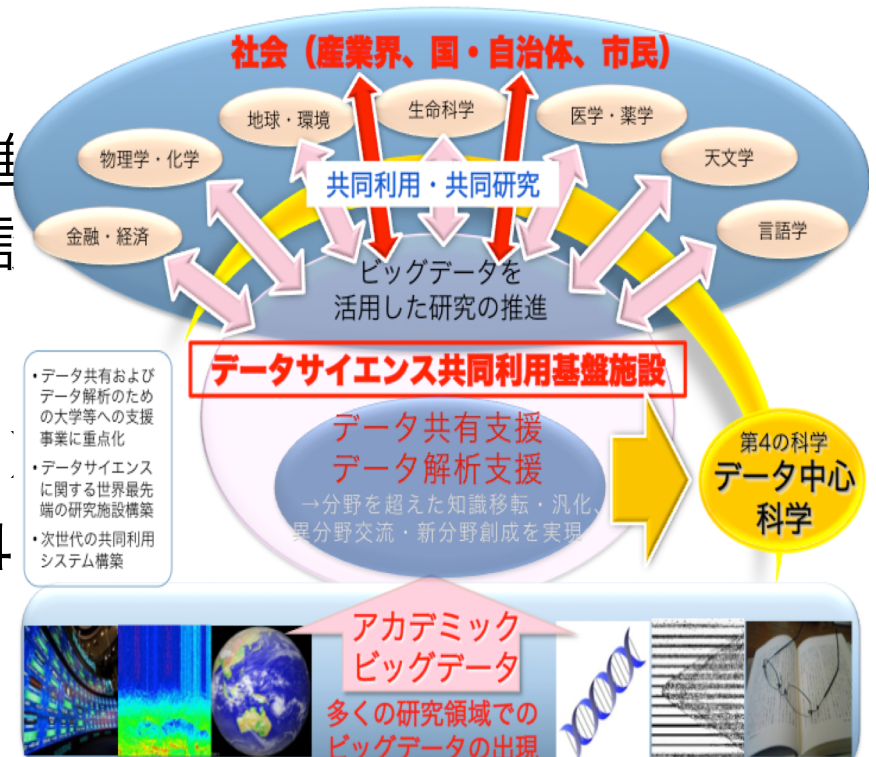
人材育成（データサイエンス、セキュリティ）
 （H30 所属研究者46名 年間予算4

オープンサイエンスの推進
 教員配置の二重構造化

←研究の規模、スピードへの対応

●長期的対応（法人第4期に向けて検討中）

システムサイエンス，システム化技術の推進

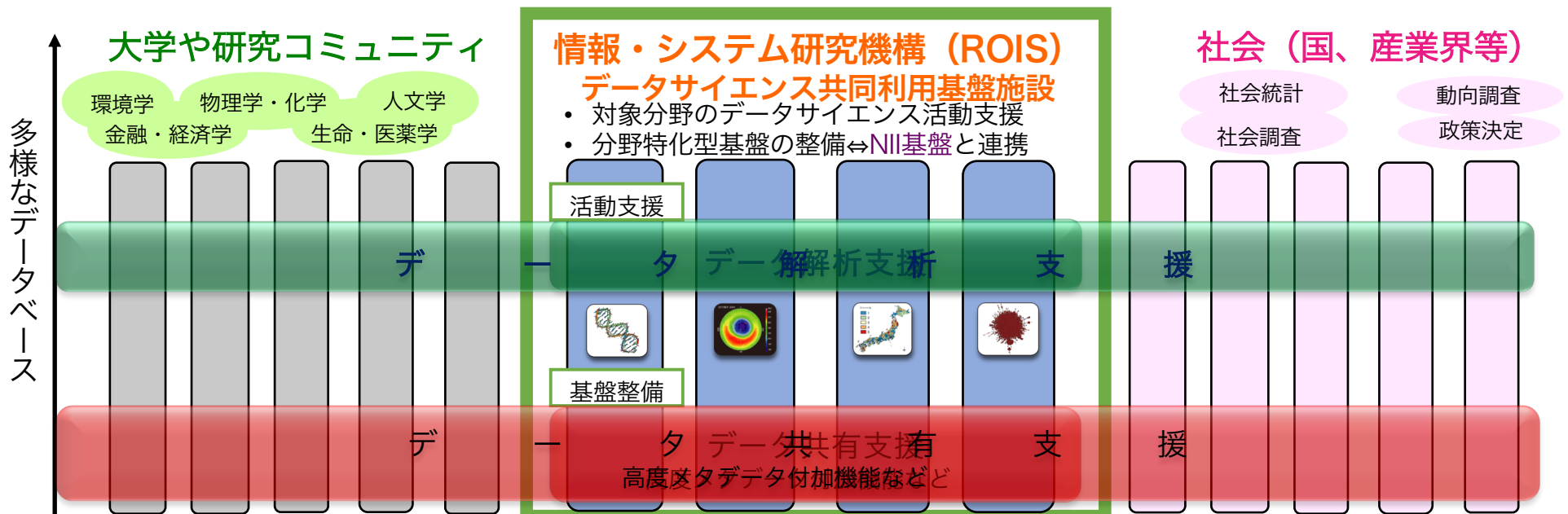


データサイエンス共同利用基盤施設の発展構想

～データサイエンスの展開：オープンサイエンス推進～

ROISデータ解析支援、共有支援を全国に展開（大学の機能強化等）
オープンサイエンスの実現→異分野融合・新分野創成を促進

学術界・社会全体のデータベース統合・相互利活用への道 ～オープンサイエンス実現へ



国立情報学研究所 (NII)

オープンサイエンス研究データ基盤

データ管理基盤 データ公開基盤 データ検索基盤

共通機能を基礎基盤として整備（全国規模で展開）



データ共有支援系

データ解析支援系

データベース名	担当センター	研究分野、コミュニティ	活動内容		解析対象	担当センター	研究分野、コミュニティ	活動内容	
<複数DBの統合化ポータル> ・ゲノム・発現統合検索 ・自然言語による統合検索 ・各種応用ツール及びコンテンツ ・各種RDF化支援ツール ・RDF 検索支援ツール	ライフサイエンス統合データベースセンター (DBCLS)	生命科学、医学、	データベース統合利用のための環境構築、データベース統合化のための基盤技術開発及び国際標準準備		<さまざまな生物種のゲノムデータ解析> ・新規ゲノム決定 ・ゲノム再シーケンス ・トランスクリプトーム解析 ・メタゲノム解析	ゲノムデータ解析支援センター (CGI)	ゲノム生物学 ゲノム医学 ゲノム創薬	次世代型DNAシーケンサーから得られる大量の配列データに基づいた多様な生命科学研究を対象に、情報科学的な解析支援を実施	
<極域科学のデータアーカイブ統合データベース> ・IUGONET:超高層大気 ・PANSY:大型大気レーダー ・EISCAT:オーロラ観測 ・EISCAT:衛星データ	極域環境データサイエンスセンター (PEDSC)	極域・地球環境	極地研による極域研究のデータ統合とデータサイエンス化から大学コミュニティへの展開を図る		<数値シミュレーション全般の理論・応用支援> ・統計的手法の研究・開発 ・データ同化による数値シミュレーションの実用応用 ・実用例：気象予報 ・エミュレータ開発	データ融合計算支援センター (CDFC)(仮称)	データ中心科学の考え方・手法を用いる分野全般	データ同化やエミュレータなどの手法を活用する研究相談・支援及び共同研究 上記に関連する講習会や体験学習の実施	
・日本人の国民性調査 ・意識の国際比較調査 ・(オンサイト解析室)	社会データ構造化センター (CSDS)	社会調査データ、公的調査のマイクロデータ、ソーシャルビッグデータ	社会状況に関するデータ(社会調査、公的マイクロデータベース、ソーシャルビッグデータ) 有効利用のための大学間連携ネットワーク基盤形成及び地域社会への貢献						
<データサイエンスに基づく人文科学(人文情報学)分野の創生> ・歴史的状況記録データベース ・日本古典籍データセット ・日本古典籍字形データセット ・江戸料理レシピデータセット ・顔貌コレクション ・近代雑誌データセット	人文学オープンデータ共同利用センター (CODH)	人文学、国語学、人文情報学	くずし字チャレンジ！ IIIF Curation Viewer IIIF Curation Finder						

データサイエンス共同利用基盤施設 公募型共同研究

DS施設公募型共同研究「ROIS-DS-JOINT 2018」採択課題一覧（第1回募集分）

一般共同研究

課題番号	継	申請者	所属機関/職名	研究課題名	DS施設側教員	DS施設側サテライト等名
001RP2018		細川 敬祐	電気通信大学/大学院情報理工学研究科/准教授	ロングイヤビン及びアイスランドの光学観測データを活用したオプティカルフローデータベースの構築	田中 良昌	極域環境データサイエンスセンター
002RP2018		深沢 圭一郎	京都大学/学術情報メディアセンター/准教授	観測・数値シミュレーション・機械学習の融合による宇宙プラズマ現象予測モデルの開発	中野 慎也	データ融合計算支援プロジェクト
003RP2018	○	Masayuki Kanai	Senshu University / School of Human Sciences / Professor	Harmonization and Documentation of International Survey Data on Wellbeing in Seven Asian Countries for Publication	Tadahiko Maeda	社会データ構造化センター
004RP2018		馬場 社太郎	琉球大学/教育学部/教授	南極岩石試料レポジトリの確立によるデータサイエンスへの応用	矢吹 裕伯	極域環境データサイエンスセンター
005RP2018	○	吉沢 明康	京都大学/大学院薬学研究所/特定助教	文字列高速検索技術とRDFを用いたタンパク質及び塩基配列同定リソースの開発とその応用	内藤 謙樹	ライフサイエンス統合データベースセンター
006RP2018		堀田 成政	京都大学/地球環境学堂/助教	地理的加重モデルによる多変量地理空間データ解析	村上 大輔	データ融合計算支援プロジェクト
007RP2018	○	高橋 彰	関西学院大学/総合政策学部/契約助手	メモリーハンティングを活用した戦後から現在の京都の景観変化に関する研究 - 京都市電のデジタル・アーカイブ写真を事例として -	北本 朝展	人文学オープンデータ共同利用センター
008RP2018		桑原 知巴	香川大学/医学部/教授	既知ゲノム情報に基づく難培養性ヒトコカス科バクテリアの培養条件設計	五斗 進	ライフサイエンス統合データベースセンター
009RP2018		鴨川 仁	東京学芸大学/教育学部自然科学系基礎自然科学講座物理学分野/准教授	全地球電気回路研究のための地上大気電場観測データのデータベース化とデータ解析システムの開発	門倉 昭	極域環境データサイエンスセンター
010RP2018		渡部 諭	秋田県立大学/総合科学教育センター/教授	犯罪統計データの解析法の開発とオープンデータに向けた検討	吉野 諒三	社会データ構造化センター
011RP2018		加藤 千尋	信州大学/学術研究院理学系/准教授	超紅外線宇宙観測データのためのリアルタイム・アーカイブシステムの構築	門倉 昭	極域環境データサイエンスセンター
012RP2018		三浦 恭子	熊本大学/大学院生命科学研究部老化・健康長寿学分野/准教授	最長寿命菌類ハダカバネズミにおける老化耐性機構の制御遺伝子の同定	坊農 秀雅	ライフサイエンス統合データベースセンター
013RP2018		陳 艶艶	福岡工業大学/社会環境学部/助教	調査データに基づく環境配慮行動のパターン解析とモデル化	吉野 諒三	社会データ構造化センター
014RP2018	○	櫻井 伸一	京都工芸繊維大学/繊維学系/教授	ポリマー材料の一軸延伸過程での電発発生メカニズム解明のためのデータ同化シミュレーション	中野 慎也、野村 俊一	データ融合計算支援プロジェクト
015RP2018	○	嶋田 健一	ハーバード大学/医学部システム薬理学研究室/博士研究員	オノログ共進化パターンに基づくゲノムワイド遺伝子相互作用の推定	千葉 啓和	ライフサイエンス統合データベースセンター
016RP2018	○	天竺桂 弘子	東京農工大学/大学院・農学研究院/講師	昆虫由来化合物の新奇な構造を創る代謝酵素CYPの同定	坊農 秀雅	ライフサイエンス統合データベースセンター
017RP2018	○	才田 聡子	北九州工業高等専門学校//准教授	インタラクティブな複合大規模データの可視化ツール開発によるデータサイエンス教育教材の提供	中野 慎也	データ融合計算支援プロジェクト
018RP2018		山本 真行	高知工科大学/システム工学群/教授	南極インフラサウンド観測データ収録公開システムの構築	金尾 政紀	極域環境データサイエンスセンター
019RP2018		海老原 祐輔	京都大学/生存圏研究所/准教授	南極点及びマクマード基地における多種類の光学観測データを用いた統合データベースの開発	門倉 昭	極域環境データサイエンスセンター
020RP2018		村田 健史	情報通信研究機構/ソーシャルノーションユニット総合システム研究開発推進センター/研究統括	歴史的行政区画データセットへの現代町丁目地理データ接続とCODHからのWeb公開	北本 朝展	人文学オープンデータ共同利用センター
021RP2018		深田 秀実	小樽商科大学/社会情報学系/教授	観光地でのビジュアルデータ活用による「観光防災情報プラットフォーム」に向けた研究	小出 哲彰	社会データ構造化センター
022RP2018		伊藤 伸介	中央大学/経済学部/教授	データ構造から見た公的統計ミクロデータの利活用のあり方に関する研究	南 和宏	社会データ構造化センター
023RP2018		河野 憲嗣	大分大学/経済学部/教授	大分県におけるインバウンド観光客の動線分析	小出 哲彰	社会データ構造化センター
024RP2018	○	姫野 順一	長崎外国語大学/外国語学部/特別任用教授	スモホ・アプリ長崎古写真ハンティングの実証とARとの結合実験	北本 朝展	人文学オープンデータ共同利用センター
025RP2018	○	堂免 隆浩	一橋大学/大学院社会学系研究科/教授	自治体連携型調査システムの開発とその成立条件に関する研究	前田 忠彦	社会データ構造化センター
026RP2018		馬場 保徳	石川県立大学/生物資源工学研究所/助教	リガシールロス分解複合微生物系におけるde novo assembly based RNA-seq解析	野口 英樹	ゲノムデータ解析支援センター

(1) 一般共同研究

データサイエンスに関する研究課題について、他機関研究者等とDS施設研究者がDS施設を利活用し、少人数で行う研究。

研究代表者は、情シス以外の教育・研究機関（国外を含む大学、大学共同利用機関法人、国立研究開発法人、独立行政法人等）や行政組織等と直接雇用関係にあり、かつ研究・開発、調査を主たる業務とする方（図書館司書や博物館等の学芸員、研究員を含む）。

【1件あたりの申請額】 上限100万円/年
【同一課題の継続】 3年を上限



(2) 共同研究集会

比較的少人数で実施する集会で、共同研究に向けた討論やデータサイエンスに関する研究交流、講習会等。共同研究集会に参加するための旅費や開催に要する経費負担。



申請者は、情シス以外の教育・研究機関（国外を含む大学、大学共同利用機関法人、国立研究開発法人、独立行政法人等）や行政組織等、民間企業と直接雇用関係にある方で研究・開発、調査を主たる業務としない方も申請が可能。

【1件あたりの申請額】 上限50万円。

未来投資型プロジェクト（新研究領域開拓）

目的・対象者		プロジェクト実施内容（H29年度の例）																																		
<p>目的：機構の第4期中期目標・中期計画期間に向けて、情報及びシステムの研究についての有望な事業を調査探索する。</p> <p>申請対象者：機構に所属する研究者（研究教育職員・特任教員・特任研究員）</p>		<p>【FS：Feasibility Study】 公募期間：平成29年4月28日～6月1日 採択数（応募数）・金額：8件（19件） 11,820,000円 採択方法：外部有識者を含む審査委員により評価。その順位を参考に、採択可否及び配分額を決定。 年度末評価を行い、次年度継続事業（FS→本研究へ）4件を決定。</p>																																		
<table border="1"> <thead> <tr> <th colspan="2"></th> <th>H28年度</th> <th>H29年度</th> <th>H30年度</th> </tr> </thead> <tbody> <tr> <td rowspan="4">未来投資型プロジェクト</td> <td>応募</td> <td>25</td> <td>19</td> <td rowspan="2">公募中</td> </tr> <tr> <td>FS採択</td> <td>13</td> <td>8</td> </tr> <tr> <td>本研究 新規)</td> <td></td> <td>6</td> <td>4</td> </tr> <tr> <td>本研究 継続)</td> <td></td> <td></td> <td>6</td> </tr> <tr> <td colspan="2">計</td> <td>13</td> <td>14</td> <td>10</td> </tr> </tbody> </table> <p>  FSから本研究への採択  本研究（新規）から本研究（継続）への採択 </p>				H28年度	H29年度	H30年度	未来投資型プロジェクト	応募	25	19	公募中	FS採択	13	8	本研究 新規)		6	4	本研究 継続)			6	計		13	14	10	<p>課題例：「動物学・流体力学・ロボット工学の融合による鳥類の形態進化研究と動物型ロボット設計への応用：「野生動物工学」分野の創成」（極地研） 研究内容：流体学的な視点に基づく鳥類の野外行動とクチバシ形態の分析結果を応用した技術試験ロボットの設計、「野生動物工学」という新たな研究分野の創設を目指す。</p>								
		H28年度	H29年度	H30年度																																
未来投資型プロジェクト	応募	25	19	公募中																																
	FS採択	13	8																																	
	本研究 新規)		6	4																																
	本研究 継続)			6																																
計		13	14	10																																
<table border="1"> <thead> <tr> <th rowspan="2">H29年度</th> <th colspan="2">共同研究者数（人）</th> </tr> <tr> <th>FS採択課題</th> <th>本研究 新規)</th> </tr> </thead> <tbody> <tr> <td>国立大学</td> <td>8 (32%)</td> <td>16 (47%)</td> </tr> <tr> <td>大学共同利用機関法人</td> <td>4 (16%)</td> <td>9 (26%)</td> </tr> <tr> <td>公立大学</td> <td>1 (4%)</td> <td>0 (0%)</td> </tr> <tr> <td>私立大学</td> <td>5 (20%)</td> <td>3 (9%)</td> </tr> <tr> <td>公的機関</td> <td>2 (8%)</td> <td>2 (6%)</td> </tr> <tr> <td>民間機関</td> <td>1 (4%)</td> <td>1 (3%)</td> </tr> <tr> <td>外国機関</td> <td>4 (16%)</td> <td>3 (9%)</td> </tr> <tr> <td>その他</td> <td>0 (0%)</td> <td>0 (0%)</td> </tr> <tr> <td>合計</td> <td>25 (100%)</td> <td>34 (100%)</td> </tr> </tbody> </table>		H29年度	共同研究者数（人）		FS採択課題	本研究 新規)	国立大学	8 (32%)	16 (47%)	大学共同利用機関法人	4 (16%)	9 (26%)	公立大学	1 (4%)	0 (0%)	私立大学	5 (20%)	3 (9%)	公的機関	2 (8%)	2 (6%)	民間機関	1 (4%)	1 (3%)	外国機関	4 (16%)	3 (9%)	その他	0 (0%)	0 (0%)	合計	25 (100%)	34 (100%)	<p>【本研究】 採択数（申請数）・金額：6件（13件） 18,000,000円 採択方法：実施報告書を基に審査委員による年度末評価を行い、本研究としての採択可否及び配分額を決定。 年度末評価を行い次年度継続事業（本研究1→2年目へ）6件を決定。</p>		
H29年度	共同研究者数（人）																																			
	FS採択課題	本研究 新規)																																		
国立大学	8 (32%)	16 (47%)																																		
大学共同利用機関法人	4 (16%)	9 (26%)																																		
公立大学	1 (4%)	0 (0%)																																		
私立大学	5 (20%)	3 (9%)																																		
公的機関	2 (8%)	2 (6%)																																		
民間機関	1 (4%)	1 (3%)																																		
外国機関	4 (16%)	3 (9%)																																		
その他	0 (0%)	0 (0%)																																		
合計	25 (100%)	34 (100%)																																		
		<p>課題例：データ中心ケミストリ：データ駆動型新規物質・知識創成に向けたケミカルデータ共有支援（機構本部） 研究内容：新規化学物質の発見と化学合成法の設計に資する、量子科学とデータケミストリを統合した化学物質・反応データ基盤の創出と共有支援を推進。</p>																																		
		<p>課題例：画像検索のための構造化問い合わせ言語による歴史的典籍画像検索システム（情報研） 研究内容：画像検索処理を構造的・統一的に記述する問い合わせ言語、「PQL: Pictorial Structured Query Language」を構築し、従来不可能だった「画像内容を考慮した問い合わせ」を可能とする画像検索システムの構築を目指す。</p>																																		

機構間連携・文理融合プロジェクト

目的・対象者		プロジェクト実施内容（H29年度の例）																																		
<p>目的：大学共同利用機関4法人間の連携を具体的に進める。文理融合に代表されるように、従来の研究分野を超えた新領域を創成する。 申請対象者：機構に所属する研究者</p>		<p>【FS】 公募期間：平成29年6月1日～6月30日 採択数（応募数）・金額：3件（3件） 2,800,000円 採択方法：審査委員により評価を行い、それを参考に採択の総合評価を行い、配分額を決定。 年度末評価を行い、次年度継続事業（FS→本研究へ）2件を決定。</p>																																		
<table border="1"> <thead> <tr> <th colspan="2"></th> <th>H28年度</th> <th>H29年度</th> <th>H30年度</th> </tr> </thead> <tbody> <tr> <td rowspan="4">機構間連携・文理融合プロジェクト</td> <td>応募</td> <td>3</td> <td>3</td> <td rowspan="2">公募中</td> </tr> <tr> <td>FS採択</td> <td>3</td> <td>3</td> </tr> <tr> <td>本研究 新規)</td> <td></td> <td>1</td> <td>2</td> </tr> <tr> <td>本研究 継続)</td> <td></td> <td></td> <td>1</td> </tr> <tr> <td colspan="2">計</td> <td>3</td> <td>4</td> <td>3</td> </tr> </tbody> </table> <p>  FSから本研究への採択  本研究（新規）から本研究（継続）への採択 </p>				H28年度	H29年度	H30年度	機構間連携・文理融合プロジェクト	応募	3	3	公募中	FS採択	3	3	本研究 新規)		1	2	本研究 継続)			1	計		3	4	3	<p>課題例：Seeking Sentence Unit of Language：クラウドソーシングを用いた手話対話における文単位認定プラットフォーム構築（情報研） 研究内容：クラウドソーシングを用いた手話対話における文単位認定の手法を構築し、日本手話-日本語自動翻訳システム開発等に向けた環境の整備を目指す。</p>								
		H28年度	H29年度	H30年度																																
機構間連携・文理融合プロジェクト	応募	3	3	公募中																																
	FS採択	3	3																																	
	本研究 新規)		1	2																																
	本研究 継続)			1																																
計		3	4	3																																
<table border="1"> <thead> <tr> <th rowspan="2">H29年度</th> <th colspan="2">共同研究者数（人）</th> </tr> <tr> <th>FS採択課題</th> <th>本研究 新規)</th> </tr> </thead> <tbody> <tr> <td>国立大学</td> <td>4 (29%)</td> <td>8 (40%)</td> </tr> <tr> <td>大学共同利用機関法人</td> <td>10 (71%)</td> <td>9 (45%)</td> </tr> <tr> <td>公立大学</td> <td>0</td> <td>0</td> </tr> <tr> <td>私立大学</td> <td>0</td> <td>1 (5%)</td> </tr> <tr> <td>公的機関</td> <td>0</td> <td>2 (10%)</td> </tr> <tr> <td>民間機関</td> <td>0</td> <td>0</td> </tr> <tr> <td>外国機関</td> <td>0</td> <td>0</td> </tr> <tr> <td>その他</td> <td>0</td> <td>0</td> </tr> <tr> <td>合計</td> <td>14 (100%)</td> <td>20 (100%)</td> </tr> </tbody> </table>		H29年度	共同研究者数（人）		FS採択課題	本研究 新規)	国立大学	4 (29%)	8 (40%)	大学共同利用機関法人	10 (71%)	9 (45%)	公立大学	0	0	私立大学	0	1 (5%)	公的機関	0	2 (10%)	民間機関	0	0	外国機関	0	0	その他	0	0	合計	14 (100%)	20 (100%)	<p>【本研究】 採択数（申請数）・金額：1件（3件） 3,000,000円 採択方法：実施報告書、発表を基に審査委員による評価を行い、本研究としての採択可否及び配分額を決定。 年度末評価を行い、次年度継続事業（本研究1→2年目へ）1件を決定。</p> <p>課題名：日本列島人の進化とその言語文化の起源（遺伝研） 研究内容：膨大なヒトゲノム情報を基に日本語祖語の渡来時期と経路を推定し、DNAデータから日本語の系統の解明を目指す。</p>		
H29年度	共同研究者数（人）																																			
	FS採択課題	本研究 新規)																																		
国立大学	4 (29%)	8 (40%)																																		
大学共同利用機関法人	10 (71%)	9 (45%)																																		
公立大学	0	0																																		
私立大学	0	1 (5%)																																		
公的機関	0	2 (10%)																																		
民間機関	0	0																																		
外国機関	0	0																																		
その他	0	0																																		
合計	14 (100%)	20 (100%)																																		

人材育成基盤：研究者や高度技術者等の養成

時代が求める研究者・高度技術者を社会・大学等に輩出

研究所が実施する教育プログラムをスケールアップ

- データサイエンティスト養成（統数研）
- トップエスイー（情報研）
- 情報セキュリティ技術者養成（情報研）



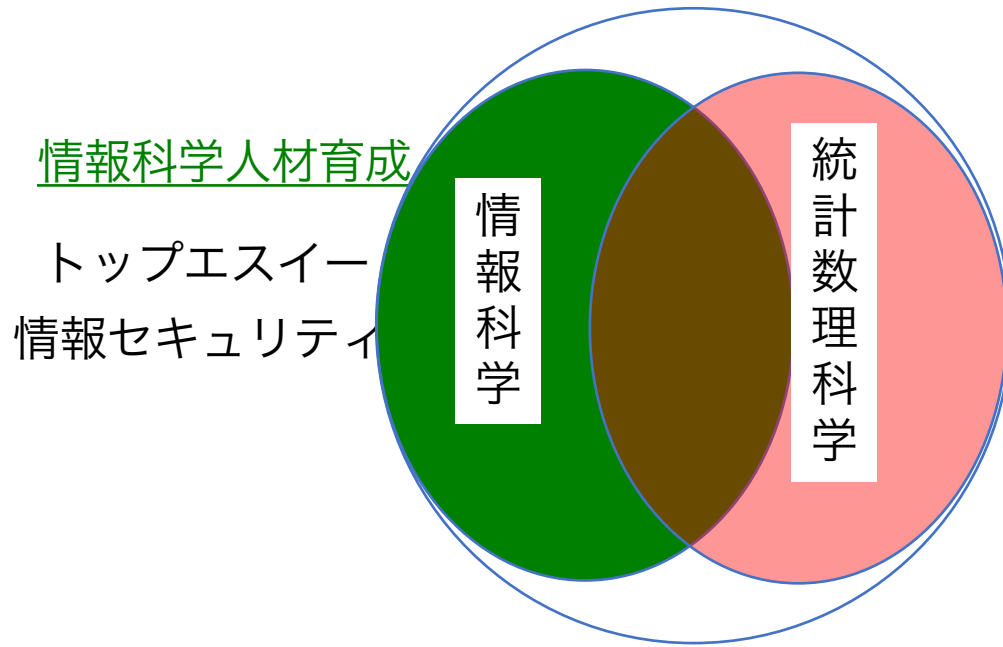
各研究所の取組
機構支援の強化

種類の増加、人材輩出数の増加、OJT強化
社会や大学等のニーズに質、量ともに応える

研究者・社会人の交流

- 研究者交流プログラム（所属機関への支援）
 - クロスアポイントメント制度を用いた双方向の人材循環
- 大学の機能強化と研究所・機構の発展双方に貢献

ROIS各研究所のデータサイエンス人材育成



Bio informatics人材育成

国際開発者会議 BioHackathon
統合データベース講習会 AJACS
SPARQLthon

統計数理学人材育成

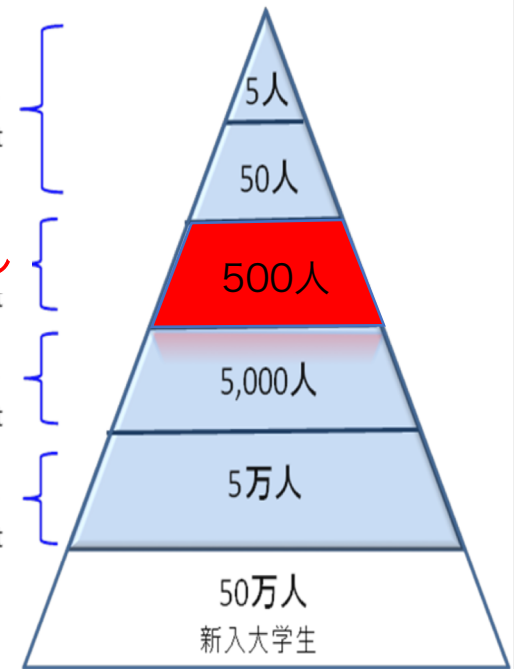
データサイエンティスト協会が定めたスキルレベル (2014年12月)

業界を代表するレベル
Senior Data Scientist

棟梁レベル
(III) Data Scientist

独り立ちレベル
Associate Data Scientist

見習いレベル
Assistant Data Scientist



データサイエンス高度人材育成プログラム (1)

ハンズオン、ハッカソン、データコンペ、アウトプット(合宿) キャンプなどの人材育成プログラムを機構全体でパッケージ化して支援 (H29～ 機構長裁量経費)

	ハンズオン	ハッカソン	データコンペ	アウトプットキャンプ	その他
DS施設	<p>AJACS(統合データベース講習会) (DBCLS) (JST/NBDC=科学技術振興機構/バイオサイエンスデータベースセンターの 主催)</p> <p>IUGONETの講習会(PEDSC) データベースや解析ツールの講習ことどもならず、どのようにデータを比較するかというような観点を含む。海外実施実績(特に途上国)あり (国内2回、海外2回)</p> <p>CODHチュートリアル (国内2回)</p> <p>データ同化ハンズオン(D融合計算支援)</p>	<p>国際版BioHackathon (JST/NBDCと ROIS/DS/DBCLSの共催)</p> <p>BLAH(Biomedical Linked Annotation Hackathon) (JSTとの共同研究の一環)</p> <p>SPARQLthon (JSTとの共同研究の一環)</p> <p>国内版BioHackathon (JSTとの共同研究の一環)</p> <p>RDFサミット (JSTとの共同研究の一環)</p>	<p>くずし字チャレンジ! 日本古典籍のくずし字を機械学習で解読するコンペティション (CODHI協力)</p>	<p>国際マイクロデータワークショップ(CSDS) 国際マイクロ統計データベースのデータ提供国(アジア8ヶ国)から毎年8ヶ国の統計局関係者を招き、約1週間、追加予定の試行版データを利用した国際比較ワークショップ</p>	<p>RDF講習会 (DBCLS) (DBCLS主催、JSTとの共同研究の一環)</p>
極地研	<p>北極域課題解決人材育成講座 ロシア極東のサハ共和国やクーツクに日本の大学院生、若手研究者、行政・企業の実務者を派遣して「ウィンタースクール」を開催</p> <p>安全管理・危機管理の合宿訓練 南極観測隊の合宿訓練(冬、夏それぞれ1週間)</p>			<p>AFoPS(アジア極域科学フォーラム)フェローシップ アジアの若手研究者を最大90日間招聘</p> <p>極域科学インターンシップ 院生に限定する代わりにアジアという制限をなくした事業で、最大90日間まで招聘</p> <p>共同利用育成研究員制度 大学院生を対象、毎年15～20名程度利用</p>	<p>共同利用研究集会 集中討議的な集会を毎年開催</p>
情報研	<p>研究教育のためのクラウド活用セミナー 現場の研究者や教職員が抱えている、研究教育や大学業務にどのようにクラウドを活用できるのかといった疑問の解消を目的としたセミナー(4回開催、うち1回はハンズオン)</p> <p>権限リポジトリ新任担当者研修 大学等における学術情報基盤を支える人材の育成</p>	<p>SINETアイディアソン 学術情報ネットワークの活用に関するアイディアソン</p>	<p>グラフコンペ 効率的なスパコン設計につながるグラフ発見を競うコンペ</p>	<p>NI国際インターンシップ 海外機関からの大学院生の受け入れ事業、最大6か月まで招聘</p>	<p>トップエスイー ソフトウェアエンジニアリングの技術・理論・ツールを使いこなすスーパーアーキテクトを育成する、社会人向けの教育プログラム</p> <p>サイバーセキュリティ人材育成 各大学のサイバーセキュリティを担当する技術職員等の技術レベル向上及び執行部への説明能力を身に着けた 橋渡し人材の育成</p> <p>JMOOCのプラットフォームgaccolによる講座配信 研究者支援に携わる者が、効果的な研究データ管理に関する基礎的な知識や方法を学習</p> <p>SPARC Japanセミナー オープンサイエンス等の学術情報流通をテーマとしたセミナー</p>

データサイエンス高度人材育成プログラム（2）

ハンズオン、ハッカソン、データコンペ、アウトプット（合宿）キャンプなどの人材育成プログラムを機構全体でパッケージ化して支援（H29～ 機構長裁量経費）

	ハンズオン	ハッカソン	データコンペ	アウトプットキャンプ	その他
統数研	<p>データ同化ハンズオン データ同化の理論から応用に渡る幅広い研究につながる生きた知識を習得</p>	<p>データ分析ハッカソン 3名1チームで、クラウド上でのデータ分析に基づいて、正解を決めにくい課題、例えば社会問題の解決や新規サービスの提案を競う</p>	<p>スポーツデータ解析コンペティション 日本統計学会等との共催、企業の協賛 データ分析チャレンジ 正解が決まっている課題、例えば高精度の売り上げ予測モデル構築する</p>	<p>夏期大学院(統計数理ブートキャンプ) 実践プログラミングまでのスキル向上を目標とする英語コース ※約4割が外国人参加者</p>	<p>リーディングDAT(Leading Data Analytics Talents)プログラム 複業レベルを問わずデータサイエンティストに必須の統計数理の知識を効率的に習得させるための系統的な講座、コースを実施 ※コース履修完了者には認定証を交付 公開講座 統計科学の最先端理論・手法から基礎的なものまでを学べる多様な一般講座 組織連携に基づくデータサイエンス講座企画 大学や産業界等の要請を受けデータサイエンス講座を講義内容と講師を合せて企画提案 共同研究スタートアップ データサイエンス高度人材を育成するためのOJTとしての側面を併せ持つ 統計教育動画配信 LMS上でのe-learningコンテンツ提供とYouTube配信 データサイエンス・リサーチプラザ 企業の技術者・研究者が一定期間滞在し、統数研の研究環境を利用して行う研究活動の支援 統数研データサイエンス・ハイスクール 高校生向けの学びの場の提供 医学・医療分野に特化および関連した統計公開講座、系統的な生物統計教育コースの開催 医療健康データ科学研究ネットワークの形成と情報交換の場の提供 医学研究者を対象としたデータサイエンスの系統的教育プログラムのe-learningサイト利用、情報交換の場の提供</p>
遺伝研	<p>DDBJデータベース利用講習会 「DDBJing」として毎年開催（H29年度は中縄高専） 他の講習会(トレーニングコース) ・バイオリソース利用法や新規実験手法についての講習会(例:ショウジョウバエ分類講習会) ・バイオインフォマティクス分野のデータアナリストの育成を目的とした「D-STEP講習会」</p>	<p>BioHackathon、SPARQLthon、Annotation Hackathon についてDBCLSとの協力</p>		<p>NIGインターン 国外の学部生・大学院生を6週間遺伝研に招聘(研究室に配属し研究に参加してもらう。最後に成果発表のためのシンポジウムを開催) 共同利用研究員育成 国内外の大学院生を共同利用研究員として受入れ人材育成 遺伝研体験入学 国内の学部生・大学院生に遺伝研の研究室に1～5日程度滞在してもらい研究生活を体験させる(毎年開催)</p>	<p>遺伝研メソッド 科学英語についての遺伝研企画の講習会 NBRPのXenopusコミュニティとゲノム解析及びゲノムブラウザに関する技術交流会 ゲノム解読やリシーケンスなどの支援・技術指導 DS施設や先端ゲノム支援研究解析プラットフォームと連携して実施</p>

新時代の学術研究へ向けての情報・システム研究機構の方向性の一端をお話しました。

各研究所の学理を基に、学術基盤の充実を図り、大学や社会と連携・共創しつつ、機構一丸となってデータサイエンス、オープンサイエンスを推進し、今まで以上に大学等の機能強化や社会のイノベーションに貢献することを目指します。

皆様のご支援ご鞭撻をお願い申し上げます。

ご静聴ありがとうございました。