



日本学術会議における検討

学術研究フォーラム第9回学術シンポジウム
「オープンサイエンスの展開」@NII
23期日本学術会議オープンサイエンスの取組に関する検討委員会
委員長 土井美和子
情報通信研究機構/奈良先端科学技術大学院大学
2018年10月17日

日本学会会議のオープンサイエンス取組経緯

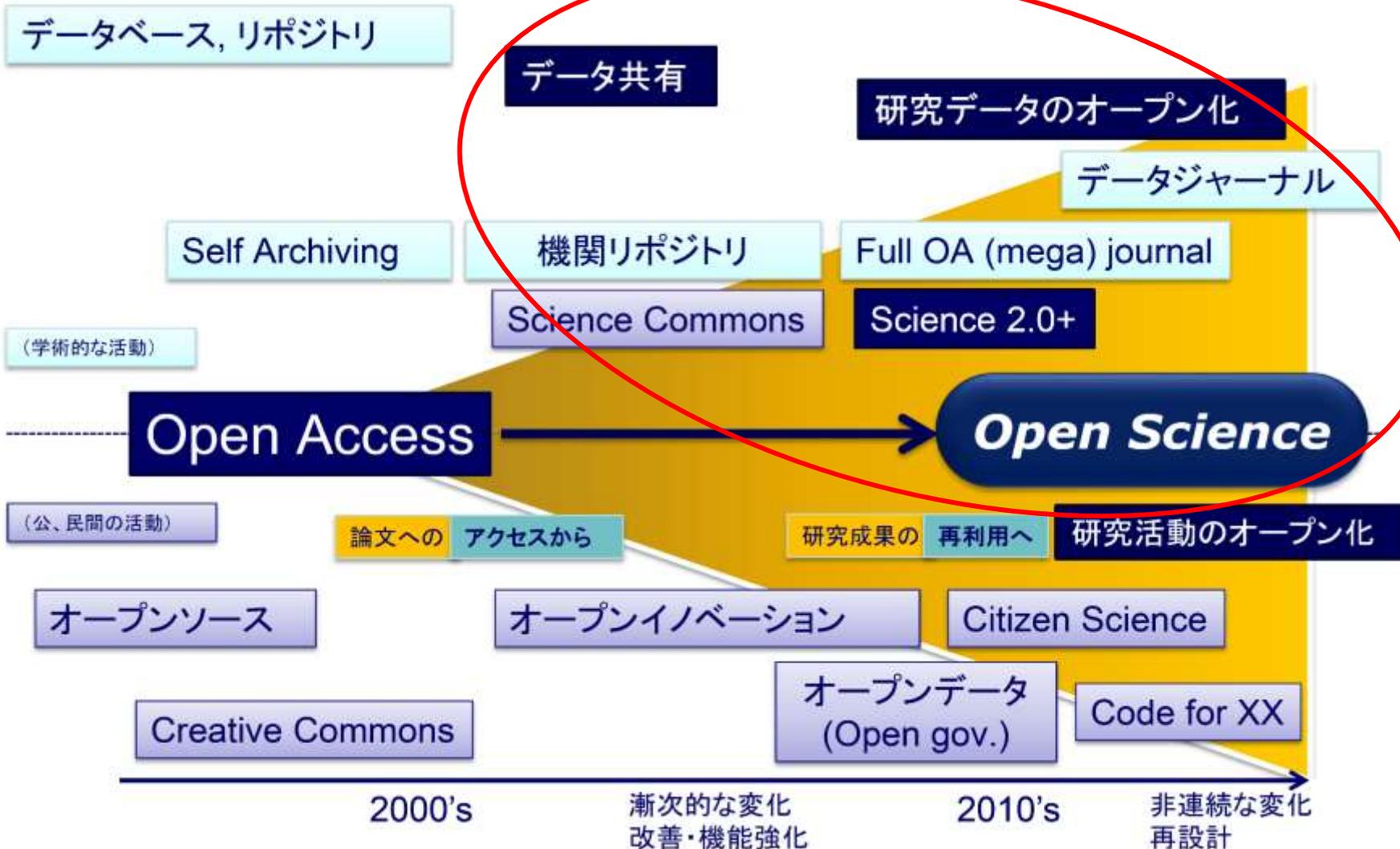
- 2015年1月 オープンサイエンスの取組に関する検討委員会発足

多数の意見聴取

学協会へのアンケート

-

対象とするオープンサイエンス



内閣府「国際的動向を踏まえたオープンサイエンスに関する検討会」報告書より引用

意見聴取リスト

- 我が国におけるオープンサイエンス推進の在り方について 真子**内閣府**政策統括官付参事官補佐
- ICSU-WDS等の国際事業と科学データ共有動向 村山NICT室長
- **ライフサイエンス**におけるデータ共有の現状と課題 高木東大教授・JST NBDCセンター長
- **材料分野**におけるオープンサイエンスについて 藤田NIMS部門長
- **微生物・環境データ**の統合化と今後の展開 黒川東工大教授
- **天文学**における世界規模のデータ共有 大石国立天文台 天文データセンター センター長
- **社会学**におけるオープンサイエンスへの取り組み:社会科学のデータアーカイブの構築 北村一橋大教授
- **人文学**とオープンサイエンス 青柳文化庁長官
- アカデミア**創薬**研究とオープンサイエンス 長野医薬品医療機器総合機構理事
- オープンサイエンスに関する日本の最近の活動事例について 林**文科省**NIST EP上席研究官
- **学術情報**のオープン化の推進について(審議まとめ)概要 渡邊文科省学術基盤整備室室長
- オープンサイエンス推進のための**研究データ基盤** 安達**国立情報科学研究所** 副所長

学協会へのアンケート

- 第1部、第2部、第3部それぞれの関連学協会へのアンケートを実施。
- 回答のあった学協会のうち、過半の学協会が論文や論文にかかわるデータ、データベース等のデジタルデータを既に公開。
- 半数程度の学協会では研究者個人がデータを保有しているが、共通のフォーマット・様式を整えれば価値を生む可能性のあるデータがある。
- 残念ながら、内閣府報告書の存在については、回答のあった学協会ではあまり認識されていなかった。
- 過去5年の間に、オープンサイエンスをテーマとしたイベントの開催事例があるのは、1割程度とあまり多くなかった。
- 半数近い学協会がオープンサイエンスに関する自由意見を寄せており、オープンサイエンスに大いに興味を持っている。

日本学会会議のオープンサイエンス取組経緯

- 2015年1月 オープンサイエンスの取組に関する検討委員会発足
多数の意見聴取
学協会へのアンケート
- 2016年7月6日 提言発表

日本学会議からの提言

(1) 研究分野を超えた研究データの管理およびオープン化を可能とする研究データ基盤の整備
コスト負担やサイバーセキュリティ対策



研究データ基盤: 研究コミュニティ毎のオープン・クローズのデータ戦略に基づく研究データのリポジトリの整備・運用、廉価なデータ保存を可能とする仕組みも必要である。

(2) 研究コミュニティでのデータ戦略の確立

対象となるデータの見極め、占有期間の設定、データのオープン範囲の決定、データ解析ツールの包含、などのオープン・クローズ戦略とガイドラインおよびデータの取捨選択の明確な基準は、各研究コミュニティが検討すべき

(3) データ生産者およびデータ流通者のキャリア設計

オープンサイエンスを推進する3つの要素

	ボトムアップ:研究者	トップダウン:政府・助成機関
データ公開	公開により研究成果の信憑性を向上。従来知見の再利用。	公的研究資金を受けた研究成果の公開。
データ管理	苦労して取得した研究成果を安全な環境で保全。	研究成果のもととなった研究資料は10年間の保存を原則。
データ検索	研究活動の効率化。分野融合型研究の加速。	市民科学。産学連携。国際的な発信力の強化。

研究データを、

- 公開しなければならないのは研究者の責任。
- 保全する環境を整備するのは研究機関の責任。
- 流通させるサポートをするのは図書館の責任。

(ICSU-IAP-ISSC-TWAS working group, Open Data in Big Data World, 2015年12月 より改変)

これらのニーズや要求を実現するためのシステム基盤が必要

研究データ基盤

これまでに研究データの共有や公開の機会がなかった分野の研究者も活用できる基盤として学際的なオープンサイエンスの実現を目指す

データ検索基盤

論文検索サービスCiNiiを研究データに発展させたサービス

再利用

研究データ管理

ユーザインタフェース

アクセスコントロール

メタデータ管理

研究データ管理基盤

管理基盤

論文のオープンアクセスには存在しない新サービス

メタデータ管理
データ検索サービス

検索・利用

データ利用者

データ登録者

実験・保存

蓄積

実験データ

成果論文

非公開

共有

公開

Hot Storage

Hot Storage

Hot Storage

Cold Storage

Cold Storage

Cold Storage

長期保存対応ストレージ領域

DOI

海外
研究データ
公開基盤
との連携

分野
リポジトリ

メタデータ収集

データの流れ
利用の流れ

成果論文

研究データ

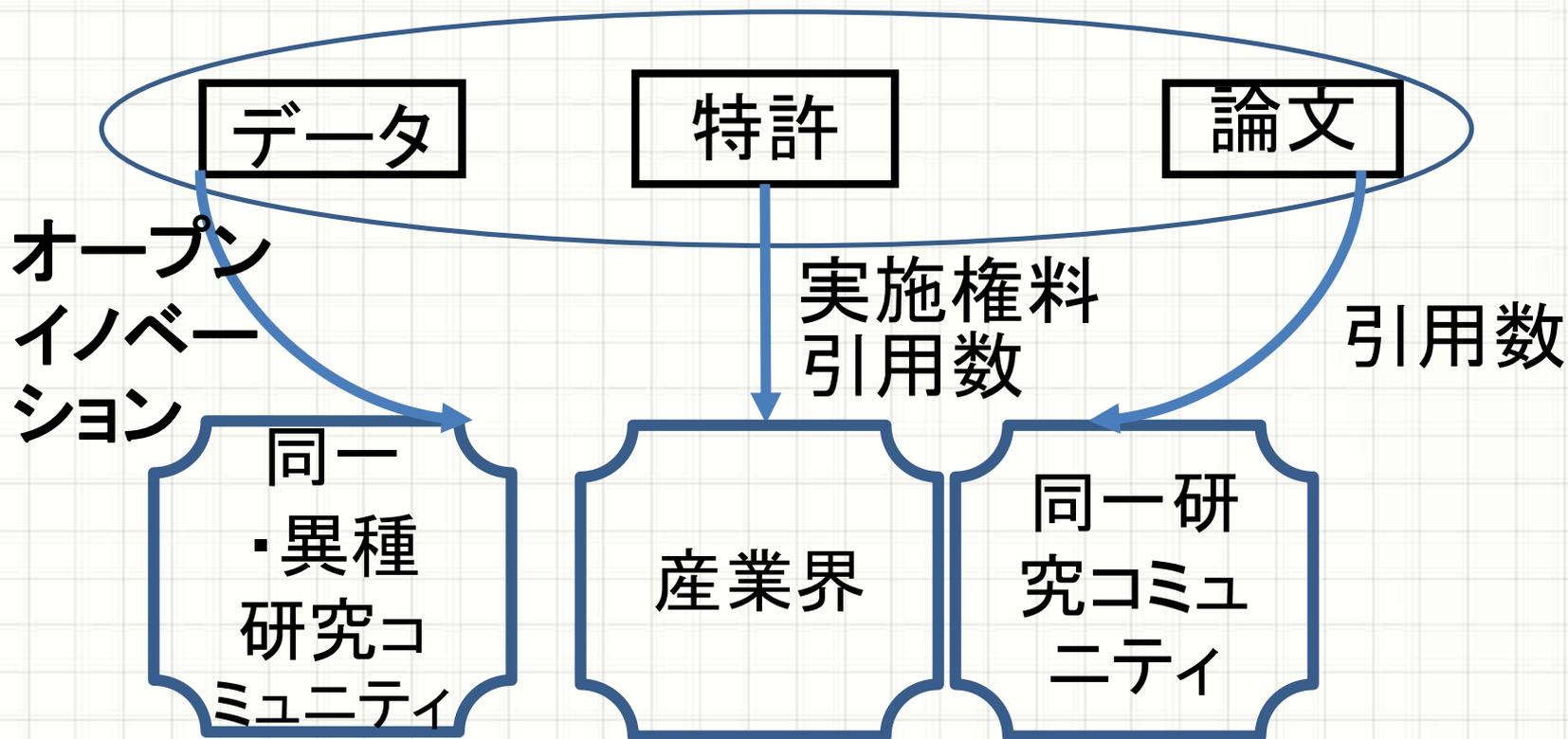
機関向け研究データ管理

研究データ用リポジトリ

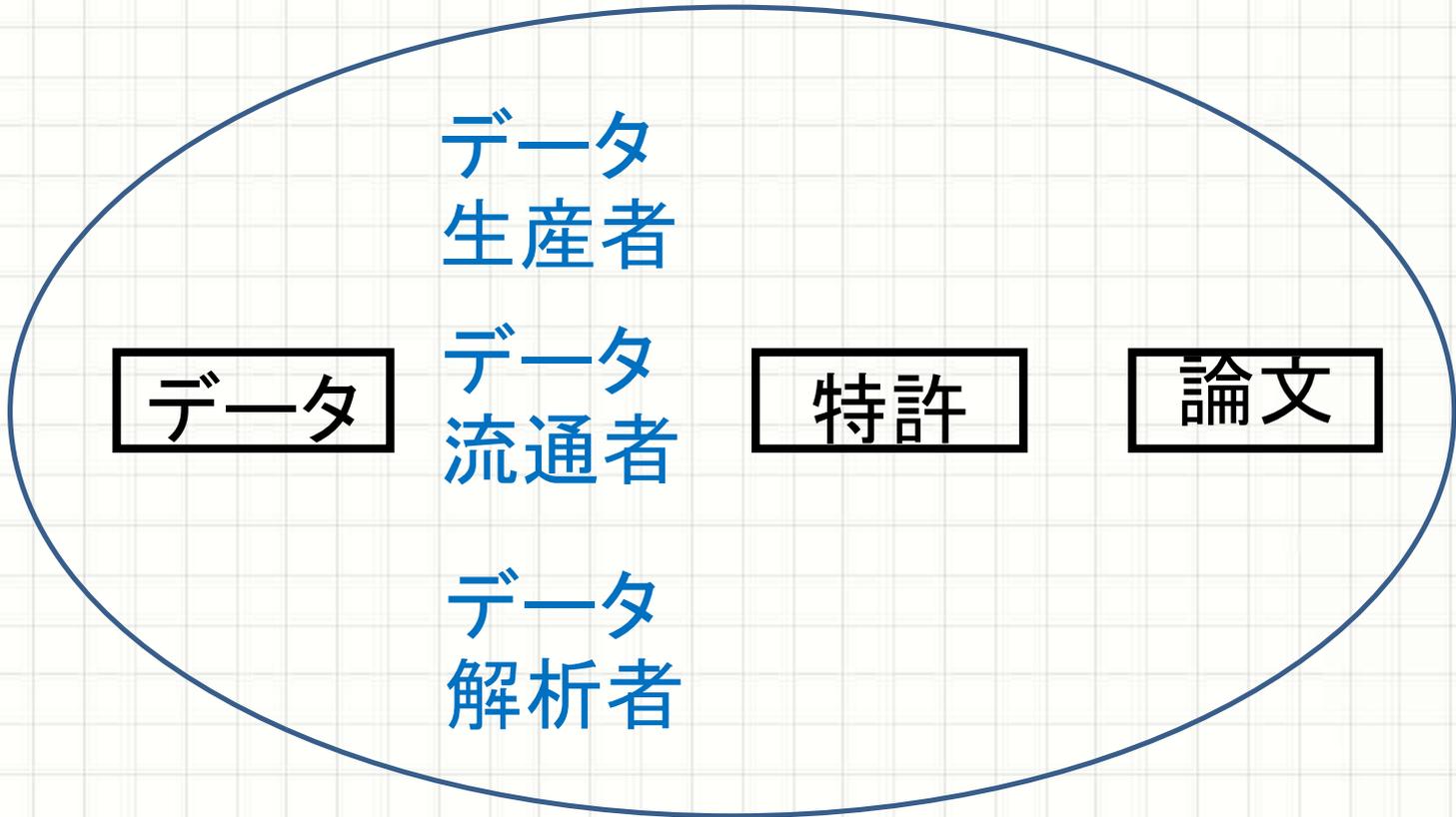
公開基盤

機関リポジトリのクラウドサービス
JAIRO Cloudを研究データに発展
させたサービス

研究のパラダイムシフト: 論文ありきからデータありきへ



研究のパラダイムシフト: 論文ありきからデータありきへ



研究のパラダイムシフト: 論文ありきからデータありきへ

個人情報保護法・
安全保障貿易管
理などの遵法性

データ

資金研究機関な
どとの契約

研究不正対策の
10年間保存

データ
生産者

データ
流通者

データ
解析者

論文

特許

マスタープラン2017に見る研究パラダイムシフト(1)

- 広域アジアを対象とした歴史地名・歴史名称を基軸とする**時空間情報基盤**の構築
- **公的統計マイクロデータ等**の研究活用のための全国ネットワーク整備
- 新しい社会科学としてのエビデンスベース人間科学の確立と**ネットワーク型大規模経年データ**の構築
- 新世代生物多様性・生態系**モニタリングのネットワーク**と拠点形成: 変動環境下における生態系機能の応答機構の解明とレジリエンスの向上を目指して
- 産官学協働研究のためのクラスター構築と**知の集積**による「食・農・バイオマス」地域資源利活用による地方振興ネットワーク形成
- 気候変動に対応する**ゲノム基盤**技術に基づく次世代ゲノム育種展開
- マリンビジョン・ネットワーキング計画: 地球環境変動に対応する**ビッグデータ解析**システム利用の広域沿岸水域生態系解析と海洋生物資源の持続的利用のための研究拠点の形成
- **統合情報**を駆使したスマート・メガスケール植物工場ネットワークによる国際競争力のある農作物生産
- **統合農業知**の可視化によるユーザーイノベーション実践科学
- 健康社会の創成に向けた**多次元脳・生体イメージングセンター**の構築: 大規模ネットワーク解析によるヒトの知能の理解、革新的疾患予防・治療法開発・創薬に向けて

マスタープラン2017に見る研究パラダイムシフト(2)

- **医学医療情報**の大規模統合解析と医用人工知能応用基盤の開発 研究計画
- **ビッグデータ**と人工知能で拓く健康長寿社会のためのシミュレーション科学
- **ゲノム**医科学研究拠点の形成
- 放射線医科学イノベーション創出に向けた**情報の統合と活用**
- **国民ポートフォリオデータベース**の構築とその分析による国民参加型健康管理の実現と医療資源の最適化を目指す、ライフスタイルサービスの創出と国家戦略
- 海陸・掘削統合観測による革新的地震・噴火予測科学 —沈み込み帯の**時空間情報科学**の挑戦—
- **e-サイエンス**に向けた革新的アルゴリズム基盤
- 行動情報学研究基盤整備計画
- IoT時代の**高度データ**処理を達成する先端ハードウェア向け組み込みシステム基盤
- **汎用的データ収集・利活用**による分野横断学術研究基盤
- ゼッタバイト時代における新たな価値創造を牽引する**超ビッグデータ**利活用社会基盤の研究計画
- **アカデミック・ビッグデータ**活用研究拠点の形成
- 高感性情報科学技術の高度化とそれに基づく**学術情報基盤**の構築

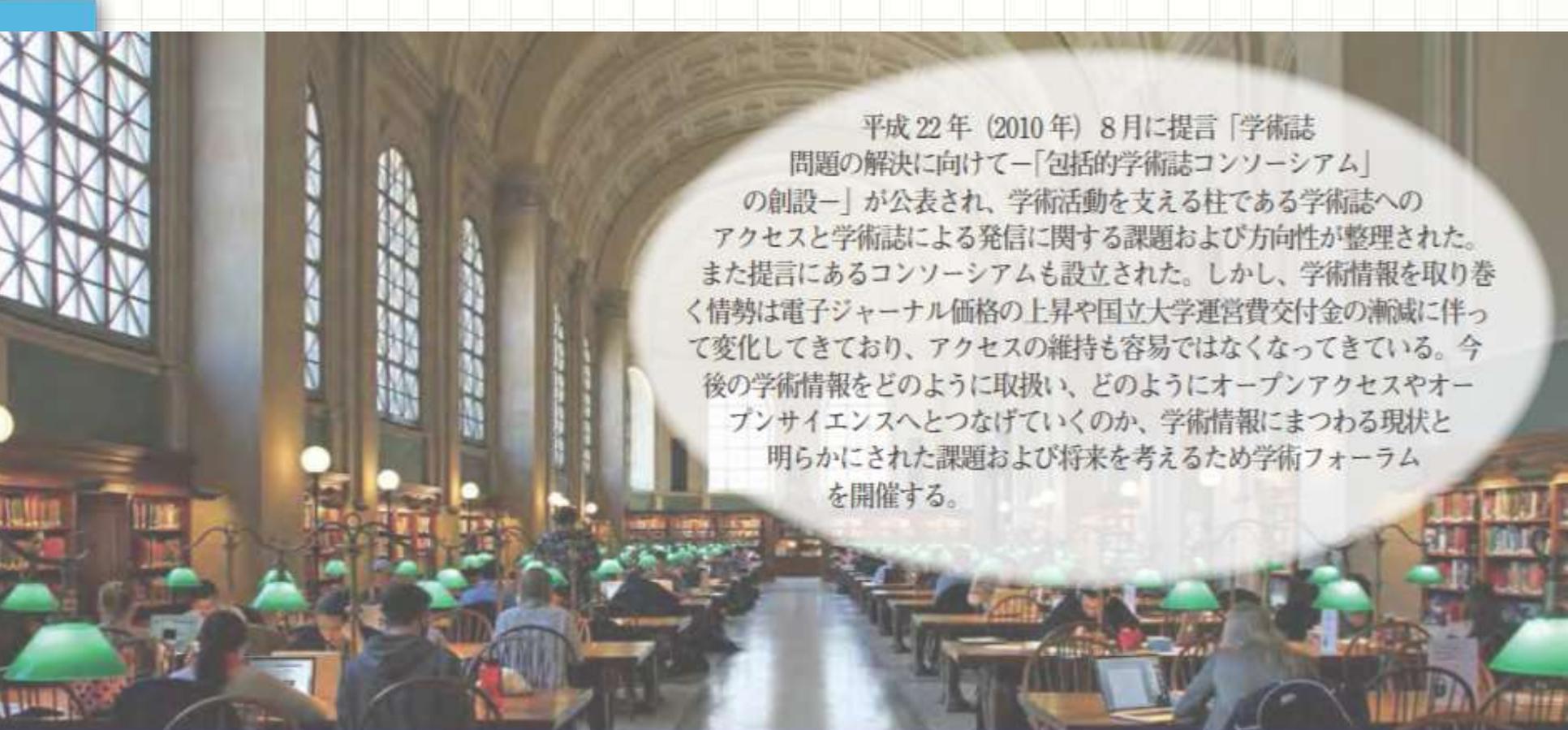
マスタープラン2017に見る研究パラダイムシフト(3)

- 学習ログの科学的分析に基づく高度教育情報基盤の開発
- **異分野の智**の結集により持続成長可能な社会実現を推進するハイブリッド新材料の創出
- 統合的リスク**情報システム科学**の確立と社会実装を加速するネットワーク型研究基盤構築
- 社会のインタラクティブ合意形成を実現する**知の統合プラットフォーム**研究開発拠点 KCP-Complex の形成
- 海洋環境の持続可能で安全な利用に資する**情報インフラ**の構築
- ヒューマンセントリック思考による経験知集約材料創製
- **電子ジャーナル・バックファイル等へのアクセス基盤**の整備
- 身体芸術の文理融合型学際研究と国際身体芸術**アーカイブズ**・コンソーシアムの設立
- 国立沖縄自然史博物館の設立－東・東南アジアの自然の解明と**ビッグデータ**自然史科学の実現－
- あらゆる分野の因果推論を支援する**データ**解析・可視化研究コミュニティの構築－1億総明晰社会実現に向けて－

マスタープラン2017に見る研究パラダイムシフト(4)

- **オープンサイエンス**推進のための研究データ基盤
- 融合社会**脳研究センター**構想
- 国家と社会のための科学技術総合マネージメント**データジャーナル**による異分野融合
- 日本語の歴史的**典籍の国際共同研究ネットワーク**構築計画
- わが国の産業・社会の基盤資源としての昆虫類の**生物情報データベース**および大規模標本の整備
- 「行間が読めて説明ができる」自然言語処理を実現するための**知識**と推論システムの構築

179件中39件(22%)



平成22年(2010年)8月に提言「学術誌問題の解決に向けてー「包括的学術誌コンソーシアム」の創設ー」が公表され、学術活動を支える柱である学術誌へのアクセスと学術誌による発信に関する課題および方向性が整理された。また提言にあるコンソーシアムも設立された。しかし、学術情報を取り巻く情勢は電子ジャーナル価格の上昇や国立大学運営費交付金の漸減に伴って変化してきており、アクセスの維持も容易ではなくなっている。今後の学術情報をどのように取扱い、どのようにオープンアクセスやオープンサイエンスへとつなげていくのか、学術情報にまつわる現状と明らかにされた課題および将来を考えるため学術フォーラムを開催する。

危機に瀕する学術情報の現状とその将来

日時：平成29年5月18日(木) 13:00～18:00

アクセス：日本学術会議(東京都港区六本木7-22-34) 東京メトロ千代田線「乃木坂駅」下車5番出口(青山霊園方面)より徒歩1分

お申込み：日本学術会議ウェブサイト(<https://form.cao.go.jp/scj/opinion-0003.html>)にて

平成29年5月17日(水)17時までにお申込み下さい。参加無料、定員250名、先着順

お問合せ先：日本学術会議事務局 企画課学術フォーラム担当、電話：03-3403-6295

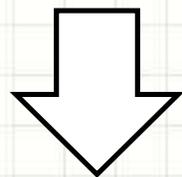


日本学術会議主催
学術フォーラム

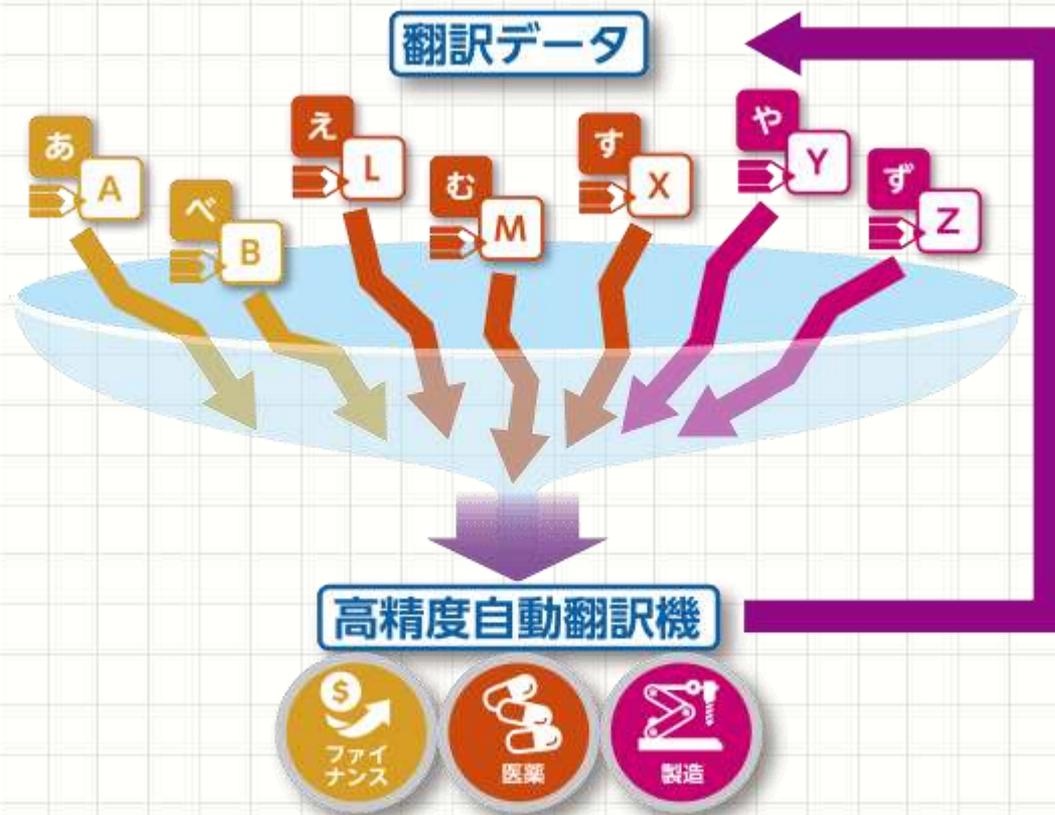
PROGRAM

- 学術情報とその将来に関する学術フォーラムのpart2を来年4月頃を開催する

対訳コーパスは散らばっている



一か所に集める



翻訳バンクの製薬分野での事例① ～アストラゼネカ～

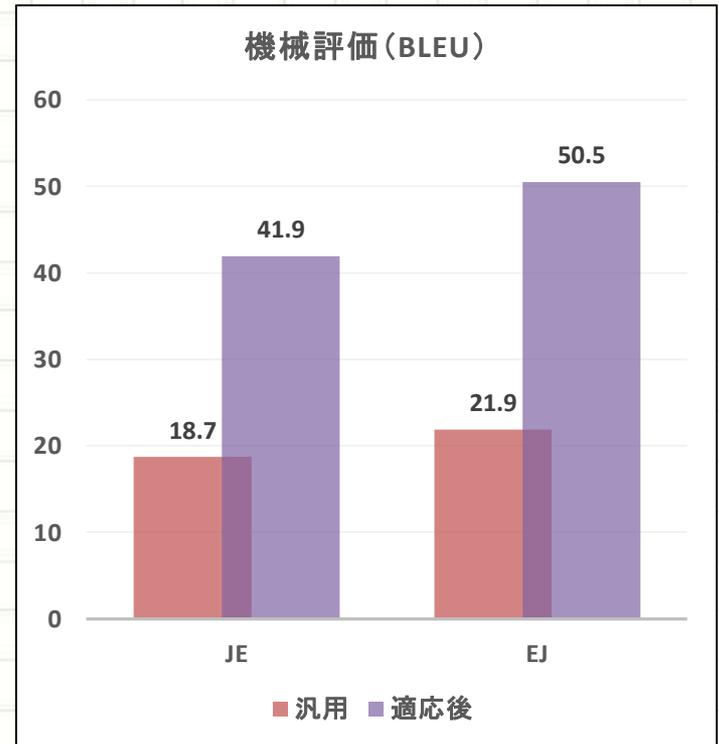
AI自動翻訳を研究

アストラゼネカ 情通機構と協力

アストラゼネカ(大阪市北区、ステファン・ウォックストラム社長、06・4802・3600)は17日、情報通信研究機構と協力して医薬業界向けに人工知能(AI)を使った自動翻訳システムを研究すると発表しました。同機構を持つAI翻訳エンジンを医薬業界向けに最適化し、実用性を検証する。申請関連の書類を英語から日本語に翻訳する効率を高め、国内に新薬を申請する期間の短縮につなげる狙い。

アストラゼネカは現在、60を超える新薬の開発プロジェクトを国内に抱える。このうち80%以上が、薬の効果や安全性を複数の国や地域で同時に検討する国際共同治験。このため英語の資料を日本語に翻訳する機会が増えている。

翻訳の手間から国内申請が滞ると、海外と比べ新薬の普及に時間がかかる懸念がある。このため翻訳の効率化が求められていた。



<https://www.nikkan.co.jp/articles/view/00470047>

翻訳バンクの製薬分野での事例②

～MSD～

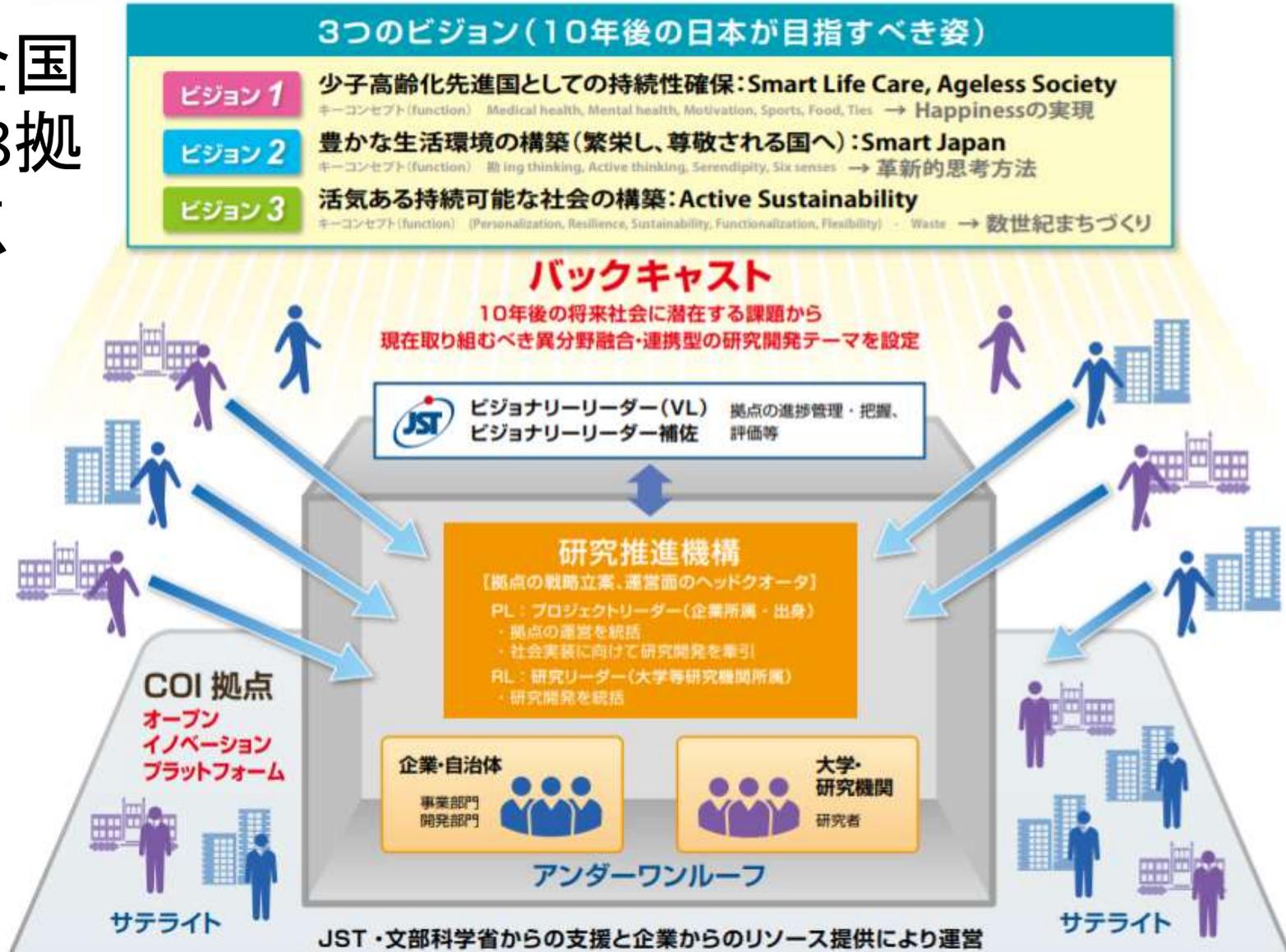
原文1	Nonspecific symptoms of cancer (e.g., weight loss, anorexia, fever) often develop first.
Before	ガンの症状は まず 体重減少や食欲不振、発熱などが起こる。
After	癌の非特異的な症状(体重減少、食欲不振、発熱)が最初から CNS 由来の中枢神経機能不全に起因する。
原文2	The main symptoms are weight loss, anorexia, and fever.
Before	高ナトリウム血症の主な徴候は、脳細胞の収縮による中枢神経系の機能不全に起因する。
After	高ナトリウム血症の主な徴候は、脳細胞の収縮による中枢神経系の機能障害に起因する。



JST COI(センター・オブ・イノベーションプログラム)にみるオープンサイエンスの例

拠点の推進イメージ

- 全国18拠点



弘前大学COI拠点の全体像

《ヘルスケア分野に革新をもたらす3本の戦略的研究課題設定》

I 健康ビッグデータを用いた
疾患予兆法の開発

II 予兆因子に基づいた
予防法の開発

III 認知症サポートシステム
(意思決定支援)の開発

『AI等最先端科学研究(超多項目健康BD解析)』×『地道な健康教育・啓発活動(環境づくり)』の融合

※研究フィールド「青森県」は日本一の短命県:課題先進地域

岩木健康増進プロジェクト

1人あたり**2000項目**の超多項目健康ビッグデータ

いきいき健診プロジェクト

65歳以上高齢者**2400人**の健康データ(認知症)

腸内細菌
口腔内細菌

弘前大学が保有する**世界無二の超多項目健康ビッグデータ**で、予兆から予防、行動変容までトータルでの革新的な研究開発を行う。

BigData



認知症
生活習慣病

疾患危険因子の特定

疾患予測アルゴリズム

健康
教育・啓発

短命県返上+健康長寿社会の実現



×

50年以上に及ぶ世界的な**九大・久山町**研究

世界的長寿エリアでの**新・京丹後スタディ(1000名)**

健康人の**2000項目健康ビッグデータ**をもつ弘前大学だからこそできる革新的チャレンジ!

「健康づくり×健康寿命延伸×まちづくり」に経済活動(BIZ)を合体・融合させ、「**真の社会イノベーション**」を創造する

世界的に類をみない、健康人の超多項目健康ビッグデータ

《構造イメージ図》

分野の垣根を越えた

多因子的解析を可能にする網羅的データ

2,000項目 × 1,000人 × 14年

ほぼすべての関係者が
集結する『場』

医学部各講座、他学部(人文・理工他)、他大学・研究機関、公的機関、企業、自治体、市民

医学部各講座、
全学部等が幅広く関与

Iwaki Pure Big Data

【岩木プロジェクトデータ項目(2,000)】

1 ゲノム解析 2 フリーラジカル

1 体格・体組成 2 内臓脂肪 3 体力 4 栄養状態 5 歩行速度
6 巧緻性検査(ペグボード) 7 重心動揺検査 8 歩行分析(アコムアイ)
9 筋力 10 骨密度・骨代謝 11 関節 12 頸椎MRI 13 肝・胆
14 内分泌 15 消化器系 16 呼吸器 17 心臓(心エコー)
18 脈管(動脈硬化) 19 腎 20 視力・眼底検査写真 21 聴力検査
22 泌尿器 23 神経内科(認知) 24 神経系 25 皮膚科診察
26 アレルギー 27 口腔衛生 28 腸内細菌 29 呼気ガス 30 微量元素
31 アミノ酸分析 32 脂肪酸分析 33 ベントジシン 34 ホモシステイン
35 アディポネクチン 36 レプチン 37 セロトニン 38 コルチゾール
39 PAI-1 40 フィブリノゲン 41 FDP 42 エコール 43 メタボローム解析
44 免疫・炎症 45 酸化ストレス(ビタミン、カロテノイド、8-OHdG)

1 転倒 2 睡眠 3 食事 4 口腔ケア状況 5 飲酒
6 喫煙 7 病歴・服薬 8 服用コンプライアンス
9 生殖、母子手帳
10 生活の質QOL:SF-36(身体機能、日常役割機能(身体・精神)、
体の痛み、社会生活機能、全体的健康感、活力、心の健康)
11 PHCS:自覚的健康管理能力(ヘルスプロモーションを主眼とした健康関
連習慣や健康関連行動の変容に関する評価指標)
12 ボディーイメージ 13 民間医療利用

1 職業・学歴: 農家etc 職業別METs- 中卒・高卒・大卒
2 家族: 同居人数、配偶者の有無
3 日常生活の状態等(会話、就業状況、年金受給状況、社会関連性、
生活時間など)

①
【遺伝学分野】
分子生物学的データ
(DNA)

②
【健康科学分野】
生理・生化学データ
(性別・血圧・体力・肥満・
共生細菌・診療データ)

③
【人文科学分野】
個人生活活動データ
(就寝時間・会話の頻度・食事・
趣味・ストレス)
プレゼンス
アイデンティティ

④
【社会科学分野】
社会環境的データ
(労働環境・経済力・学歴)

一つの測定項目と他の2000項目との
関連性が検討できるメリット!

例)1,000名の腸内細菌データだけでは
大きな意味を持たないが、2,000項目と
の関連性ではイノベティブな知見をもた
らす。

⇒ 性、年齢、肥満、体格、体力、動脈硬
化、喫煙、飲酒、食生活、運動、便通、口
腔内細菌、口腔環境、ピロリ菌、認知症、
糖尿病、高血圧、アレルギー、肝機能、腎
機能、心機能、肺機能、睡眠、微量元素、
呼気ガス、免疫能、ストレス、ロコモ、メタボ、
好中球機能、リンパ球、サイトカイン、ビタ
ミン、ホルモン、脂肪酸、アミノ酸、服用薬剤、
泌尿器疾患等との関係が明らかになる。

《分野・領域横断的な研究体制》

※岩木ビッグデータではひとりの人間の分子生物学的データから社会環境的データまでをすべて関連づけた網羅的解析が可能
※14年間実施し延べ“約2万人”以上。H28-29は別途65歳以上高齢者2.4千人対象とした認知症健診も実施。

多大学間連携による最強BD解析チームが本格稼働へ

※バイオインフォマティクス・生物統計・臨床統計分野の第一級専門家が一大集結

“(データの)標準モデル化” “予防・先制医療” “個別化医療・精密医療” “健康社会モデル”

京都大学



ビッグデータ解析による 新たな疾患概念の構築

岩木ビッグデータの統合的な解析手法を開発し、従来の疾患概念とは異なる新たな疾患概念を構築することにより、次世代の個別化医療・予測医療の実現を目指す。

大学院
医学研究科
ビッグデータ
医科学分野

奥野 恭史 教授
(BD解析タスクチームリーダー)

DB構築 ▶ 疾患予測AI開発へ

腸内細菌と肥満の関連解析へ

東京大学



ビッグデータを用いたヒト 属性と腸内細菌組成と の関連解析

超多項目の健康ビッグデータの特性を生かし、ヒトの属性と腸内細菌組成に注目して解析を行う。

大学院
医学系研究科
生物統計学分野

松山 裕 教授

弘前COIを中心とした 強力なBD解析体制



統括: 中路特任教授(弘前大)

東京大学



ゲノムデータ解析による 疾患予兆発見の開発

スーパーコンピュータを用いた大規模ゲノム解析を活用し、ゲノムデータと生活習慣・社会環境データを対応させることにより、革新的な疾患予兆法を開発する。

井元 清哉 教授
(BD解析タスクチームサブリーダー)

医科学研究所
ヘルスインテリ
ジェンスセンター
健康医療データ
サイエンス分野

名古屋大学



ビッグデータ解析による 生活習慣病の予兆法開発

岩木ビッグデータをその特性を生かしながら統計学を活用して解析し、効果的な生活習慣病の予兆式を開発して、早期発見・予防につなげる。

医学部附属病院
先端医療・
臨床研究支援センター **中枋昌弘 病院講師**

東京大学



ビッグデータ解析による MCI予兆法の開発

岩木ビッグデータをその特性を生かしながら統計学を活用して解析し、生活習慣・社会環境も考慮した革新的なMCIの予兆式を開発して、MCIの早期発見・予防につなげる。

大学院
医学系研究科
生物統計
情報学講座 **平川晃弘 特任准教授**

認知機能と睡眠の関連解析へ

生物統計 × バイオインフォマティクス × 機械学習 × スーパーコンピューティング × シミュレーションの融合

参画企業との戦略的データ共有・共同解析を本格展開

《COI参画企業との共同研究によるBD解析が本格化》



※参画企業・大学のデータ解析環境(共有化システム・制度)を整備し、本格的運用開始

弘前COIをコアとした**戦略的多拠点間データ連携**の加速化へ

※『**九大×弘大×京医大**』をコアに、+ 2 (名桜・和歌山) + 3 (順天他)を強力推進

※コアとなる連携基盤を固めつつ、次なる連携先(名大・東北大他 + α)と交渉加速化

《COI健康・医療データ連携機構PJ》 大学間(データ)連携の拡大イメージ



他機関

※他拠点の成果物を、本学フィールドでの実証連携することも可能