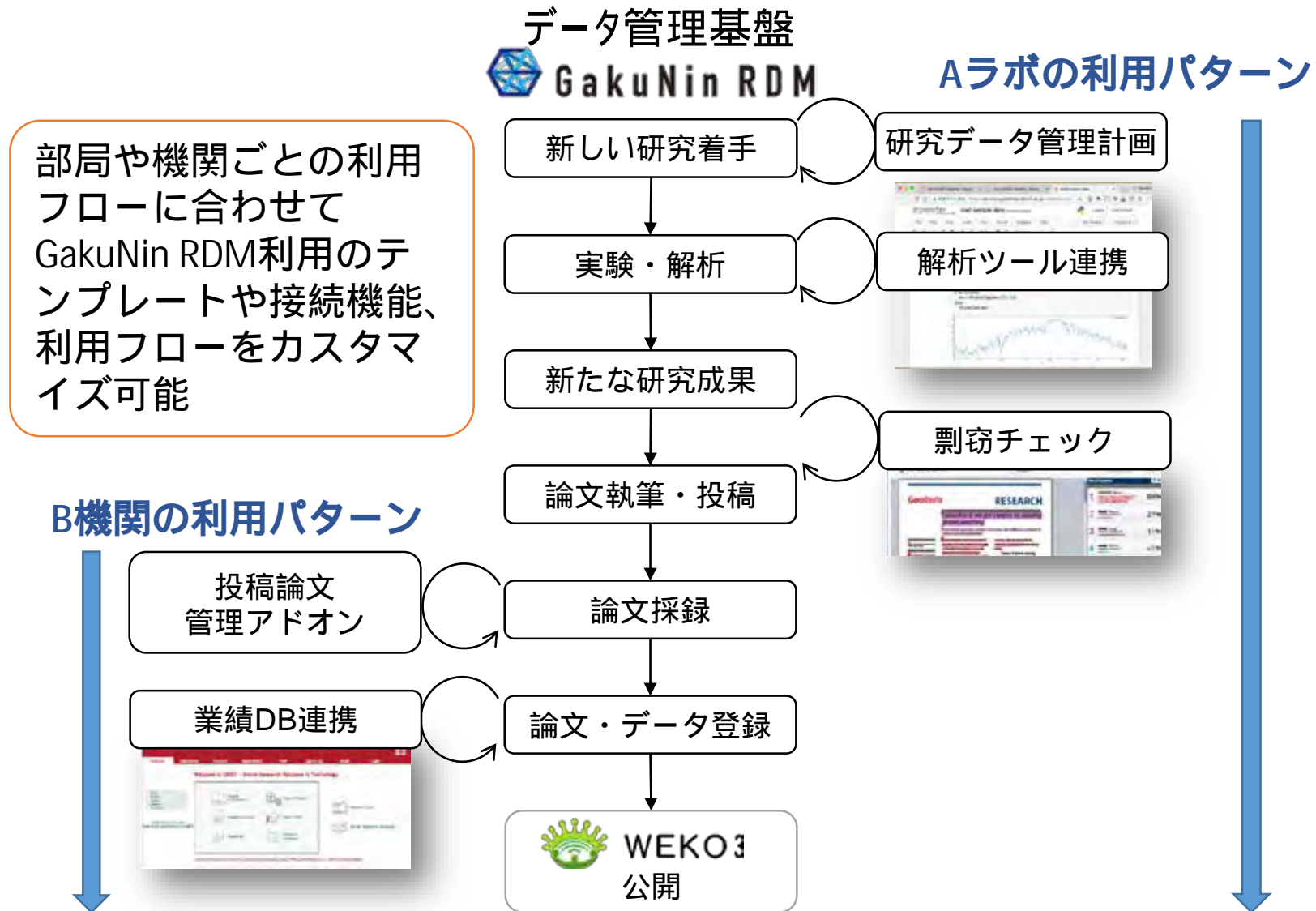


## 機関の利用モデルに沿った展開



# 研究データ基盤システム（データ管理基盤）の運用状況



**北海道大学**  
GRDMを学内オンプレストレージ  
NextCloudに接続して、研究データ  
管理のための適切な環境を構築



**名古屋大学**  
GRDMを利用して、東海国立  
大学機構としての地域の研究  
データ管理のフレームワークを  
確立



**金沢大学**  
機関としての研究データ管理のため、  
実験的にGRDMを利用中



**京都大学**  
機関としての研究データ  
管理のため、実験的に  
GRDMを利用中



**東京大学**  
GRDMのデータ保管機能、共有  
機能を使って、論文のエビデンス  
データを管理するための適切な  
環境を構築（研究公正管理）



**九州大学**  
電子ラボノートとGRDMを  
繋いで、ラボノート中の研究  
データを管理するための  
適切な環境を構築

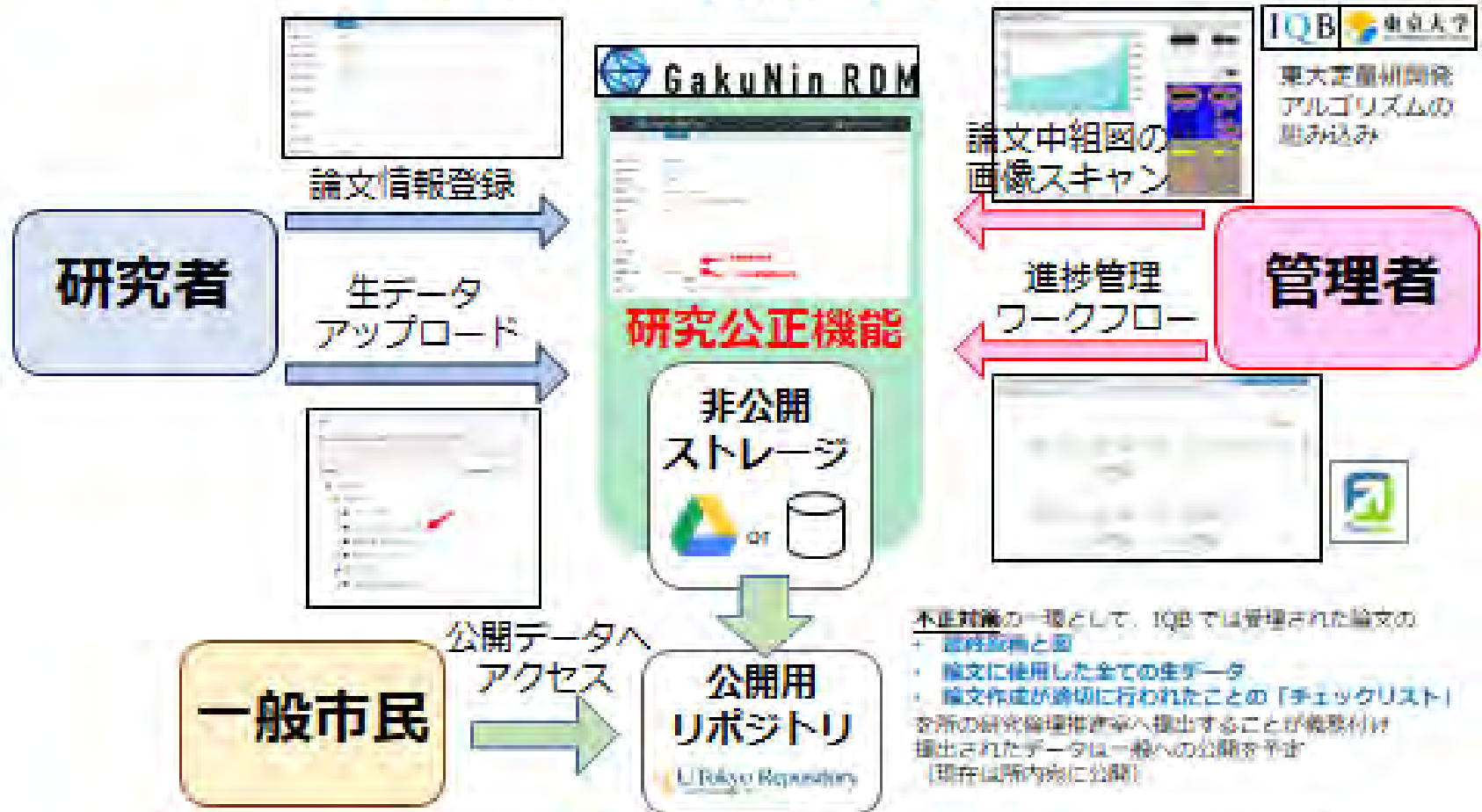
これら6大学以外にも北見工業大学が帯広畜産大学、小樽商科大学とのデータ共有にGRDMの利用を検討  
また、広島大学はDropboxにGRDMを繋いで、「研究証跡確保の機能」の利用を検討

**なお、JAIRO Cloudは 605大学・機関が利用（2019年6月5日現在）**

# 東京大学（定量生命科学研究所）での運用例

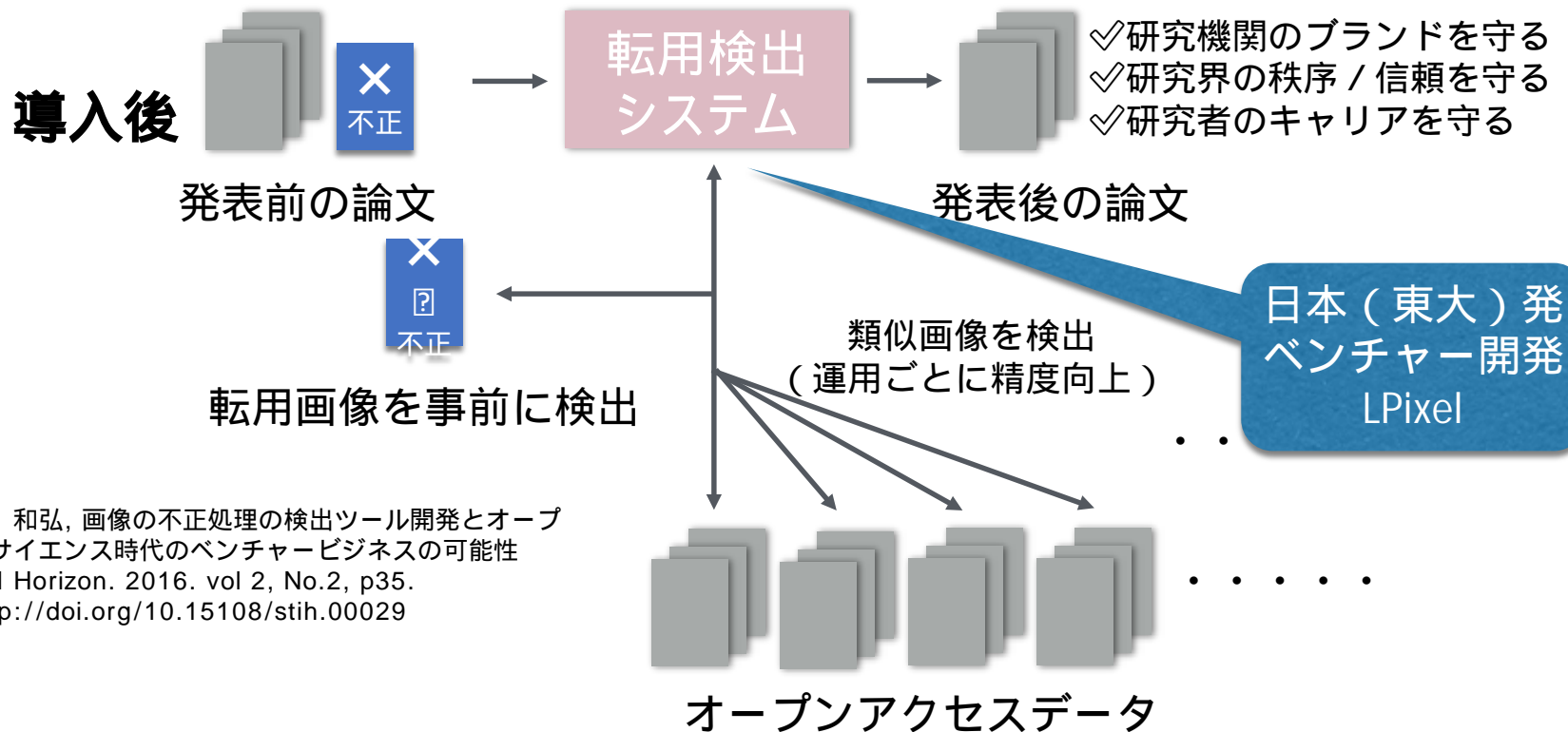
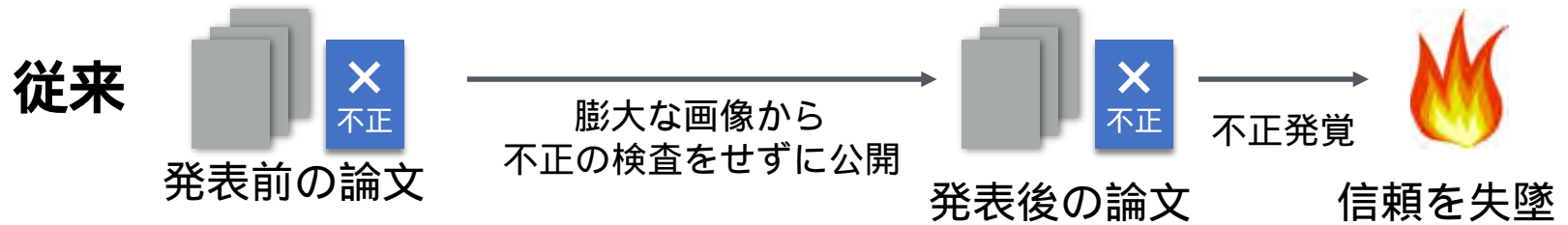
## GakuNin RDMベースの研究公正管理システム

RIMS: Research Integrity Management System



東京大学 定量生命科学研究所  
『全論文データの登録・保管・公開を支援するフレームワークの開発』

# 研究データの蓄積を活用した不正防止の実例 (画像の剽窃を防ぐ新たな安心チェック体制)



林 和弘, 画像の不正処理の検出ツール開発とオープンサイエンス時代のベンチャービジネスの可能性  
STI Horizon. 2016. vol 2, No.2, p35.  
<http://doi.org/10.15108/stih.00029>

# ムーンショット研究開発プログラムにおける先進的な研究データマネジメント

【統合イノベーション戦略 2019（令和元年6月閣議決定）】

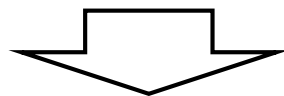
## 第 部 7 . データ基盤の構築

○ 公的資金による研究開発活動により生み出されたデータを適切に収集・管理・公開する体制構築を本格的に開始する。具体的には、研究データ基盤システムの整備、ムーンショット型研究開発における先駆的活用等による先進的なデータマネジメントを推進する。研究データマネジメントに必要な人材の育成・確保を推進する。

## 第 部 第 1 章 知の源泉（2）研究データ基盤の整備・国際展開

○ 公的資金による研究活動により生み出された研究データの利活用とそれによる先進的な研究マネジメント支援を促進するため、ムーンショット型研究開発制度において、先行的に研究データ基盤システム\*の活用を図るなど、先進的なデータマネジメントを推進する。【科技、文、経】

\* 別項目で研究データ基盤システム（NII Research Data Cloud）と規定。



## 先進的なデータマネジメントのイメージ（例）

- プログラム・マネージャー（PM）のマネジメントの下で、研究のプロセスで生まれたデータの共有・一部公開等を柔軟に行うことにより、新たな知の結合を行い、独創的な研究やイノベーションの創出につなげる。
- いわゆる「失敗を許容するマネジメント」のために、論文の仮説を立証するものではないが、一定の体系性を持つ実験・観測データ等を、PMが把握できるようにする。
- クラウド領域、共有・公開システム等の機能を導入。

内閣府、文科省、経産省、JST、NEDO及びNIIとの間で継続的に協議。

# 各研究機関におけるデータポリシー

ガイドライン項目番号	記載項目など	海洋研究開発機構 データ・サンプルの取扱いに関する 基本方針 (データポリシー)	国立環境研究所 データの公開に関する基本方針 (データポリシー)	宇宙科学研究所 データポリシー	物質・材料研究機構 研究データポリシー
	制定日 構成	2007年5月 1. 目的 2. データ・サンプルの定義 3. データ・サンプルの帰属 4. データ・サンプルの管理、保管及び活用 5. データ・サンプルの知的財産としての取扱い 6. データ・サンプルの公開 7. データ・サンプルの処分	2017年4月 1. 目的 2. 公開するデータの範囲 3. データの品質・管理 4. データの帰属・利用条件 5. データの公開期間 6. 免責	2018年3月 1. 目的 2. 適用 3. 本ポリシーが対象とするデータの定義 4. 非公開データ、公開データの考え方 5. 公開データのポリシー 6. 公開データ利用の際のルール 7. データ保存の考え方	2018年8月 1. 目的 2. 対象とする研究データ (1)データの形式 (2)データの種類の 3. 基本的な考え方 (1)データ管理計画 (2)研究データの収集・保管 (3)研究データの品質の確保 (4)研究データの提供・供用 (5)研究データの公開 (6)研究データの廃棄 (7)研究データの帰属 (8)免責
3(1)	目的	成果として取得したデータ・サンプルの取扱いと科学的・教育的利用について述べる	研究活動を通じて取得・作成したデータの公開について基本事項を定める	研究開発を実施する際のデータの取扱いに関する基本方針を述べる	研究データの収集、管理、利活用等に関する基本方針を定める
3(2)	データ定義・範囲	<ul style="list-style-type: none"> <li>調査観測データ</li> <li>シミュレーションデータ</li> <li>画像</li> <li>図面等調査研究で得られた各種情報とそれらの記録</li> <li>調査研究で得られた標本</li> </ul>	<ul style="list-style-type: none"> <li>研究成果として公開したデータ</li> <li>公益性や社会的ニーズが高く公開が適当と判断したデータ</li> </ul>	<ul style="list-style-type: none"> <li>源泉データ</li> <li>観測データ</li> <li>工学データ</li> <li>実験データ</li> <li>シミュレーションデータ</li> <li>サンプルや観測データの分析データ</li> <li>データ利用に必要なツール、ソフトウェア、アルゴリズム、説明文書もデータに準ずる</li> </ul>	<ul style="list-style-type: none"> <li>公表された研究成果物（論文、データベース、ソフトウェア等）</li> <li>未公表の研究成果物（公表データのバックデータ等を含む）で利活用を図るべきもの</li> <li>材料データプラットフォーム（以下DPF）等で利活用するデータ</li> </ul>
	非公開・共有等の対象データ	<ul style="list-style-type: none"> <li>人類共有の財産として研究・教育などの利用に広く公開</li> </ul>	個人情報保護や産業技術情報保護の観点から公開が適当でないとは判断するデータは非公開	以下の場合に非公開とする <ul style="list-style-type: none"> <li>個人情報保護や公共の安全等に支障がある場合</li> <li>データ処理が不完全で公開に問題がある場合</li> <li>取得・作成した研究チームが一定期間占有利用する場合</li> <li>利用権を他機関等に付与する場合</li> </ul>	<ul style="list-style-type: none"> <li>公益性や社会的ニーズに鑑み原則公開（インターネット上）</li> <li>知的財産保護や個人情報保護などの観点から公開が適当でないものは非公開</li> </ul>
	公開・共有の制限事項	<ul style="list-style-type: none"> <li>科学的・教育的利用は無償</li> <li>産業利用は有償</li> </ul>	<ul style="list-style-type: none"> <li>データの無断転載、2次配布は別に定める場合を除き認めない</li> <li>データの種類、利用目的によっては有償</li> </ul>	<ul style="list-style-type: none"> <li>営利、非営利を問わず無償</li> <li>一部用途制限する場合有。データ加工利用時の加工の明記（政府標準利用規約（第2版）*）</li> </ul>	<ul style="list-style-type: none"> <li>情報セキュリティ、データの適正利用の観点から条件を付与する場合有</li> <li>第三者への提供等に関し、条件を付与することがある</li> </ul>

データポリシー配下の規定である「データ・サンプル取扱規程」において、公開を制限する場合については、（１）法令に違反するおそれがある場合、（２）取扱いに危険が伴うと考えられる場合、（３）個人のプライバシーを侵害するおそれがある場合、（４）相手先に適切な管理をする能力がないと考えられる場合、（５）相手先が無断で第三者に貸与又は譲渡するおそれがある場合、（６）国の利益を損ない、または機構の適切な業務運営を阻害するおそれがある場合、（７）その他機構が必要と判断した場合と定めている。

# 各研究機関におけるデータポリシー

3(3)	データ保存/管理/運用	国内外の研究機関・研究者が利用できるように適切に管理・保管、迅速で円滑に提供	<ul style="list-style-type: none"> <li>データの品質確保と適切な管理に努める</li> <li>データのトレーサビリティの確保に努める</li> </ul>	<ul style="list-style-type: none"> <li>原則すべてを可能な限り保存</li> <li>原理的に再現できないデータは長期保管</li> <li>源泉データは、再処理可能な状態で長期保管</li> <li>再現できるがコストのかかるデータはできるだけ破棄しない</li> </ul>	<ul style="list-style-type: none"> <li>機構内外から広くデータを収集しリポジトリで保管</li> <li>データの信頼性、正確性、機械可読性、トレーサビリティなど品質確保に努める</li> </ul>
	管理計画				データ管理計画を定め、収集・管理利活用の事項を示す
	セキュリティ				研究成果物等の取扱いや情報セキュリティ、個人情報保護等に関する機構の諸規定その他の関係規定に定める
	リポジトリ				材料DPFまたはデータ管理計画で指定するリポジトリで保管
3(4)	メタデータ	利用しやすい形での情報提供と利用者ニーズを取り入れる仕組みを作る		メタデータもデータに準じる	
	識別子			<ul style="list-style-type: none"> <li>データを簡単に見つけ、使いやすくするサービスを無償で提供</li> <li>永続的識別子を付与(DOI等)</li> </ul>	
	フォーマット			適切なデータ処理や説明を行ってデータを公開	電磁気的方式で記録されたもの
3(5)	帰属	特別な取り決めがある場合を除き機構に帰属	<ul style="list-style-type: none"> <li>研究所が取得、作成したデータは別に定める場合を除き研究所に帰属</li> <li>他機関との共同で取得されたものは機関間の取り決めによる</li> </ul>		<ul style="list-style-type: none"> <li>機構で取得されたデータは特別な定めがある場合を除き機構に帰属</li> <li>他機関と共同で取得されたデータは機関間の取り決めによる</li> <li>他機関から提供を得たデータは別に定める</li> </ul>
	知財	<ul style="list-style-type: none"> <li>対象データは機構が指定</li> <li>機構の「知的財産に関する基本方針」に基づき活用・管理</li> </ul>			機構の研究開発/知財戦略の観点からデータアクセスについて制限措置をとること有
	免責		研究所は公開データ利用に関する損害の責任を負わない	研究所はデータ利用者の行為の責任を負わない(政府標準利用規約(第2版)*に出版)	機構は機構の正当な手続きによるデータの取扱い及び利用に伴って生じる損害の責任を負わない
3(6)	公開猶予期間	一定期間優先使用可	合理的と認められる範囲で公開を遅らせる	<ul style="list-style-type: none"> <li>取得から約1年</li> <li>利用範囲と非公開期限を決定</li> </ul>	合理的な範囲で設定可
	公開時期	公開猶予期間終了後に速やかに公開	<ul style="list-style-type: none"> <li>可能な限り速やかに、かつ継続的に公開</li> <li>公開を打ち切る場合有</li> </ul>	公開できる範囲で長期間保管(最低30年)	<ul style="list-style-type: none"> <li>可能な限り速やかに、継続的に公開</li> <li>機構の判断で公開打ち切り有</li> </ul>
	引用		データを論文等に掲載する場合の引用を明記	出版の明示(政府標準利用規約(第2版)*に出版)	データに基づく成果の公表やデータ引用先表示を求める
ガイドラインに記載のない事項	データ廃棄	保管・管理に値しない場合や合理的な期間を経過したデータ・サンプルの処分(廃棄・譲渡)		<ul style="list-style-type: none"> <li>コストやリソースの観点から破棄する場合有</li> <li>比較的容易に再現可能なデータは破棄する可能性有</li> </ul>	情報資産の取扱いに関する関係規定に基づき、機構の判断で破棄する場合有
	その他	関係諸規定の整備と利用者へのルール順守を明記		研究所が協力して他機関が取得したデータについても当該機関が本ポリシーの趣旨を尊重して対応することを期待	

# 各研究機関におけるデータマネジメントの実例

		JAMSTEC		NIMS	AIST		理研
事業内容		津軽海峡東部海洋レーダーデータサイト	船舶観測データ・サンプリングの計測・取得	材料データプラットフォーム	人体形状の計測	歩行データの計測	FANTOMプロジェクト*
データ公開状況	生データ	一部公開：レーダーデータは利用申請をすることでダウンロード可能。水温データはWebサイトでグラフを公開、ダウンロード不可。	公開：航海・潜航に関する品質管理を施したデータ	データごとに適切な公開条件を付して公開（利活用を図るために登録されるデータ）	公開：人間計測データ	公開：歩行データ 非公開：転倒歴など	公開：独自技術（Cap-Trap法、CAGE法）より得られた完全長cDNA配列やRNAの5'配列データ
	加工データ	公開：レーダーデータから算出した海表面の流向・流速データ	-	データごとに適切な公開条件を付して公開（研究成果物として公表されたもの）	非公開：解析・二次データ 有償公開：統計量	非公開：転倒スコアなど	公開：哺乳類電子の機能情報（遺伝子の位置や発現機構など）データファイルならびに専用ビューワーを介して提供
	メタデータ	公開	公開	公開	公開	公開	公開
メタデータの内容		<ul style="list-style-type: none"> <li>データ取得日</li> <li>データ取得方法</li> <li>データ取得者</li> </ul>	<ul style="list-style-type: none"> <li>データ取得日</li> <li>データ位置情報</li> <li>データ取得方法</li> <li>データ処理方法</li> <li>データ処理者</li> <li>データ形式</li> <li>キーワード</li> </ul>	<ul style="list-style-type: none"> <li>データ作成日、公開日等</li> <li>公開と利用の条件</li> <li>データ作者情報</li> <li>データ来歴、装置、試料等の情報</li> </ul>	<ul style="list-style-type: none"> <li>データ取得日</li> <li>データ取得方法・項目</li> <li>実験対象者の性別・年齢</li> <li>データ形式</li> </ul>	<ul style="list-style-type: none"> <li>データ取得期間</li> <li>データ取得方法・項目</li> <li>実験対象者の性別・年齢・身長・体重</li> <li>データ形式</li> </ul>	<ul style="list-style-type: none"> <li>使用サンプルの詳細情報</li> <li>サンプルの分類情報</li> <li>配列データの取得方法</li> <li>データ形式</li> </ul>
データ公開猶予期間		30分ごとの準リアルタイム	航海メタデータは原則2ヶ月 データは原則2年。ただし、定常観測データは処理が済み次第公開	データポリシーおよびデータリポジトリ運用規則に沿ってデータ登録者が設定する期間	既に公開済み	既に公開済み	論文発表と同時公開（既に公開済み）
データ公開猶予理由		-	学会発表、論文作成、特許作成	学会発表、論文作成、特許作成など	-	-	論文作成のため

\* ) 理研主導による国際共同研究プロジェクト ( FANTOM = Functional ANnoTation Of Mammalian genome ) 。哺乳類を対象として、網羅的に遺伝子の機能 ( ゲノム中の遺伝子の位置や発現機構 ) を解析。



# 各研究機関におけるデータマネジメントの実例

	JAMSTEC		NIMS	AIST		理研
第三者機関によるデータ利活用の有償・無償基準	<ul style="list-style-type: none"> <li>・科学的・教育的目的利用は無償</li> <li>・産業利用は有償</li> </ul>	<ul style="list-style-type: none"> <li>・科学的・教育的目的利用は無償</li> <li>・産業利用は有償</li> </ul>	原則無料。データポリシおよびデータリポジトリ運用規則に沿って有償等の可能性あり	無償	無償	無償
公開媒体	Webサイト	<ul style="list-style-type: none"> <li>・Webサイト（メタデータ・データファイル）</li> <li>・オフライン提供（生データ、高解像度映像、サンプル）</li> </ul>	データリポジトリ	ホームページ	ホームページ	<ul style="list-style-type: none"> <li>・専用Webサイト（理研内）</li> <li>・国際塩基配列DBコレクション</li> <li>・公開リポジトリサービス</li> <li>・NBDC（JST内）の生命科学系データベースアーカイブ</li> </ul>
データ帰属先	<ul style="list-style-type: none"> <li>・レーダーデータはJAMSTECに帰属。</li> <li>・水温データの一部はJAMSTECに帰属、一部は観測を実施している他機関に帰属。</li> </ul>	原則JAMSTECに帰属	NIMSに帰属。ただし共同研究契約等がある場合はその契約に依る。	原則AISTに帰属	原則AISTに帰属	理研に帰属するが、データはCC BY 4.0で公開

## 表面流マップの準リアルタイム公開・・・漁業者にも有用な情報提供



### 津軽海峡東部海洋レーダーデータサイト

MIO Ocean Radar data Site for the Eastern Tsugaru Strait (MORSETS)  
<http://www.godac.jamstec.go.jp/morsets/j/top/>

津軽海峡東部海洋レーダーデータサイト  
 MIO Ocean Radar data Site for the Eastern Tsugaru Strait (MORSETS)

2018年11月02日 13時00分

流速 (cm/s) 0 20 40 60 80 100 120 140 160 180

津軽海峡東部海洋レーダーデータサイト

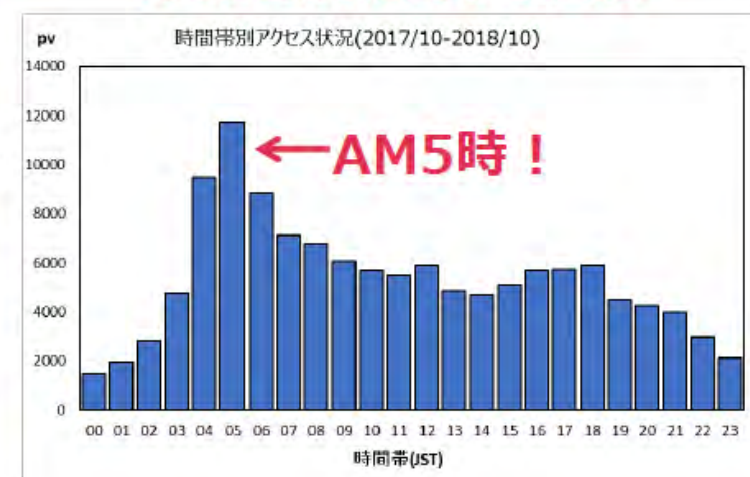
地点番号	地点名	水温(°C)	最終更新日時
①	朝根浜	17.7	2018年11月02日 12時00分
②	尻屋	17.9	2018年11月02日 11時00分



表面流マップは、研究データという位置づけのみならず、おそらく漁業者や現場を航行する船舶などにとっても有用な情報。



## Web公開に対する反応



# NIMSの研究データマネジメント事例

## アカデミア

- きらりと光る方法論
- 世界がハッとする成果

やり方・課題  
解決シナリオ

## 産業界

- ビジネス展開の主役
- 国際競争力強化

実験装置・計測機器からのデータ収集



文献からのデータ収集



テキストマイニング

大学等の成果からのデータ収集



機関レポジトリシステム  
(電子図書館)

アカデミックデータ収集  
(信頼できる標準データ)

材料情報統合データ  
プラットフォーム

高機能物質・材料データベース



高速サーバー  
(解析ツール)  
MI<sup>2</sup>データプラットフォームを核に実用版を構築



研究者・技術者コミュニティーによる研究推進

化学企業A  
(カーボン素材)

自動車企業B  
(車体素材)

鉄鋼企業X  
(高性能鋼)

電機企業Y  
(燃料電池)

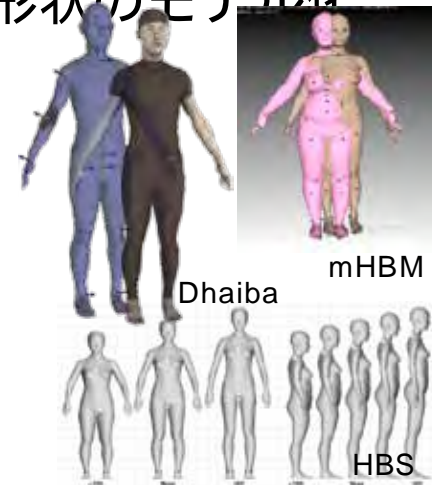


企業内データの  
徹底活用

# 人間系データバンク

## • 人体寸法・体形データベース

- 収集データを国際標準化ISO TR 7250-2 (世界の人体寸法DB)制定のインシヤティブ
- 国際協力により計測技術の供与とデータの交換で蓄積
- 3次元形状DB (収集) をもとに少ない測定点から人体形状のモデル化 (相同モデリング mHBM、体形統計処理 HBS)
- データを活用するシミュレータの構築(Dhaiba)



## • 歩行・歩容データベース

- 歩行研究で標準データ (裸足歩行) を計測 (n > 500)
- 年間200万件のアクセス
- コンサルティングによる支援
- 転びやすい人の歩行特徴の解明糸口
- 日本転倒予防学会推奨靴などの開発



## • データポリシー

- 人間計測データ(オープン), 解析・二次データ(クローズ)
- 歩行データ (オープン), 転倒歴, 転倒スコア (クローズ)

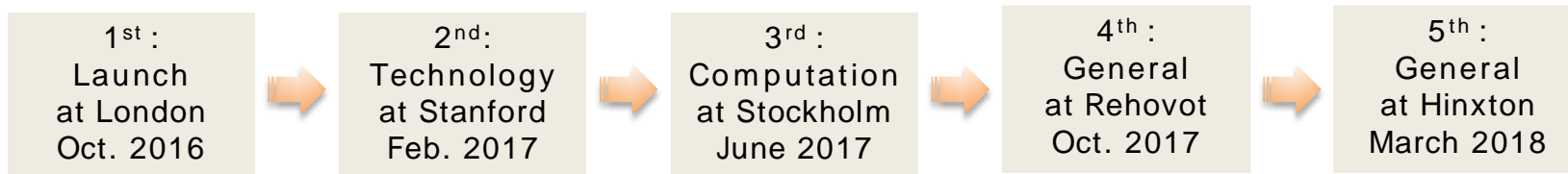




### 背景・目的・期待される成果

- ヒトゲノム情報の多くは細胞の集合（バルク）に由来する平均値的な情報であり、生命現象を本質的に理解し将来の医療をさらに発展させるためには、生体の究極の基本単位での数百種類にのぼる細胞種を「1細胞レベル」で解析することが必要とされる。
- こうした背景をふまえ、**ヒトの全細胞（約37兆個）について、種類・状態・系統などを最新の1細胞解析手法によって分類し、カタログ化することを目的とした国際プログラム「セルアトラス（ヒト細胞百科事典、HCA）プロジェクト」が欧米の主要な研究機関（欧州：サンガー研究所・カロリンスカ研究所、米国：ブロード研究所、等）を中心に提案され、参画機関の研究者間の会議で議論が進んでいる。**
- 本プロジェクトの推進のためには、系統的かつ高精度でヒトを構成する全ての細胞について一つ一つの細胞固有の遺伝子発現状態を網羅的解析を可能とするための技術として、**日本の持つ1細胞ゲノミクス解析技術が重要な役割を果たすことが欧米の研究機関から強く期待されている。**

各国で HCA Meeting が開催され議論が進んでいる



### 研究が進められている解析対象

米国では、NIHがこの二年間でセルアトラス関連のプロジェクトに約200億円の予算をを投じている。欧州でも、特定の臓器等を対象としたパイロットスタディが進行しており、わが国の参画が強く求められている。



- 免疫系
- 肝臓
- 腎臓
- 肺
- 皮膚
- 腸
- 心臓
- 神経系（等）

## FANTOMプロジェクトとは

FANTOM = Functional Annotation Of Mammalian genome

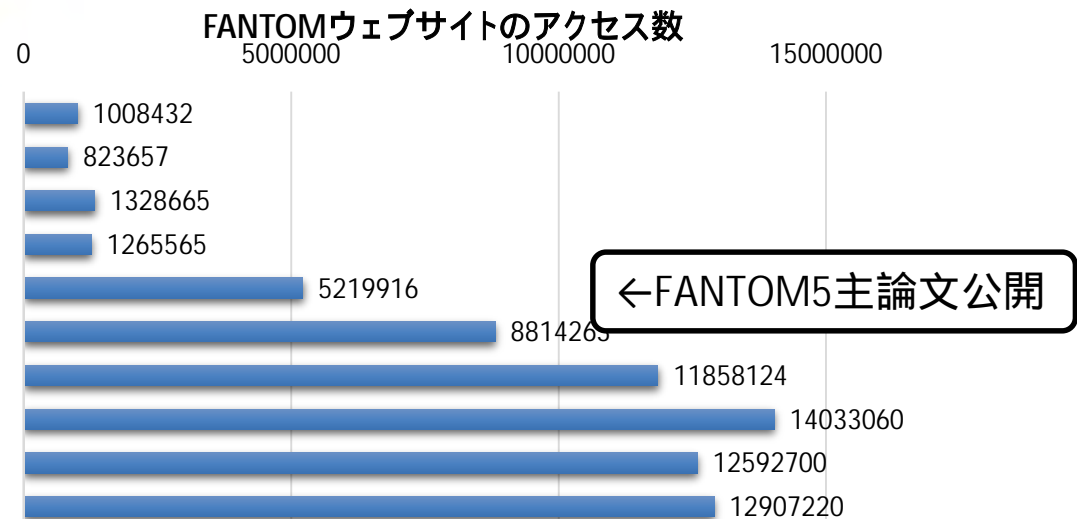
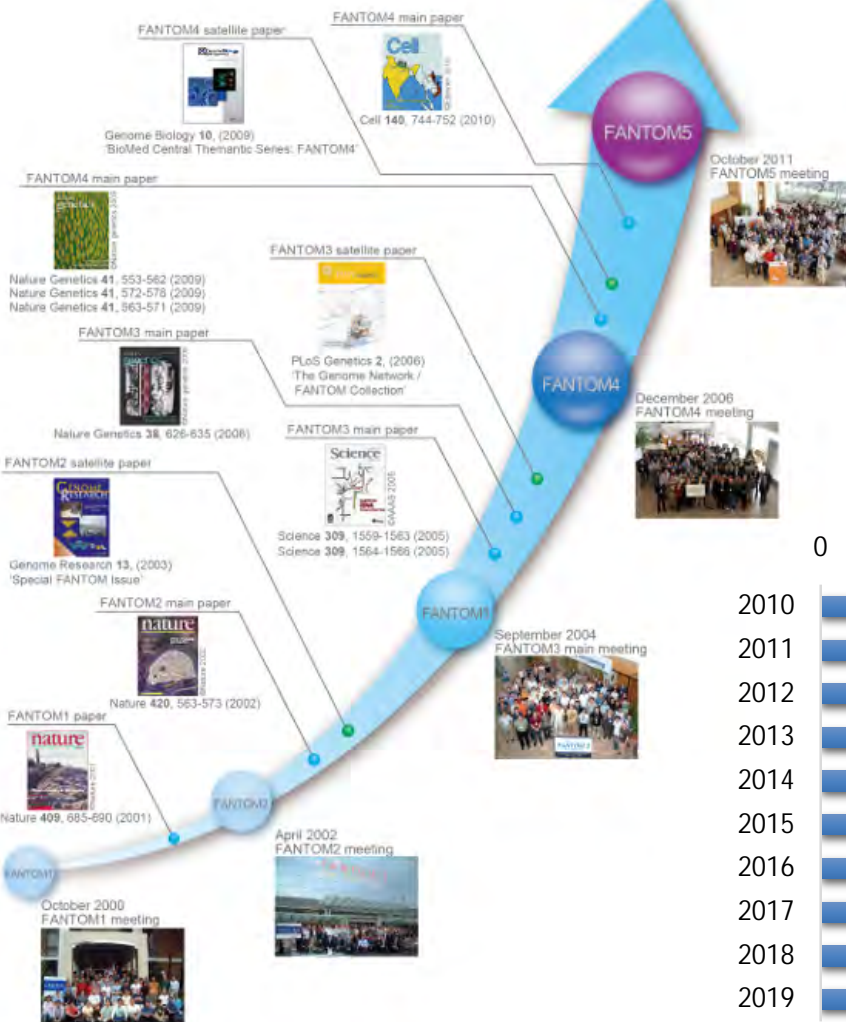
理化学研究所の主導による国際共同研究プロジェクト

過去5期に渡って実施

FANTOM1 ~ 3: 理化学研究所で配列決定したマウス完全長cDNAを用いた機能解析

FANTOM3 ~ 5: 理化学研究所の独自技術であるCAGE法を用いた転写制御についての解析

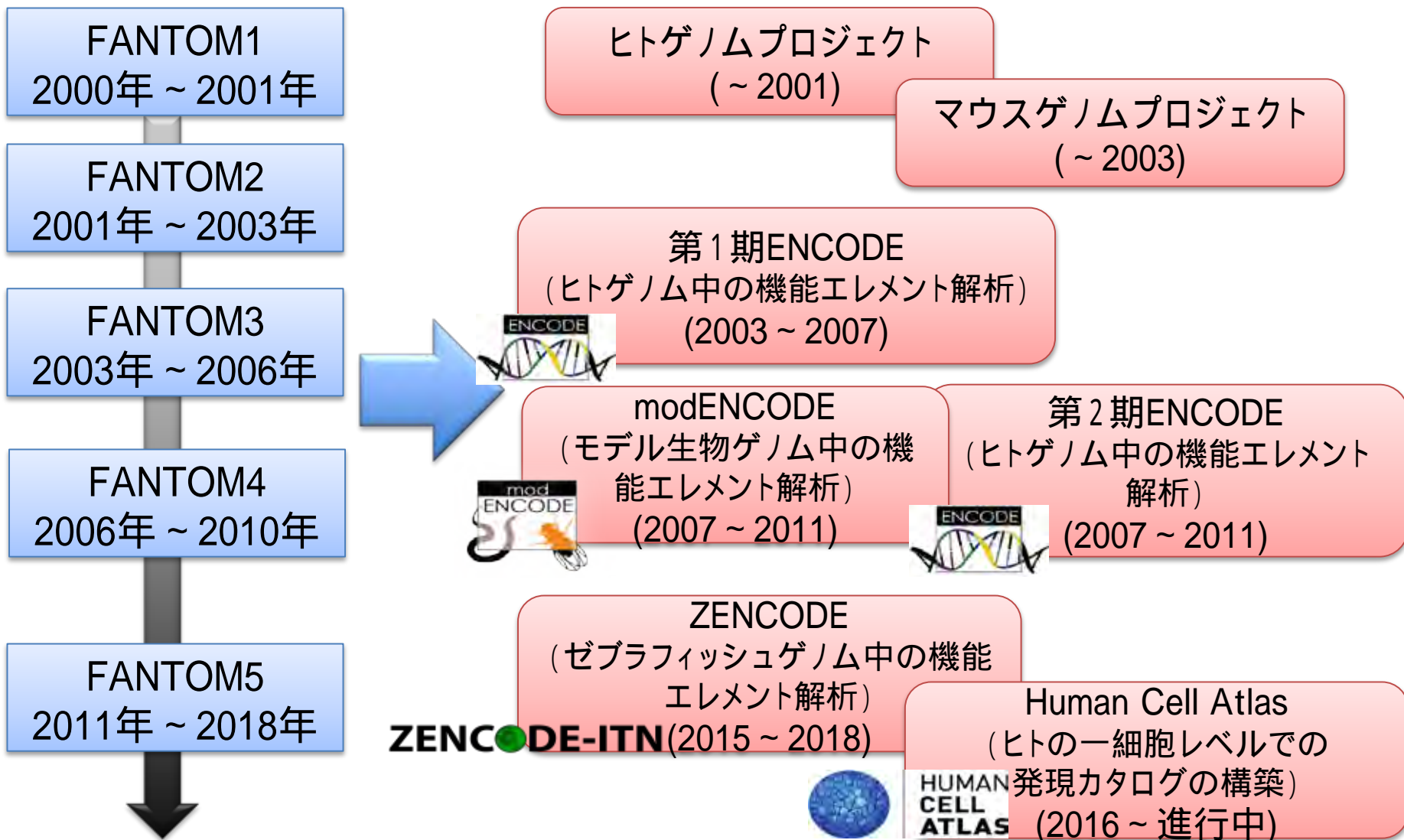
研究成果、解析に用いたデータやその解析結果はリソースとして全世界に無制限で公開  
(Creative Commons Attribution 4.0 International License)



2019年は1~3月実績より推定

出典：理化学研究所

## FANTOMプロジェクトの波及効果

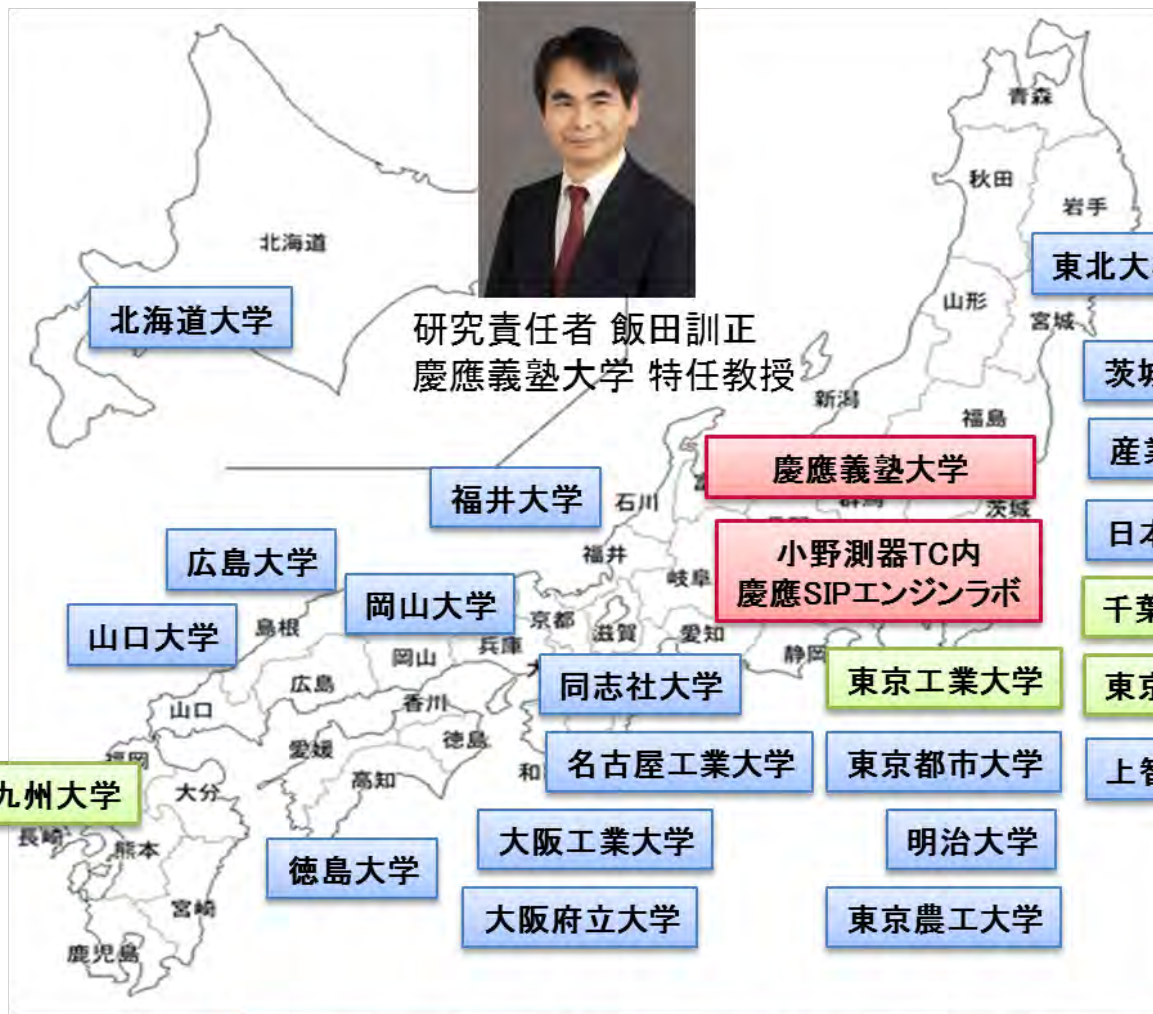


様々な国際共同研究プロジェクトに招待、参画へ

# SIP革新的燃焼技術のデータマネジメント事例（ガソリン燃焼チーム）



研究責任者 飯田訓正  
慶應義塾大学 特任教授

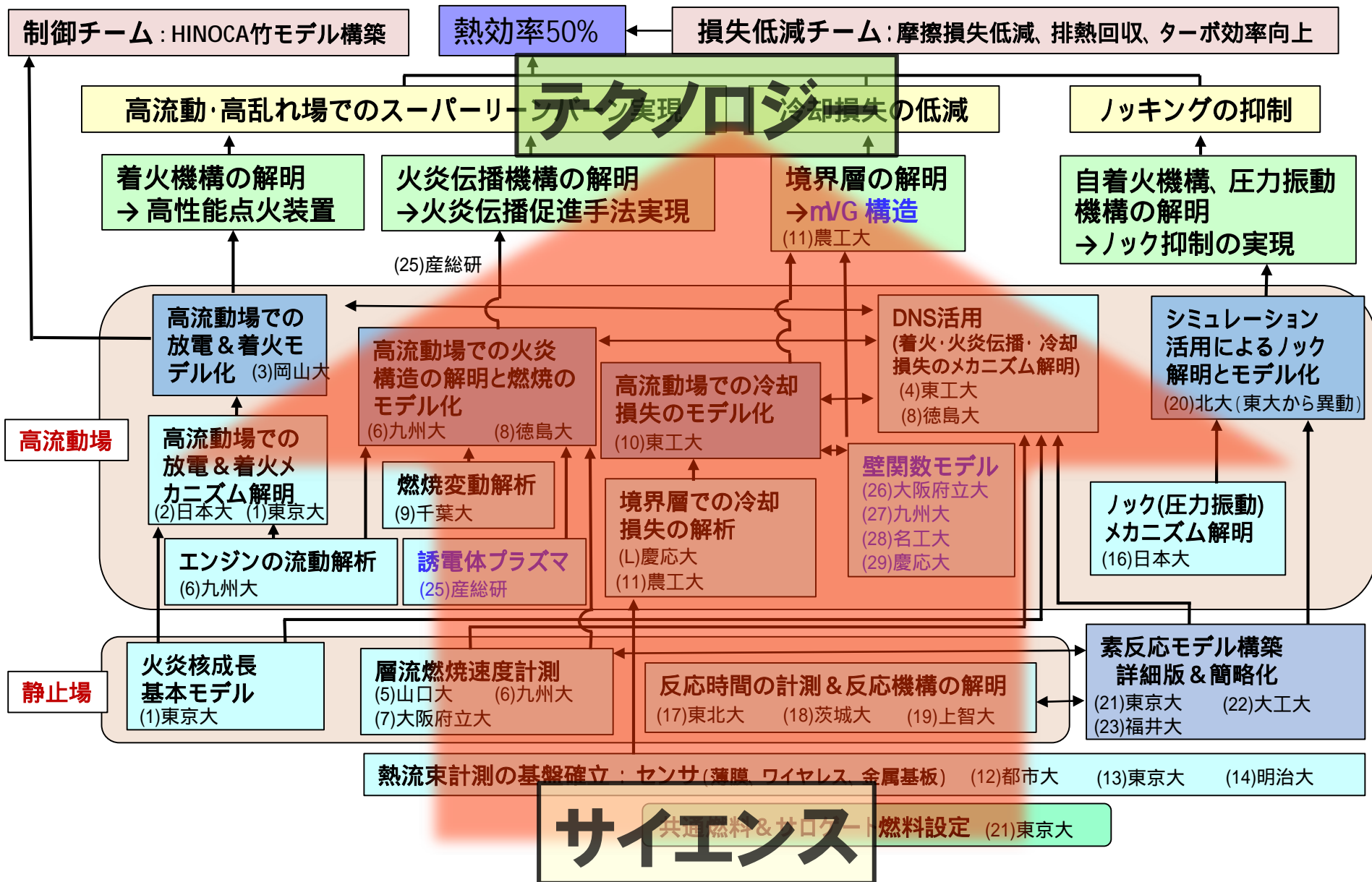


- リーダー大学, 研究拠点
- 班長大学
- クラスター大学・研究機関

・日本の知恵を結集し、世界を牽引する要素技術を創出  
 ・産業界の発展にも貢献



# SIP革新的燃焼技術のデータマネジメント：チーム内編成



# SIP革新的燃焼技術の研究データマネジメントから示唆されるポイント

1. 柔軟なアクセスコントロール	PMがプロジェクトメンバーの研究データへのアクセスを柔軟にコントロールできること
2. 幅広いデータへの対応	研究データの種類は、生データ、加工データ、メタデータのいずれも対応できること * ) プロジェクトメンバー間であれば、生データでもある程度理解できるため、生データであっても共有できること
3. コミュニケーションツールとの連動	遠隔にいる複数のプロジェクトメンバーどうしが、データを閲覧しながら円滑にコミュニケーションが取れること * ) プロジェクトメンバー間であれば、データを見ればある程度理解できるとは言え、やはり直接話してデータについて議論する機会を柔軟に設定できることが重要
4. ストレスフリーなメタデータ付与とそれに基づく検索機能	各データにメタ情報をストレスなく付与することができる仕組みと、それを基に効率よく検索できる仕組みがあること * ) ガソリン燃焼チームでは、研究者に対して、決められた画一的な項目のメタ情報を記入させるのではなく、1000文字のフリースペースを与えて、研究者が必要と思うキーワードを好きなように記入させる方法を採用
5. セキュリティ	十分なセキュリティが担保されていること
6. ユーザーフレンドリーな操作性	以上の仕組みがストレスなくユーザーフレンドリーであること * ) システムに不得手な研究者にいかにも直感的に使ってもらえるようにするのが重要
7. アジャイルな開発体制	以上のシステムを研究者に提供したときに、使っていく上で新たなニーズや不具合が出てくるので、これらにアジャイルに対応できる体制を整えること