

機密性と再現性は ガバナンスに含まれますか？

藤原 一毅

国立情報学研究所 クラウド基盤研究開発センター

2025-06-18

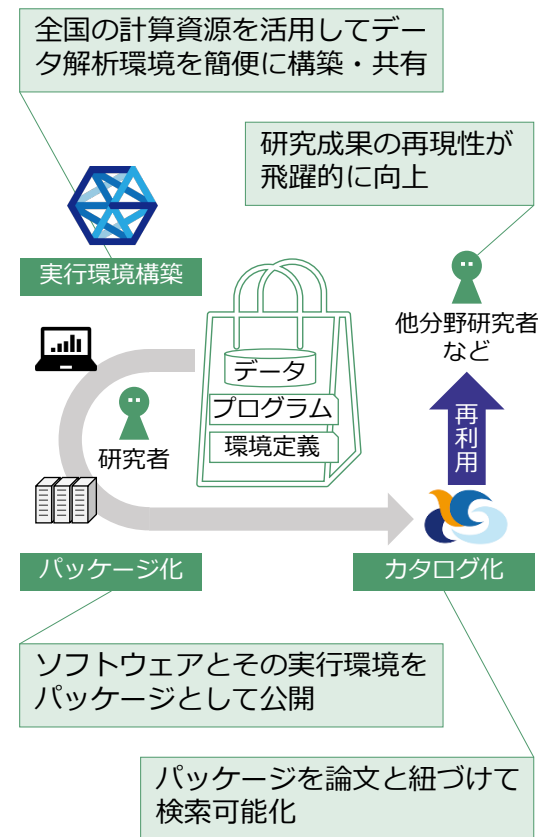
機密性の話

NII RDC 7つの新機能のうち3つ

【活用】

コード付帯機能の開発

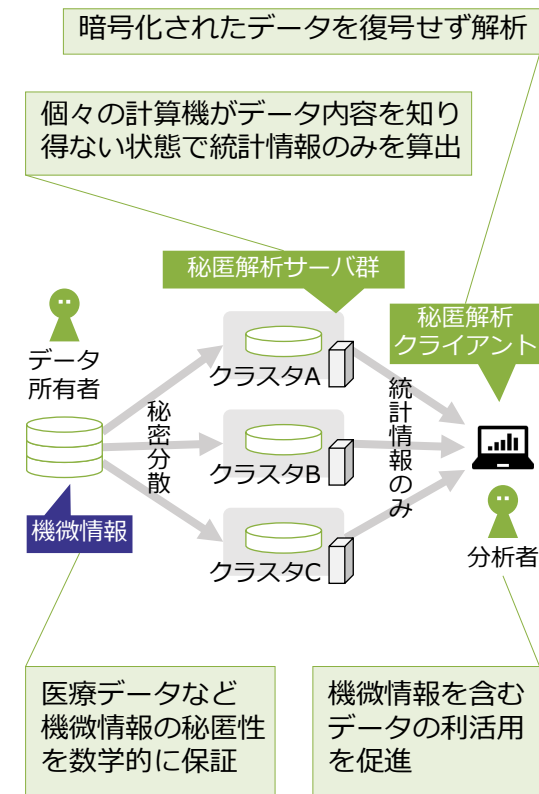
データ・プログラム・解析環境の
パッケージ化と流通機能を提供し、
研究成果の再現性を飛躍的に向上



【保護】

秘匿解析機能の開発

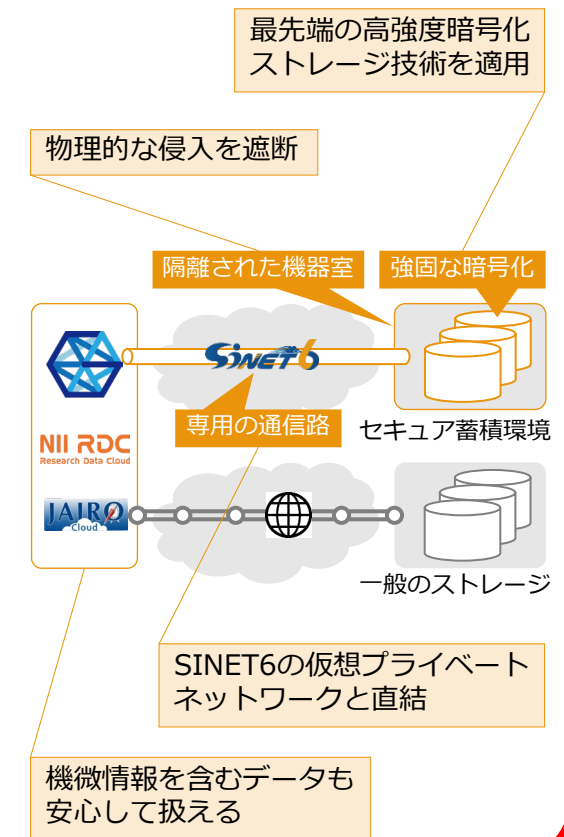
秘密計算技術で機微な情報も安心して
解析できる環境の提供で、新しい
データ駆動型研究の世界を開拓



【蓄積】

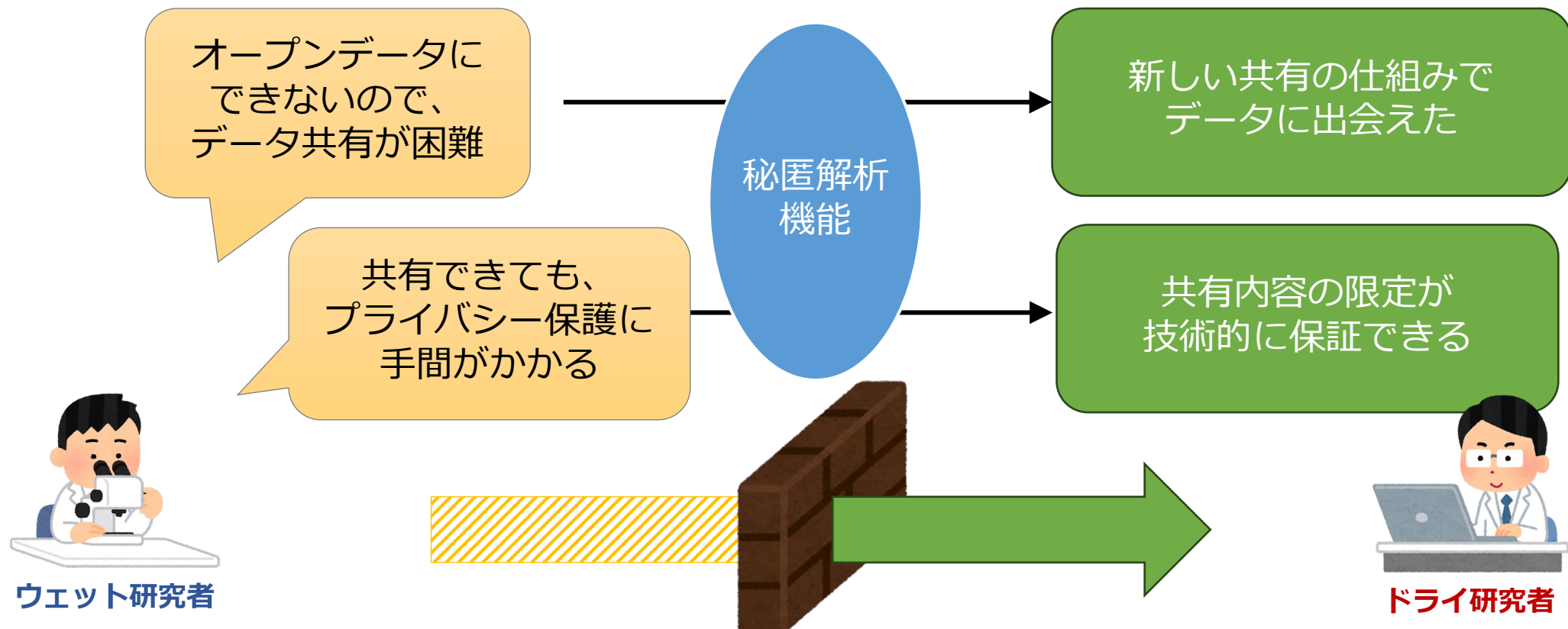
セキュア蓄積環境の開発

専用HWと高度な暗号化技術による
超鉄壁ストレージを提供し、データの
共有と保護の両立を実現



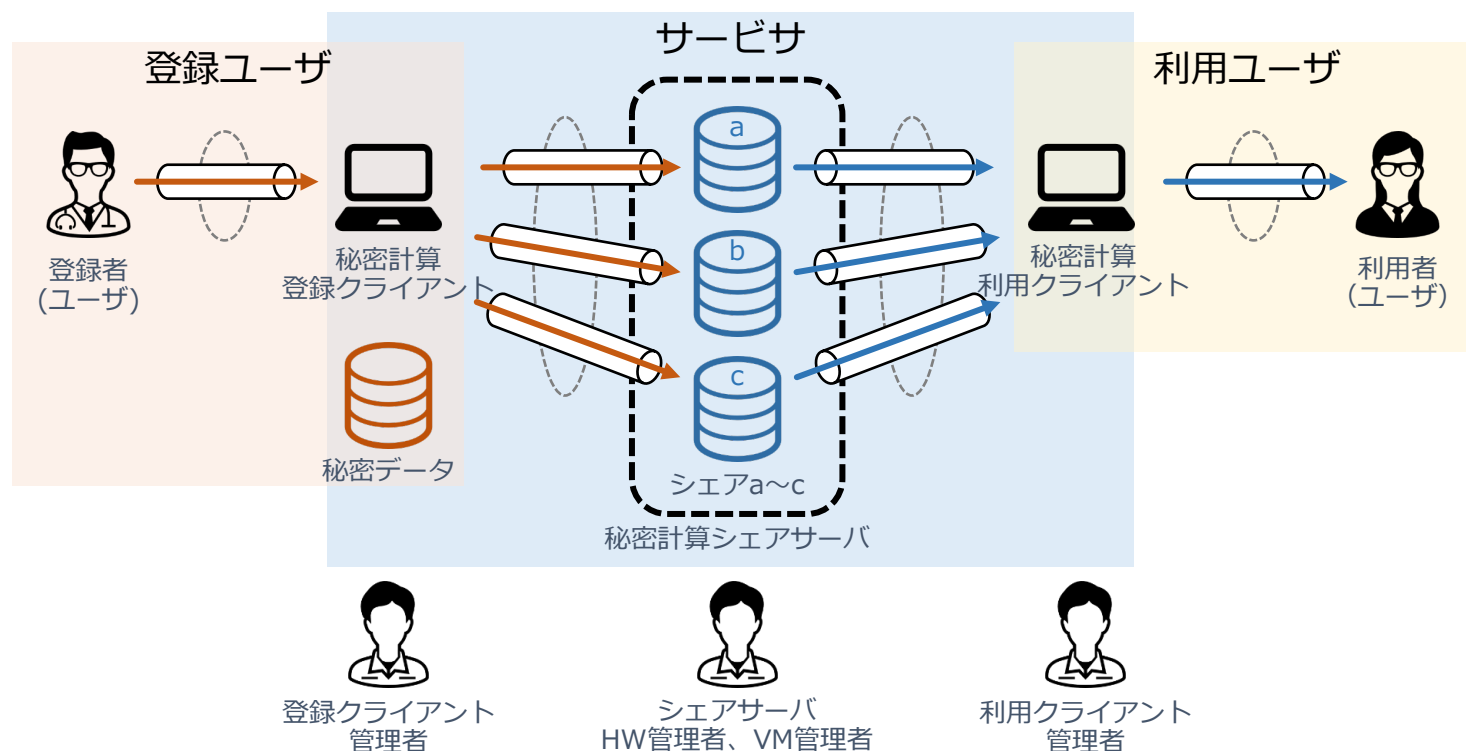
NII RDC 秘匿解析機能

- 一般公開できない機微な研究データの利活用を促進するために…
- 秘密分散にもとづくマルチパーティ計算（MPC）技術により、利用者に元データを開示することなく、統計分析のみを可能とするサービス



秘匿解析機能の仕組みと使い方

- メタデータは公開＝検索可能とする
- データ本体は非公開だが、秘匿解析（試用）は誰でも可能とする
 - 有用なデータに出会えたら所有者に連絡し、共同研究へ
 - オープンにできなかったデータも利活用の機会が広がる

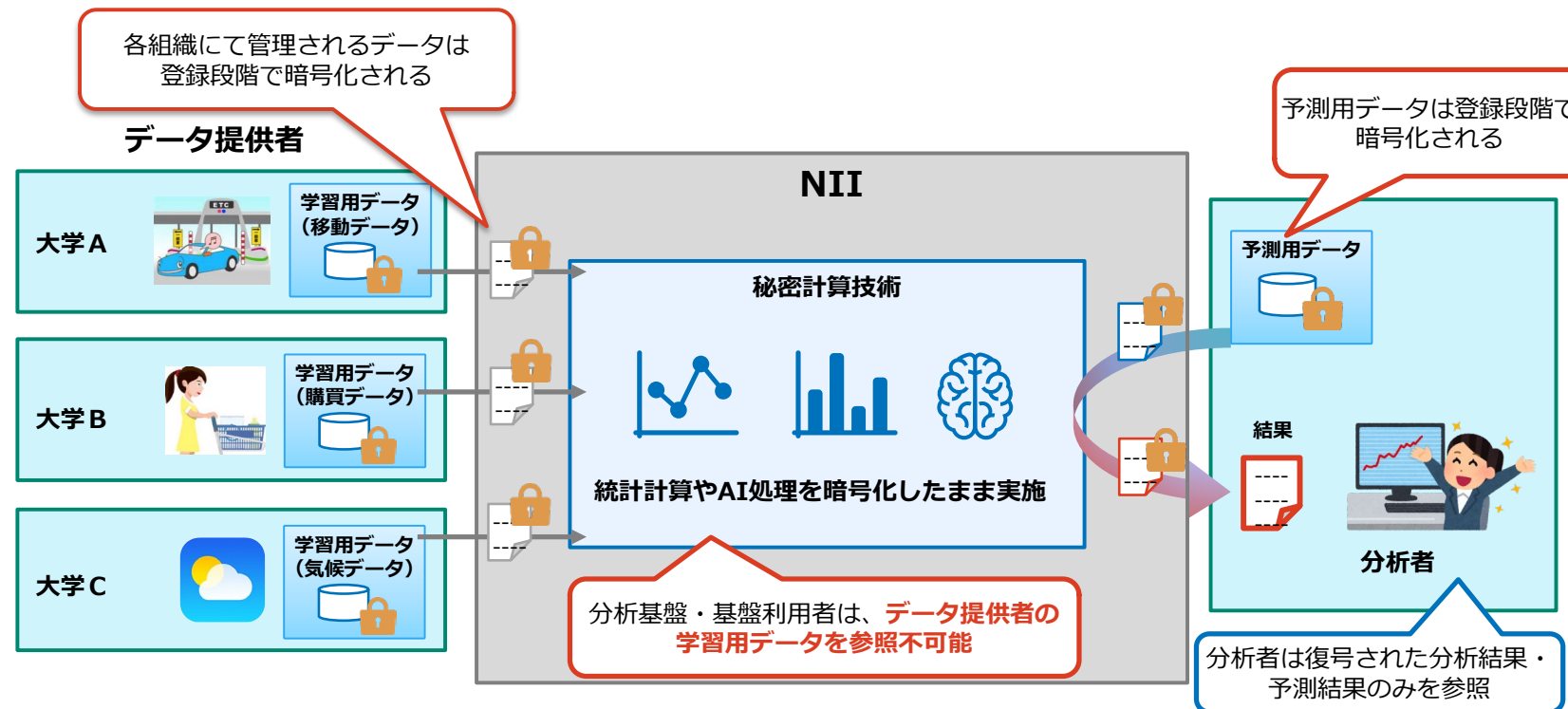


想定シナリオ

- データを安全に他組織へ試用させる
 - データ保有者がデータを秘匿解析環境に登録し（メタデータは公開する）、研究データを探している研究者にデータの中身は見せずに統計分析等の試用をさせる
- 複数の組織が持つデータを安全に統合分析する
 - 各組織のデータを安全に持ち寄り、データの中身を互いに見せないまま全体統計を計算する

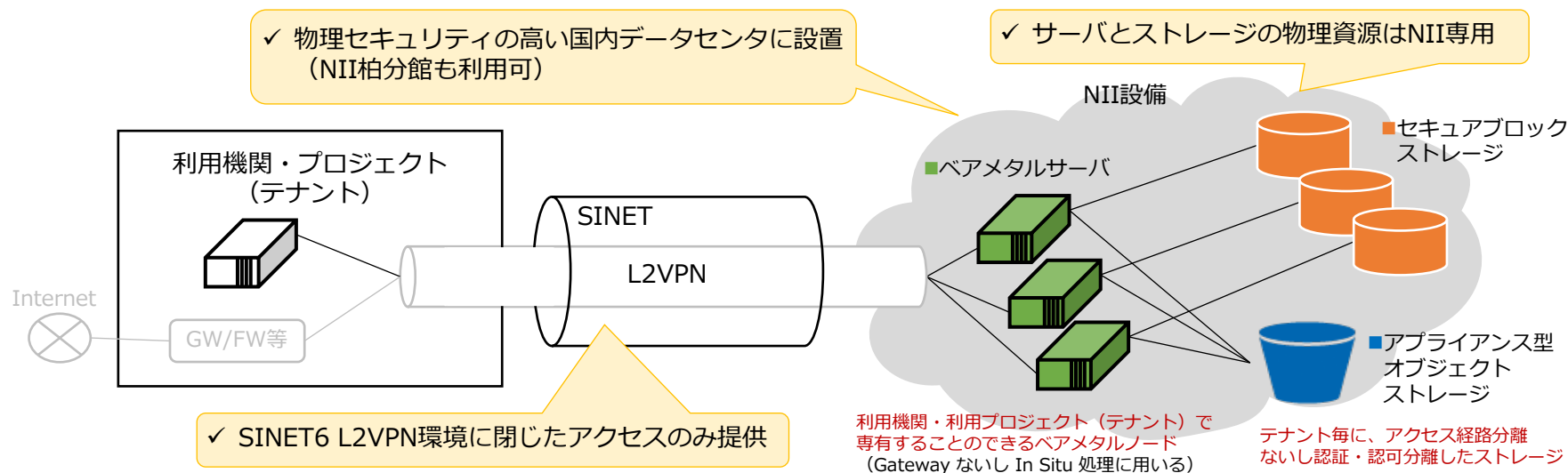
秘密計算システムの大学向けトライアル

- NIIの計算機上にNTTの秘密計算システムを構築し、大学等の研究者に無償提供
- 研究者は、システムに手持ちのデータを登録し、秘密計算ライブラリを用いたデータ分析プログラムをJupyterで書いて実行
- 秘密分散に基づくマルチパーティ計算の実用性を評価していただく
→ 有用な秘匿解析機能の開発へ



NII RDC セキュア蓄積環境

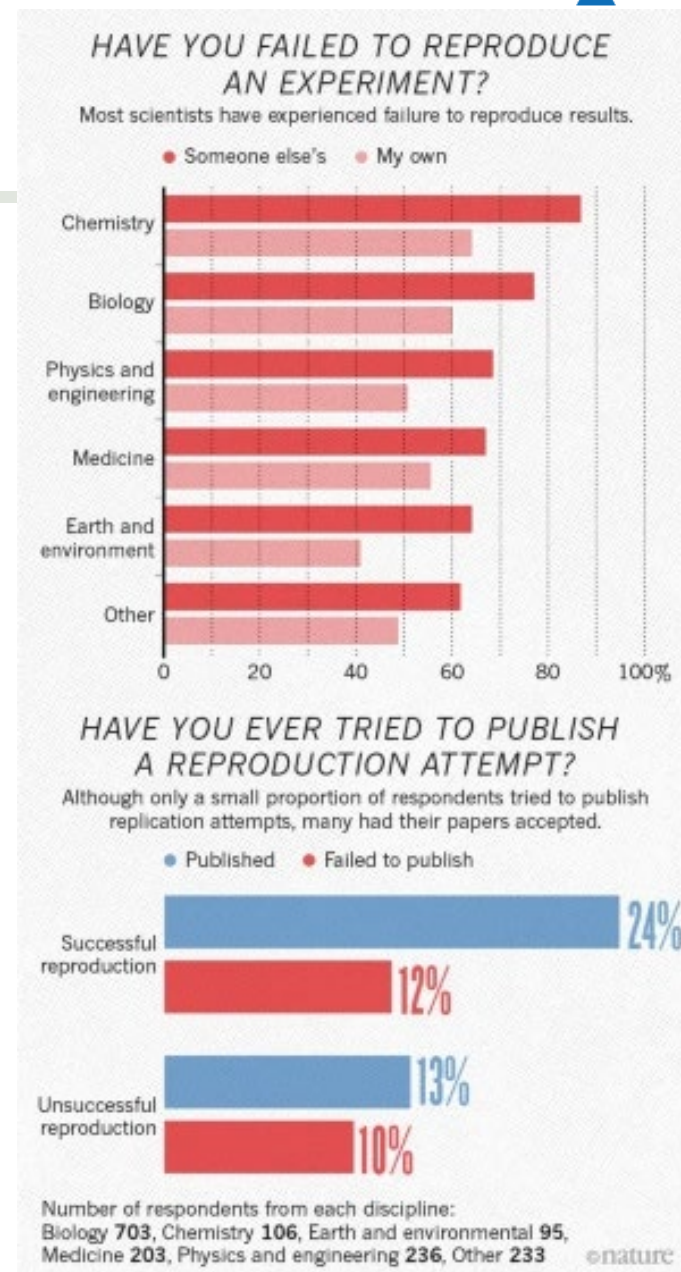
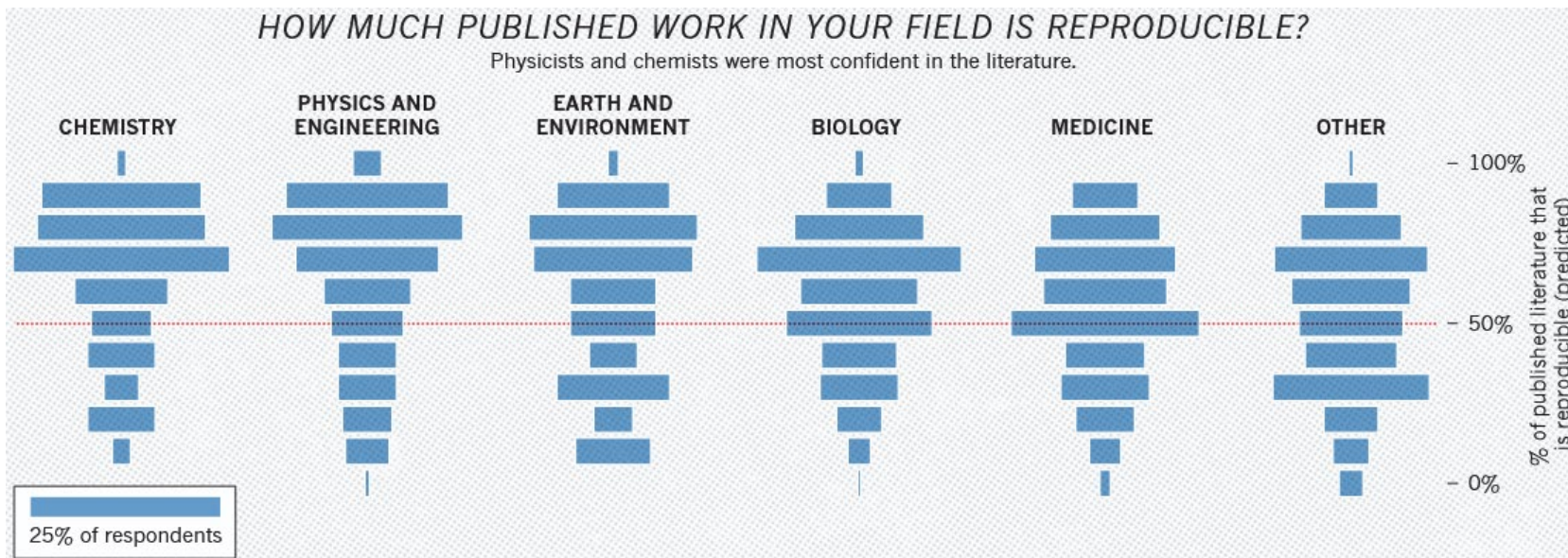
- 分離されたネットワークやデータベース、専用ハードウェアと高度な暗号化技術によるセキュアなストレージを提供
 - SINET6の仮想プライベートネットワークと直結。装置はNIIデータセンターに所在するが、ネットワーク的には利用機関内に存在
- 機微情報を扱う研究プロジェクトでデータの共有と保護を両立
 - 秘匿解析機能で、生データを扱う登録クライアントの動作環境としても利用
- 実証実験参加機関募集中



再現性の話

研究再現性の危機

- 『1,500人の科学者が再現性の蓋を開ける』
- 研究者の70%が、他の研究者の実験結果を再現できなかった経験がある (*Nature*, 2016) ^[1]

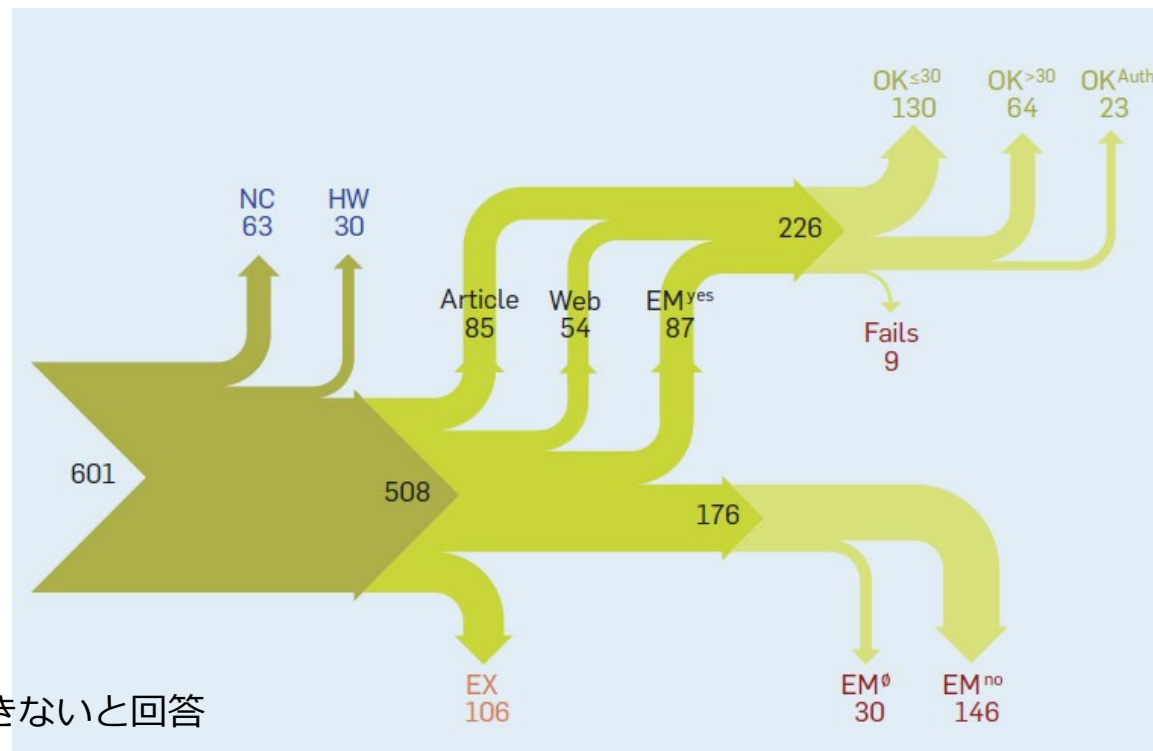


[1] Monya Baker, "1,500 scientists lift the lid on reproducibility", *Nature* vol.533, pp.452-454 (2016) <https://doi.org/10.1038/533452a>

計算再現性も危機

- ACMの論文601件を対象に、そのソースコードの入手可能性とビルドの可否を検証 [2]
- 32.3%が30分以内でビルド成功、
48.3%が時間無制限でビルド成功、
54.0%が著者ならビルド可能

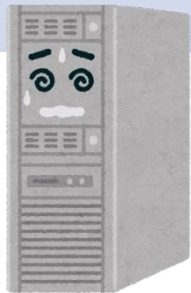
HW	特殊なハードウェアが必要なため除外
NC	コードによる裏付けがない結果のため除外
EX	著者リストの重複のため除外
BC	結果がコードによって裏付けられている
Article	論文自体でコードが見つかった
Web	Web検索によってコードが見つかった
EMyes	メール要求後、著者がコードを提供
EMno	メール要求に対して、著者がコードを提供できないと回答
EM $\emptyset$	メール要求から2か月以内に著者から返信なし
OK $\leq$ 30	コードが利用可能で、システムの構築に30分以内で成功
OK $>$ 30	コードが利用可能で、システムの構築に30分以上で成功
OKAuth	コードは利用可能だが構築に失敗。著者が妥当な労力でビルド可能と回答
Fails	コードは利用可能だが構築に失敗。著者がコードが壊れている可能性を認めた



計算がなぜ再現不能なのか

技術的問題

- 同一の実行環境を再構築できない
- ハードウェアのアーキテクチャやソフトウェアのバージョンが異なる
- 数値計算の安定性、並列演算の順序依存性



人的／社会的問題

- GUIによるファイル管理
- スプレッドシートを用いた手作業による前処理／後処理
- 手打ちのコマンドラインに含まれるパラメータ



永続性の問題

- データが入手不能になる
- ソフトウェアが入手不能になる
- 研究者が異動、死亡

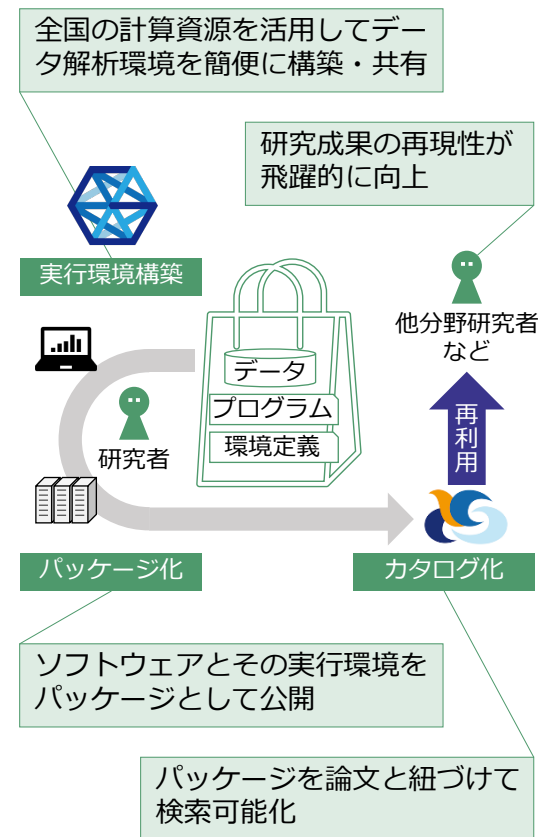


NII RDC 7つの新機能のうち3つ

【活用】

コード付帯機能の開発

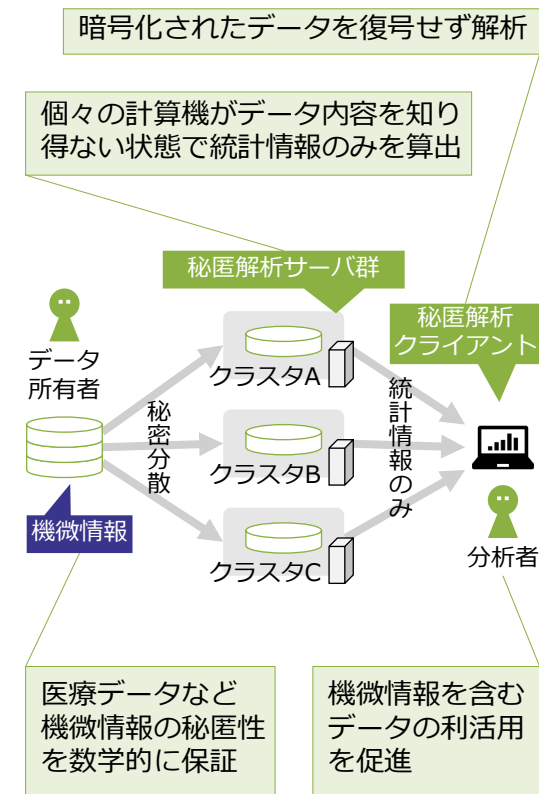
データ・プログラム・解析環境の
パッケージ化と流通機能を提供し、
研究成果の再現性を飛躍的に向上



【保護】

秘匿解析機能の開発

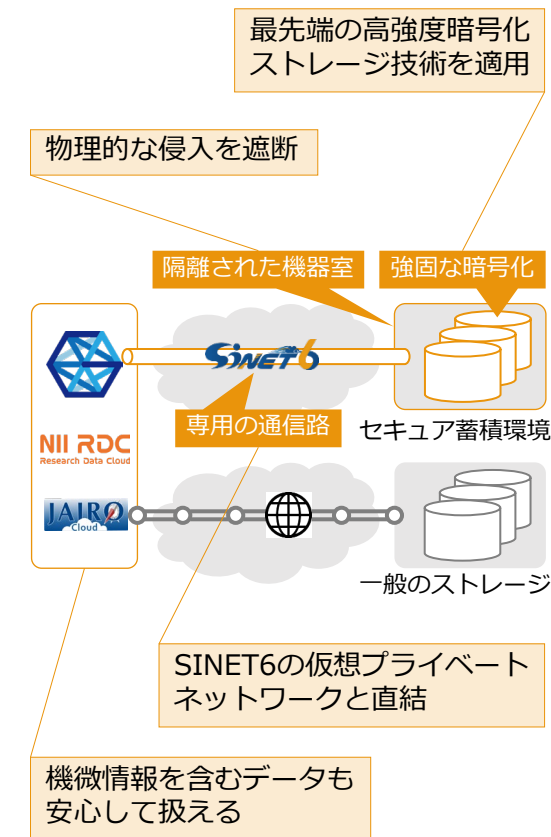
秘密計算技術で機微な情報も安心して
解析できる環境の提供で、新しい
データ駆動型研究の世界を開拓



【蓄積】

セキュア蓄積環境の開発

専用HWと高度な暗号化技術による
超鉄壁ストレージを提供し、データの
共有と保護の両立を実現



GakuNin RDM データ解析機能

- ① 個人用のデータ解析環境をNIIクラウドに※¹自動構築。
GakuNin RDM 内のデータをすぐに解析へ
- ② 多様なデータ源や計算資源と連携※²
- ③ 解析結果をプロジェクトメンバーと共有
- ④ プロジェクトを「計算再現パッケージ」として公開
- ⑤ 他の研究者が発見・再利用。発展的な研究へ

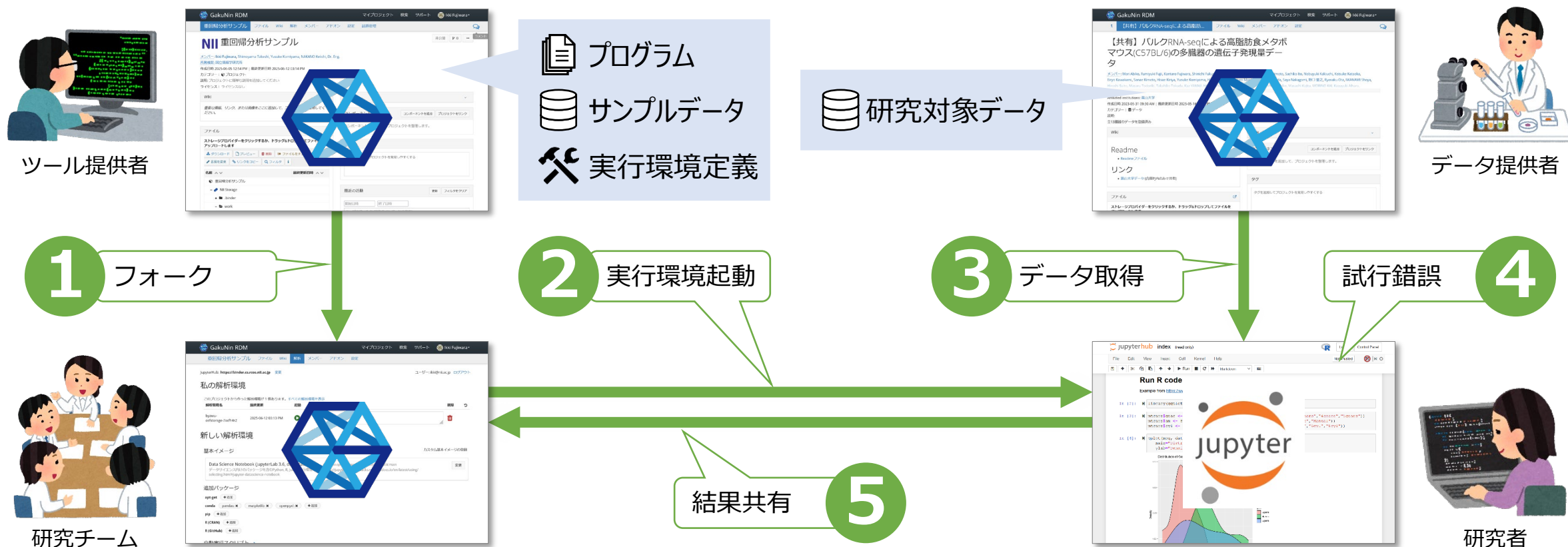
再現可能な
データ駆動科学の
普及を目指す

※¹ 機関所有計算機や商用クラウド等も利用可
※² 開発中の機能も含む



データ解析機能 > 利用イメージ

- ツール提供者のプロジェクトをフォークし、チームメンバーを招待
- 研究者ごとに実行環境を起動してデータ解析を実施。結果をチームで共有



データ解析機能 > 最新情報

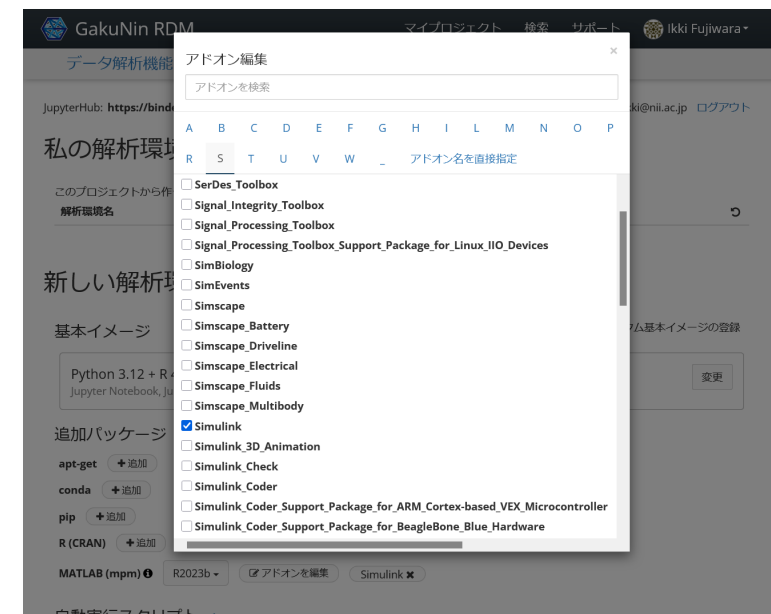
- ユーザーからのフィードバックに基づき、UI/UXのアップデートを継続的に実施



New! カスタム基本イメージ登録機能



New! 用途メモ機能



New! MATLAB アドオン選択画面

計算再現パッケージ機能

- GakuNin RDMのプロジェクトをパッケージ化して公開基盤へエクスポート
 - データ、プログラム、解析環境定義を含む。メタデータはRO-Crate形式
- 後続研究者が自分のプロジェクトとしてインポート
 - 先行研究者のデータ解析過程を再現し、発展的な研究をすぐに開始



実行証跡 (Governed Run) 機能

- GakuNin RDMプロジェクト中の解析プログラム (Jupyter Notebook) の実行記録を**実行証跡パッケージ**として保存
 - パッケージは、解析プログラムを再実行可能とするために必要な情報（実行環境定義、解析プログラム、入力データ、出力データなど）を含む
- 他のユーザーは、実行証跡パッケージから元の実行環境を再構築し、プログラムを再実行し、**解析過程を再現・検証**できる
 - ユーザーが手動で再実行するほか、データガバナンス機能からの要求に応じて自動で再実行することも可能

機密性と再現性は ガバナンスに含まれますか？

紹介したサービスのまとめ

事業名	サービス名	ステータス	参照
NII RDC コード付帯機能	GakuNin RDM データ解析機能	提供中	https://support.rdm.nii.ac.jp/usermanual/DataAnalysis-01/
	JDCat 分析ツール	提供中	https://jdcats.jp/analysis.html
	計算再現パッケージ機能	開発済み／提供時期調整中	https://rcos.nii.ac.jp/service/cs/
	実行証跡（Governed Run）機能	開発済み／提供形態調整中	
NII RDC 秘匿解析機能	秘匿解析機能	実証実験中／参加歓迎★	https://rcos.nii.ac.jp/service/sc/
	秘密計算システムの大学向けトライアル	終了	https://rcos.nii.ac.jp/news/2023/01/20230123-0/
NII RDC セキュア蓄積環境	セキュア蓄積環境	実証実験中／参加歓迎★	https://rcos.nii.ac.jp/service/ss/

★問い合わせ先 cs-support@nii.ac.jp



Center for Cloud Research and Development
National Institute of Informatics

cs-support@nii.ac.jp