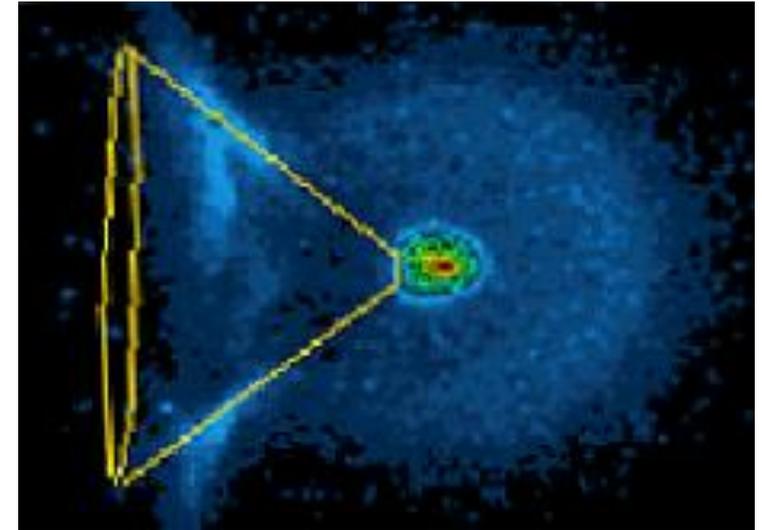
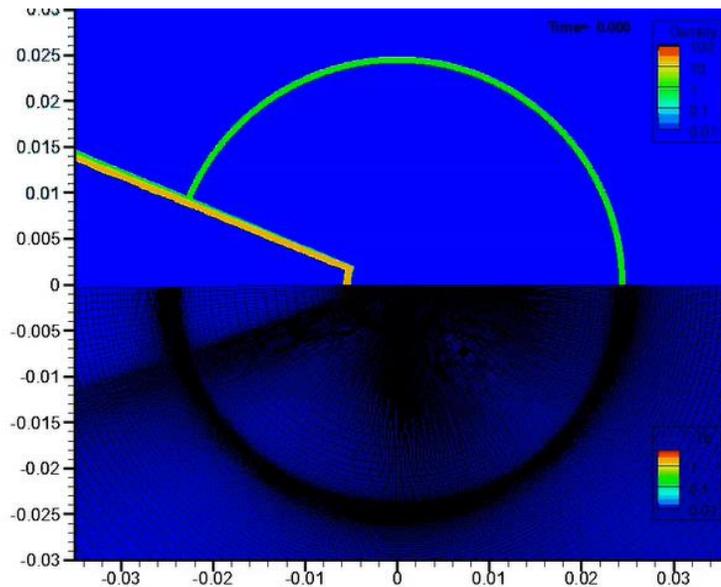


大阪大学レーザー科学研究所における共同利用・共同研究でのデータ活用と管理

～実験、シミュレーションのデータの信頼性、安定性、利便性、価値を高めるために～



大阪大学レーザー科学研究所 長友英夫
2022年3月14日 Cyber HPC Symposium 2022

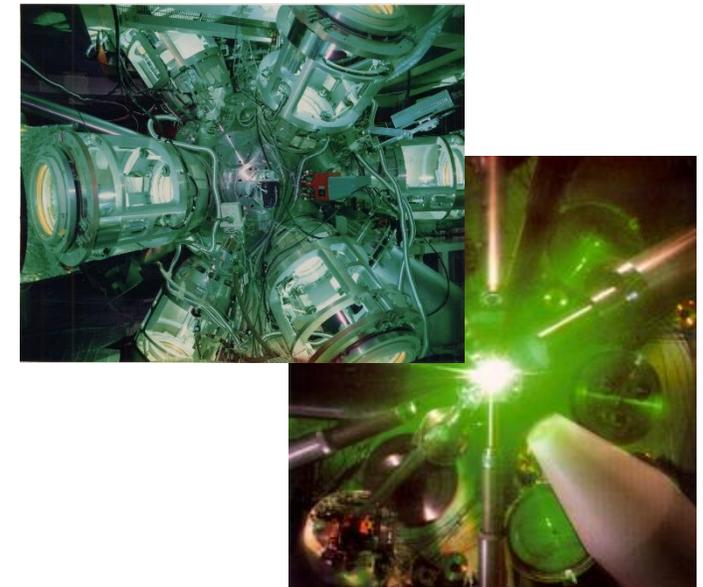
本発表は、大阪大学レーザー科学研究所全国共同利用・共同研究データベースワーキンググループ、先端研究基盤共用促進事業(先端研究設備プラットフォームプログラム)「パワーレーザーDXプラットフォーム」、大阪大学データリテリフロンティア機構 学際共創プロジェクト における研究・開発成果を含みます。ONIONの利用においてはCMCの担当者から様々なアドバイスを頂きました。関係者の皆様に御礼申し上げます。

概要

- レーザー科学研究所(旧レーザー核融合研究センター、旧レーザーエネルギー学研究中心)における実験データベースの沿革
- 先端研究基盤共用促進事業(先端研究設備プラットフォームプログラム)「パワーレーザーDXプラットフォーム」開始(2021年)
- レーザー科学研究所におけるシミュレーション研究とデータ活用
- レーザー科学研究所におけるデータ利活用の将来

背景(阪大レーザー研のデータベース環境の動向)

- レーザー科学研究所では大型レーザー装置があり、レーザー核融合、高エネルギー密度科学、レーザープロセッシングなど様々な研究に活用されている。
- 実験には多くの人に関わり、得られるデータを共有している。大型レーザーが稼働した1980年代に所内の研究者用データベースが構築された。Oracleベースの実験データベース、およびレーザーデータベースが構築され、データが蓄積・管理されてきた。
- シミュレーション研究では個別のシミュレーション研究だけでなく、それら実験に対応した実験解析に携わってきた。(レーザー研のスパコンは実験データ解析システムの一部として導入)
- 全国共同利用拠点として、全国の共同研究者が利用できる実験データベース、SEDNAの構築開始(2009年)
- 先端研究基盤共用促進事業(先端研究設備プラットフォームプログラム)「パワーレーザーDXプラットフォーム」開始(2021年)



実験データベースSEDNAを中心に、 共同利用拠点に相応しいデータ共有の仕組みを構築

実験データベースSEDNAの紹介(藤岡慎介:阪大レーザー研 所内説明会資料 2011)

SEDNA以前 (Oracleベースのシステム)

- データベースにアクセスできるのはレーザー研内からのみ
- USBメモリ等を使ったデータ受け渡し多数
- 電子メールの添付書類等での連絡多数
- 規格化されていない指示
- 大量の紙配布物

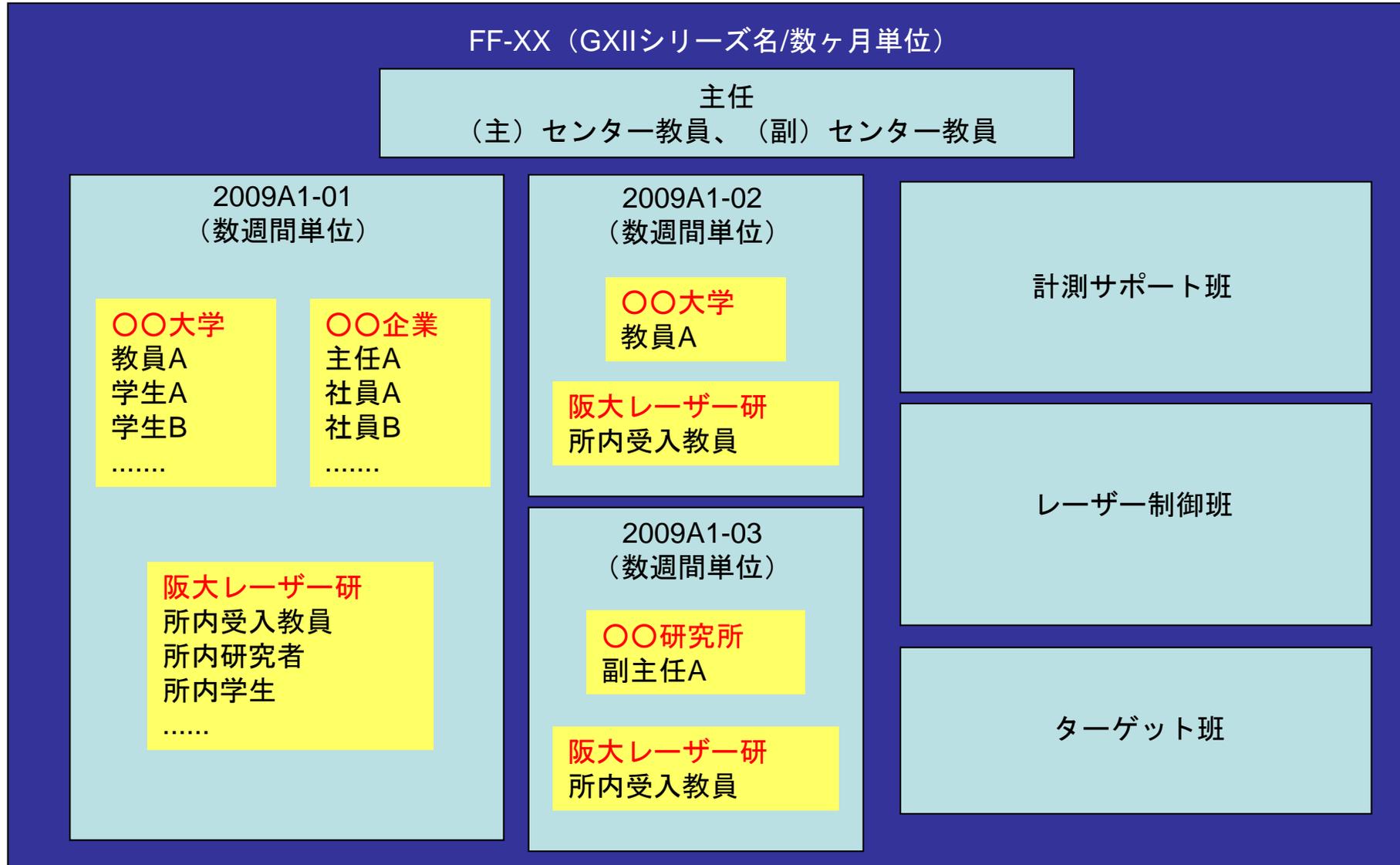
SEDNAが目指したもの

- 全国の共同研究者が所属研究室で実験提案・データ解析出来る環境
- 実験に関わるプロセスを規格化し、ミスや誤解を予防
- オンライン化により紙媒体での情報共有を低減
- 安心してデータ及び解析結果を保管できるセキュアな環境
- データ、データの流れ、解析の経緯を管理(漏洩・捏造の予防)
- 利用しつつ、状況に応じて改善・改良出来るシステム

SEDNAの夢

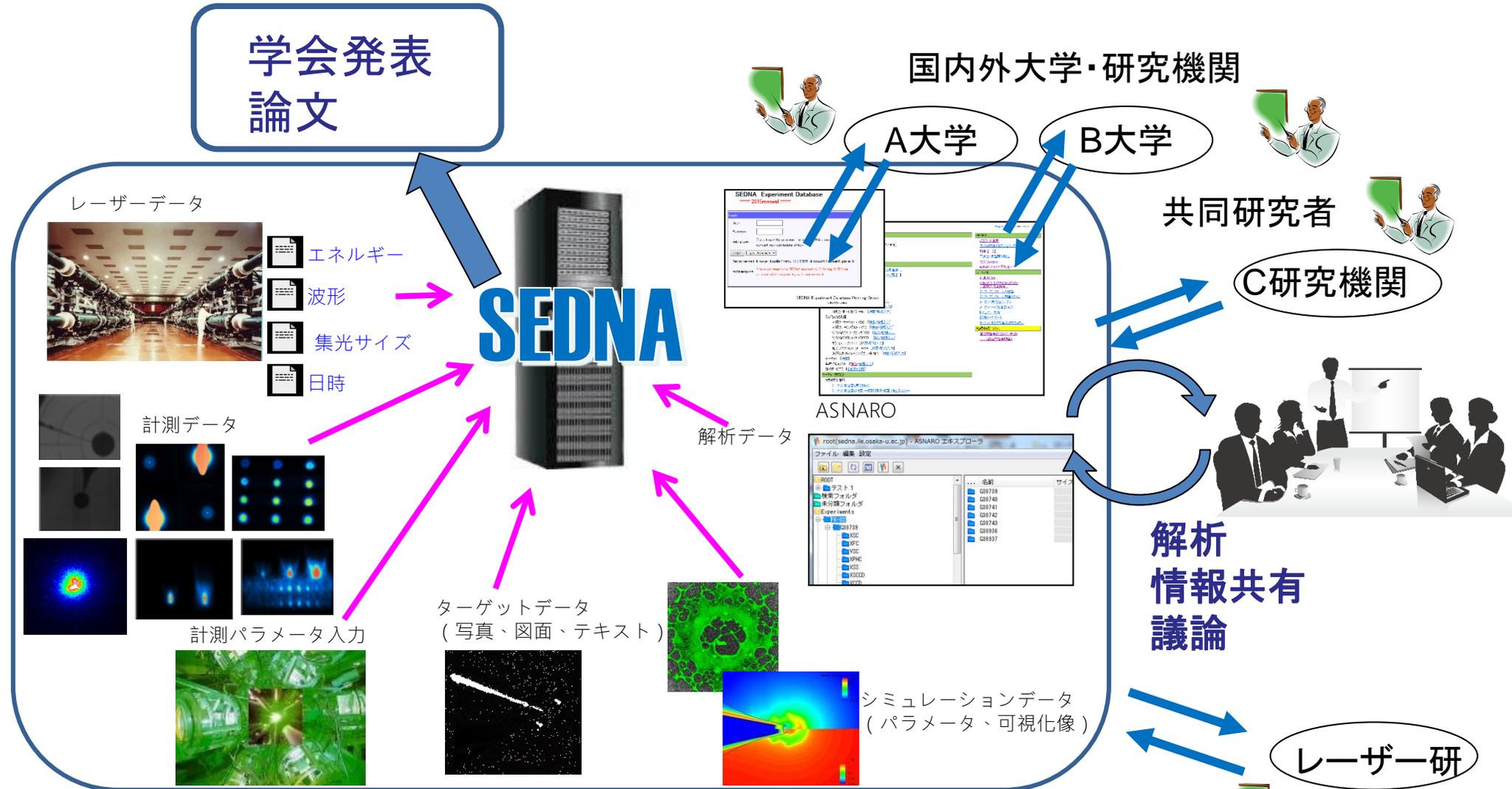
- 天体観測衛星のごとく、計測から数年を経たデータをアーカイブ化し、コミュニティで半永久的に共有できる環境の構築

SEDNAの経緯: 全共化後、外部研究者を含めた実験が増大 セキュリティーを担保したレーザー研外とのデータ共有が重要



Linux (UNIX)のファイルシステムにおけるツリー構造では適切な権限のコントロールができない。

実験データベースSEDNAをハブとして複数機関の複数の研究者が参加するバーチャルな研究環境を提供

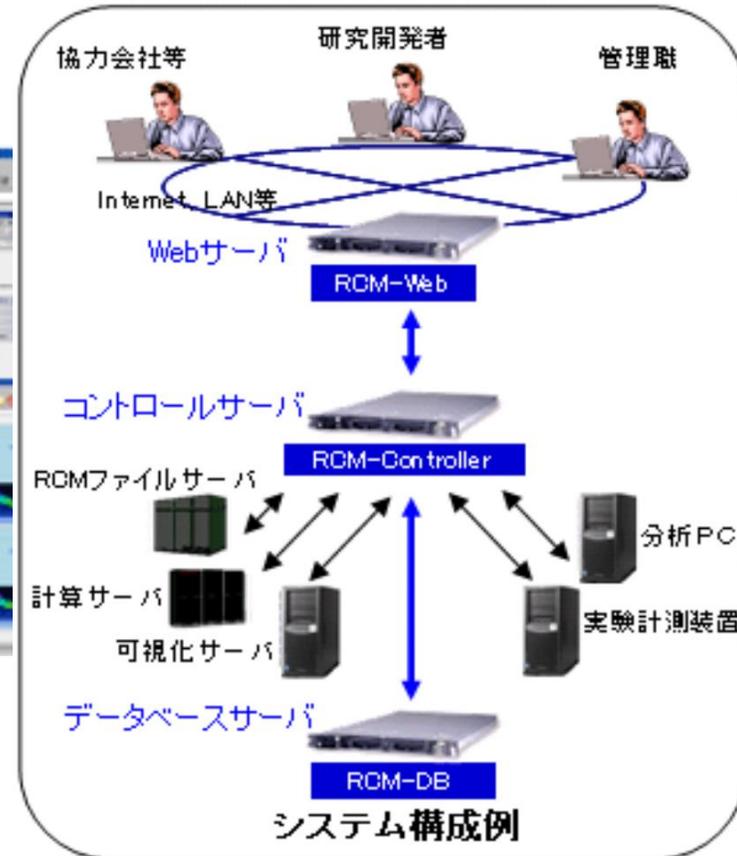
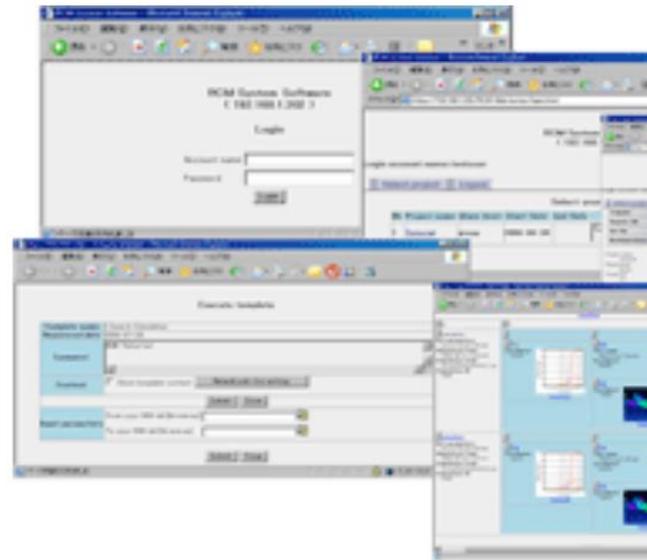


様々な共同研究に対応するため、データアクセス権限を細かく設定できる
データの不正改ざん防止のため、修正履歴をすべて記録できる

研究フロー統合管理システム(RCM)を活用した基幹データベース サーバーでデータを一元管理している

RCM(R&D Chain Management) System Softwareは、大規模データ処理に直面する研究者の方に、より高い次元の研究活動を行っていただくための研究開発基盤を構築可能な、多機能 Web サービスミドルウェアです。

RCM-Web, RCM-Controller, RCM-DB の 3 つのソフトウェアから構成されています。



- 基幹部分すべてに XML 技術を採用することで、データベースからユーザーインターフェイスまで柔軟に、且つ短期間に構築・再構築。
- 独自開発の XmlViewer により、複雑かつ膨大な情報を直感的に、よりフレキシブルに閲覧・編集。
- 研究フローの変更を前提としたワークフロー機能により、自動化とシステム構築に柔軟に対応。
- 高度なソート機能により、過去のデータを有効活用し、研究者のひらめきをサポート。
- 権限機能を活用し、研究環境をそのまま個人研究メモ、情報共有システム、情報公開システムという複数の役割を同時に実現。
- 過去の作業をトレーシングし、研究の再現性と信頼性を向上。

キャトルアイサイエンス社のWebサイトのRCM紹介
<https://www.i4s.co.jp/rcm/rcmabs.html>

SEDNAはキャトルアイサイエンスのRCMを導入することによって実現した

12C-ムレーン法30機 (30 deg. cone)			
1	11:45	11:53	T1
34213	#1 TARGET SHOT		1588 #1 TARGET SHOT
LFXE 発射ショット			H02, H04 => 500J, d=0
E01, E02, F04, E07, E08, F09 => 300J in 2μ, d/R=5 with RPP1			
F03, G05, G06 => 180J in 2μ, d=700 without xPP			
F10, G11, G12 => 300J in 2μ, d/R=5 without xPP			
16755 Core-CD shell (504.2 umf, 6.6 umf, tip-top 53 um, 33 nm Al coating) (D, 0, 0)			
最大注射の予想時刻は1.66 ns, LFXEの入射予想時刻は1.63 ns			

レーザー情報

ターゲット情報

発注書図面

Search No. Details Prev Next

ショットNo. 34213 ショット名 #1 TARGET SHOT 008 10/12/10 11:53

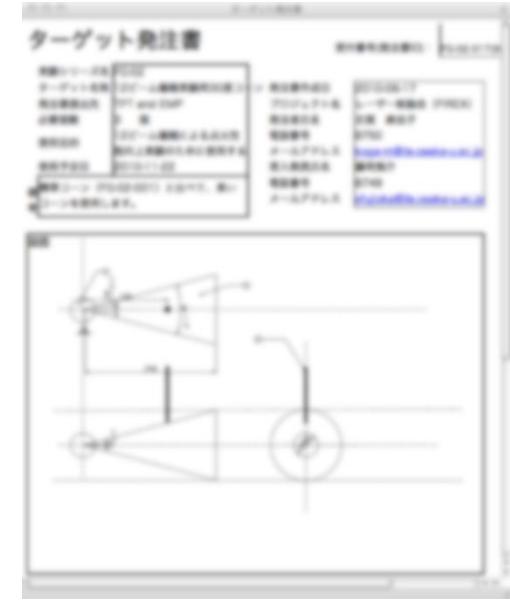
発射サイズ	ターゲット名	PHL	*	LFXE
FD-02	T1	PHL	*	LFXE

出力エネルギー一覧

	on RPP	Energy	Balance	Zurfo on RPP	Zurfo Back Scatter	Zurfo on Target	Energy	Balance
E01	767.33	766.24		343.83	38.43	315.38		
E02	715.68	766.53		337.22	31.99	318.24		
E07	768.89	765.81		329.76	6.53	363.38		
E08	639.29	632.98		319.63	12.56	294.86		
F03	586.31	585.24		383.81	23.53	348.54		
F04	953.28	943.69		316.96	0.24	293.68		
F09	792.18	784.18		345.79	7.38	318.81		
F10	782.51	694.49		391.21	31.61	319.91		
G05	457.15	452.58		223.58	47.59	269.62		
G06	512.31	527.81		289.29	27.38	292.51		
G11	798.34	795.31		327.28	21.83	363.82		
G12	787.62	779.71		362.47	27.58	278.27		

Search #1 TARGET SHOT

target_id	00705
target_id_end	00717
target_name	12C-ムレーン法30機30機
target_number	008
date	2011-02-04
submit_date	2011-02-04
remarks	
facility_id	
submit_group	
project_id	
objective	
assemble_remark	
memo	
registration_id	000001



SEDNAに求められている要件

- データの変更履歴の記録
- ユーザーインターフェース(UI)を自分たちで作成、改修可能
- データアクセス権限の柔軟な設定

パワーレーザーDXプラットフォーム

研究者の所属とバックグラウンドを問わない新共創プラットフォームへの デジタルトランスフォーメーション

代表機関 大阪大学レーザー科学研究所
 実施機関 東京大学物性研究所, 量子科学技術研究開発機構関西光科学研究所
 京都大学化学研究所, 理化学研究所放射光科学センター
 協力機関 北海道大学, 宇都宮大学, 光産創大, 広島大学, 九州大学, 宮崎大学, 産業技術総合研究所
 JAEA敦賀総合研究開発センター レーザー・革新技術研究所, レーザー技術総合研究所

背景

- 多様なパワーレーザー施設・設備で基礎学術から産業利用に広がる広範な学際研究が展開。
- 異なるプロジェクトで独自に開発・運用されてきた多様なパワーレーザーをプラットフォーム化し、パワーレーザーのリモート化とスマート化と利用のワンストップ化による、利用者視点での選択肢の広がりアクセシビリティの向上が、パワーレーザーに根ざした学術の飛躍と産業創成に必要。
- 新領域開拓やイノベーションの創出には、デジタル技術やオープンサイエンス思想の導入による、研究者の所属やバックグラウンドを問わない新たな研究基盤共用体制への発展が不可欠。
- 海外でパワーレーザー施設のネットワーク化(米国LaserNetUS, 欧州 LASERLAB-EUROPE)が進展。

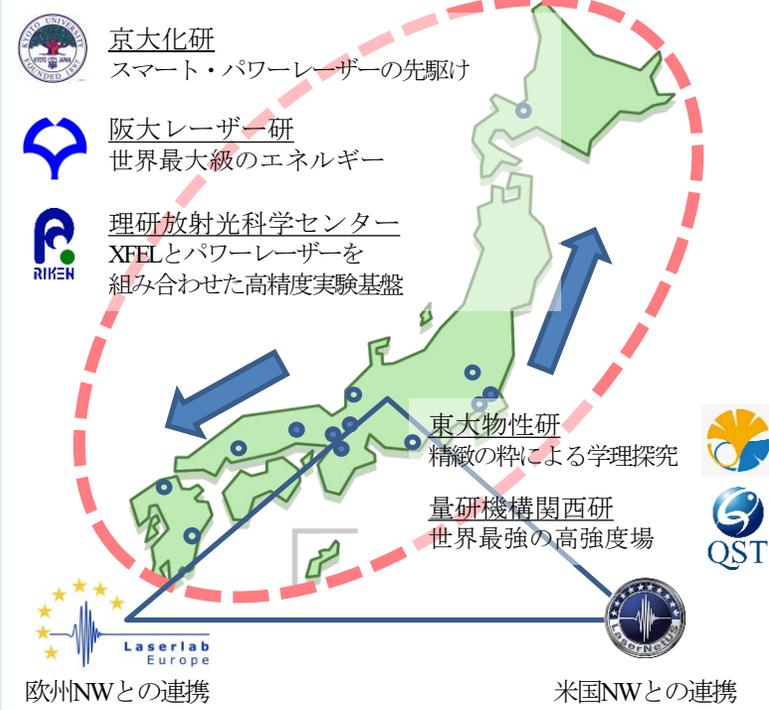
これまでの取り組みと課題解決への土台

- 融合光新創成ネットワーク事業による光技術の尖鋭化と高度化及び光科学研究者の育成とネットワーク化, JAEA-OU間の連携融合事業, QST-OUの包括的協力の締結, RIKEN-OUの光科学連携センターの設立, パワーレーザーに関するマスタープラン2020提案, 日本学術会議からの提言2020
 → **パワーレーザーのプラットフォームの基盤**
- パワーレーザー利用研究に関する先端国際拠点事業の実施, 同分野での日米間科学技術交流協定締結
 → **国内パワーレーザープラットフォーム形成に対する海外からの期待**
- 各種事業による代表・実施機関でのリモート化, スマート化の導入
 → **パワーレーザー施設・設備のデジタル化に向けた新たな潮流**
- 利用者として施設・設備を繋ぐセキュアなデータ共有システム(SEDNA)の長年の運用実績
 → **オープンサイエンスの土台**

目標達成に向けた戦略

- パワーレーザー“ソムリエ”を採用し、各機関での専門性を高めると同時に、プラットフォーム全体と各プラットフォームとの連携で、ソムリエの名に相応しい幅広い視野, 知識, 技術を持った研究者を育成。
- パワーレーザーソムリエを核としたプラットフォーム運営委員会の組織化。運営委員会による利用者へのワンストップサービスの提供。既存・新規利用者の施設・設備へのナビゲーション, 潜在的利用者の開拓。
- 国内プラットフォームと海外ネットワークとの互恵的連携の構築と, 国際的なワンストップサービスの実現。
- 代表機関, 実施機関でのスマート化・リモート化技術の導入と, NPO法人等を介したプラットフォーム内でのスマート化・リモート化技術の普及による, デジタルパワーレーザー施設・設備の構築。
- パワーレーザー施設・設備におけるオープンサイエンス・ポリシーの策定と, オープンサイエンスの実施に向けたガイドラインの策定。オープンサイエンスの利用拡大に向けたデジタル技術講習会の開催。
- 日本学術会議からの提言, 学術の大型研究に関するマスタープランへの提案等を踏まえ, 事業終了後も持続可能なプラットフォーム構築について, コミュニティ内で議論と合意形成。

パワーレーザー施設の連携 国内パワーレーザーのプラットフォーム構築と国際ネットワークへの拡大



① 研究の入口から出口までを支援
 パワーレーザーソムリエを核としてワンストップサービスを実現。パワーレーザーソムリエによる、新規利用者最適なパワーレーザー施設・設備へのナビゲーションと、既存利用者の研究発展に応じたパワーレーザー施設への段階的ナビゲーションを実施。産学フォーラム等を通じて研究成果と産業界を橋渡しを行うなど、研究の入口から出口まで包括的にサポート。

② 多様性を受け入れられる研究基盤共用体制の実現
 オープンサイエンスのポリシーとガイドラインを策定。オープンサイエンスに必要なデジタル技術をコミュニティに普及させることで、占有期間終了後の実験機器, 解析ツール, 実験データの共有を促進。新領域創成への閾値を下げると同時に、異動, ライフイベント等の労働環境の変化など、利用者の多様性を受け入れられる新しい研究基盤共用体制を構築。

③ デジタルパワーレーザーの整備
 装置運転のリモート化, ルーティン作業のスマート化, 装置不具合の自動検知等によって、パワーレーザーの運転における人の関与を減らし、利用者へ提供可能なサービス, 運転時間を増やす。

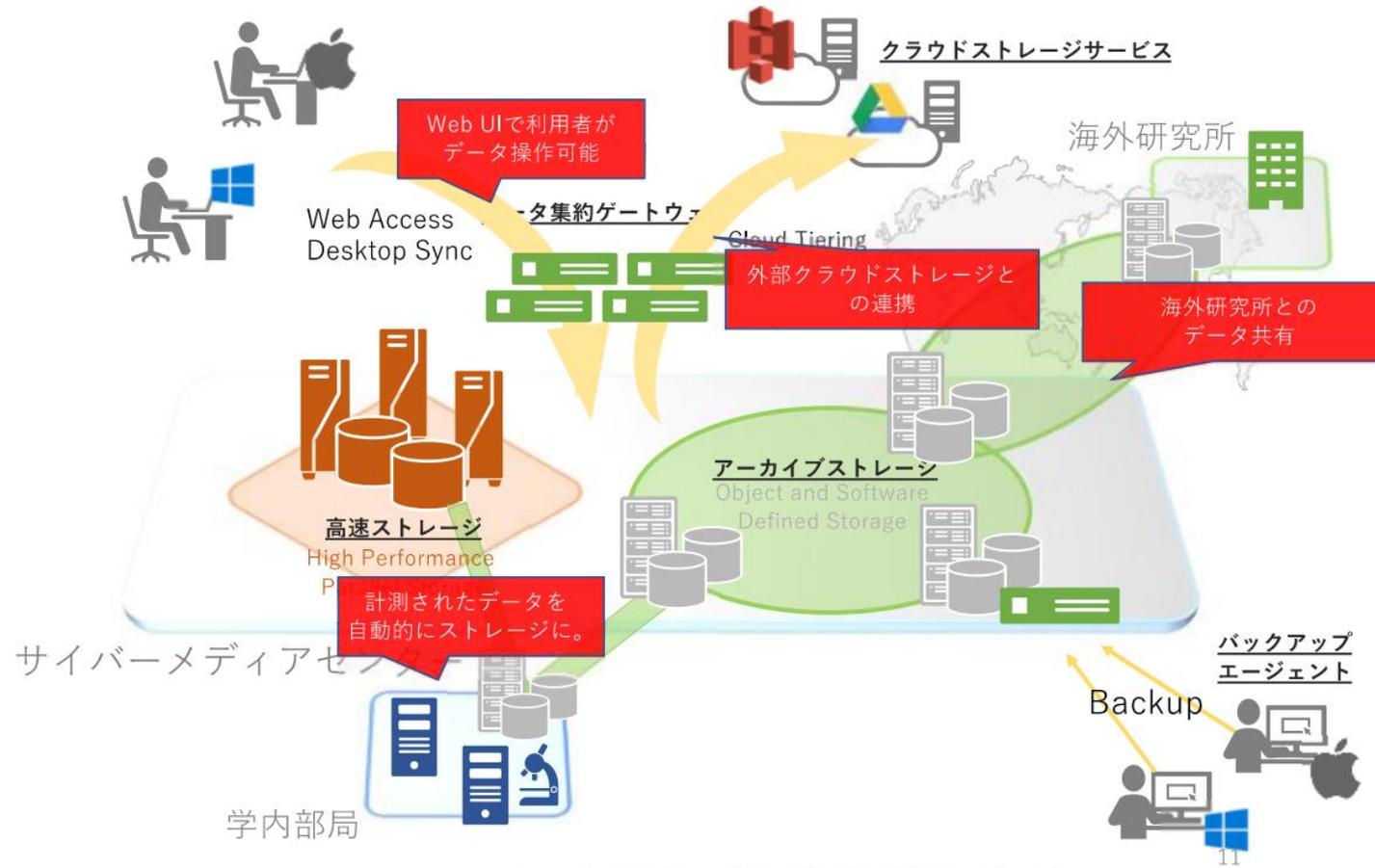
④ 時間と空間を超越した研究活動の実現
 オープンサイエンスでプログラム開発や解析に要する時間を減らし、実験から成果発表までの時間を短縮。オープンサイエンスポリシーに準拠した機器類の利用者による開発によって、研究活動のリモート化を促進。セキュアで国際的なデータ共有によって、時差を活用したシームレスなデータ解析を実現。

	R3	R4	R5	R6	R7	R8
ワンストップサービスの実現	運用体制構築	ソムリエ選出	代表機関・実施機関・フォーラムへのナビゲーション			ナビゲーション施設の拡大
国際ネットワーク化			互恵的国際連携の構築			国内・海外施設の相互利用の仲介
オープンサイエンス	ポリシー案策定	ポリシー合意形成	ガイドライン案策定	ガイドライン合意形成	ガイドラインに基づくオープンサイエンスの実施	
コミュニティへの技術波及			リモート化, スマート化, デジタルデータ処理技術に関する技術講習会の定期開催			
研究のリモート化		共同研究者間のデータ共有の強化				オープンデータ
デジタルパワーレーザー		自動アラインメント技術導入			自動ダメージ検知システム技術導入	
				レーザーのスマート化に伴う支援員の段階的転換とリモート研究の段階的範囲拡充		

大阪大学サイバーメディアセンターデータ集約基盤 ONION

阪大CMC SQUID利用説明資料より

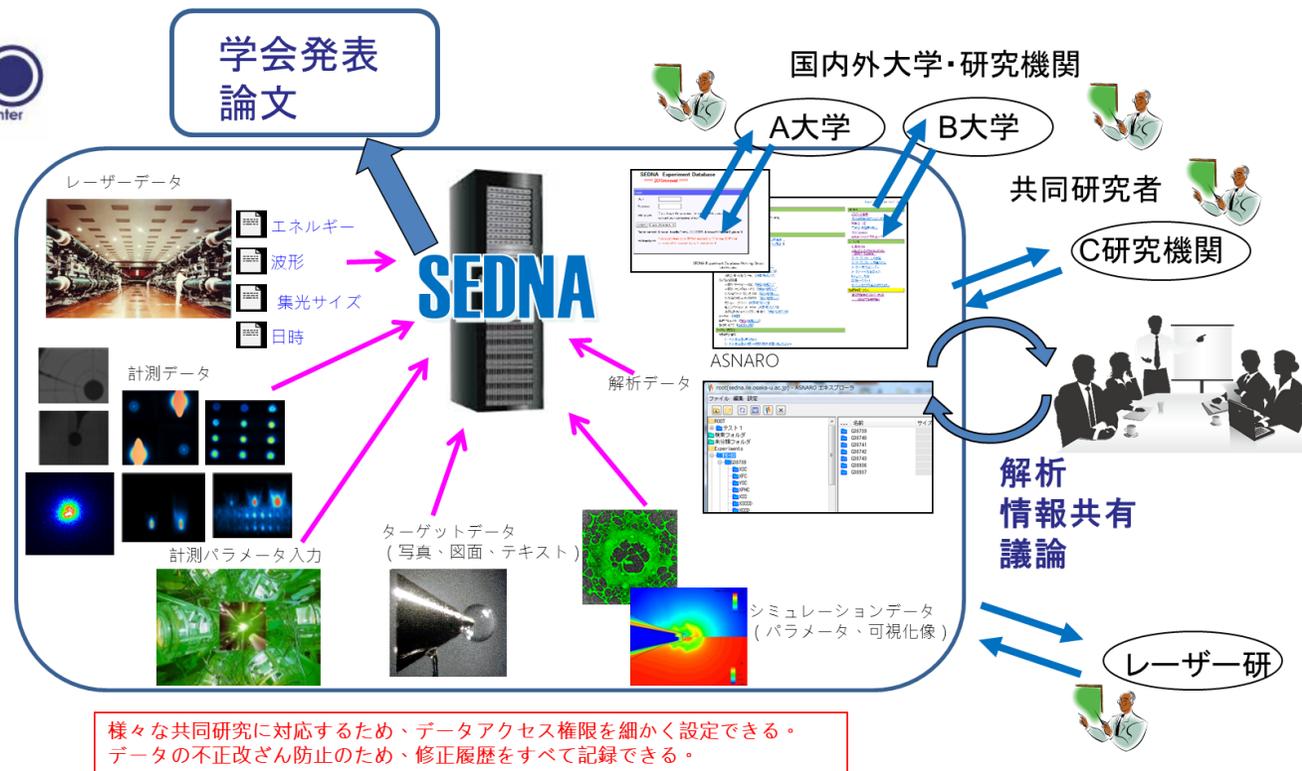
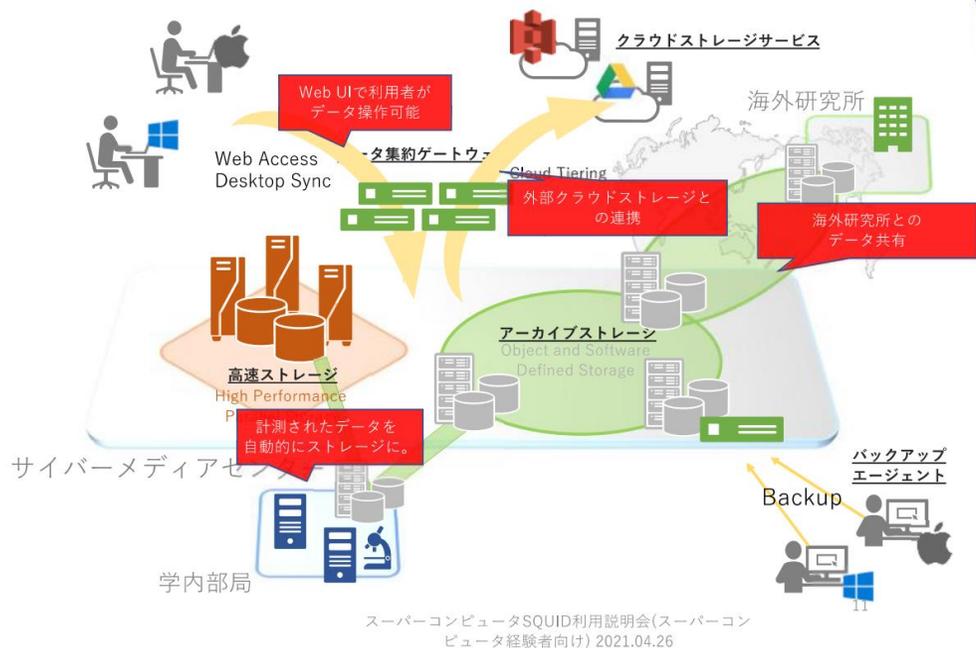
データ集約基盤 ONION 概要



スーパーコンピュータSQUID利用説明会(スーパーコンピュータ経験者向け) 2021.04.26

大阪大学サイバーメディアセンターデータ集約基盤 ONION

データ集約基盤 ONION 概要

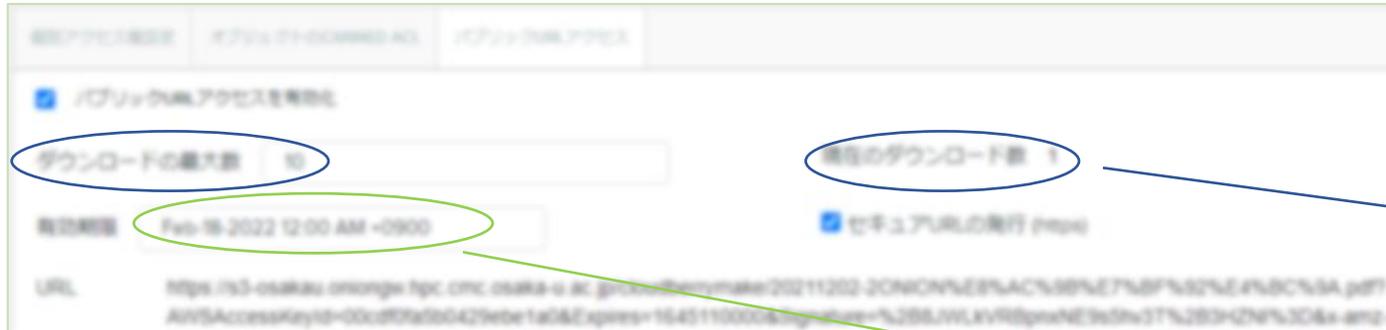


ONIONと目指してるところは同じ？

パワーレーザーDXの外部向けサービスにおいてONIONを活用できないか検討を行っている。

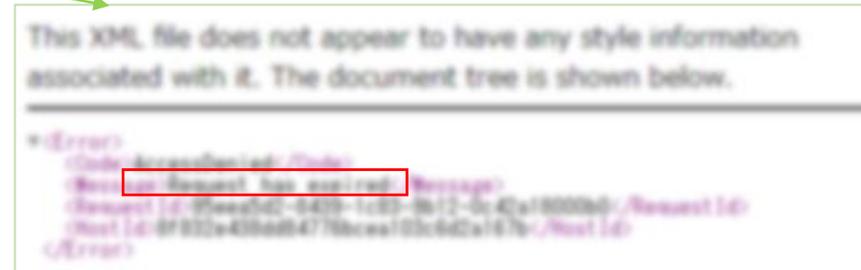
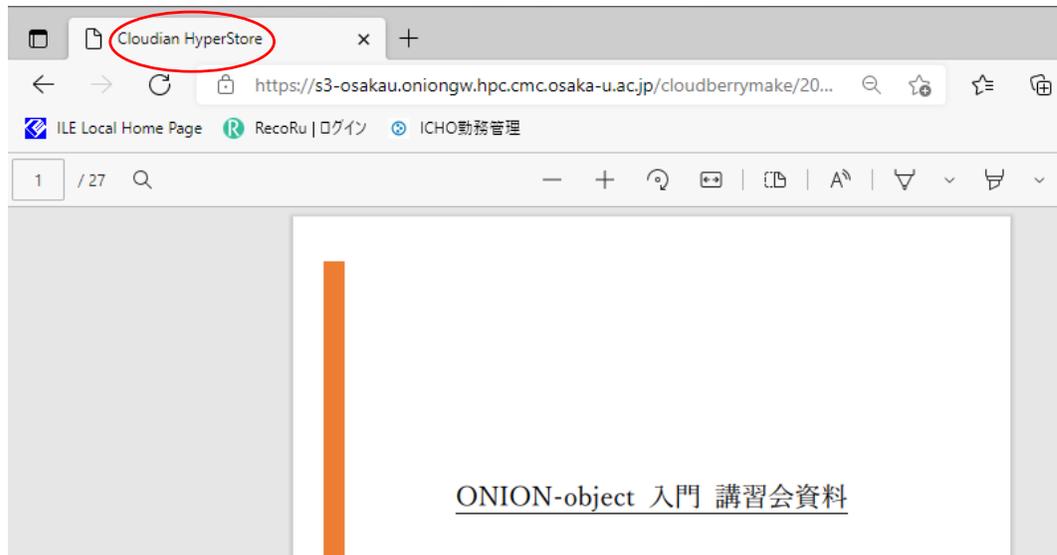
パワーレーザDXのデータ共有とONION-object

外部向けデータ公開用に利用を検討中である。



- ・ダウンロード数の設定が可能
ダウンロードされた回数もわかる

・アカウント持っていない人にもURLでファイル共有可能



- ・期限も設定でき、期限が過ぎると
Request has expiredと表示され
アクセスできない

パワーレーザーDXのデータ共有とONION-object

パワーレーザーDXの外部向けデータ公開用にONION-objectの利用を検討中である。

(検討の途中経過)

- ONION-objectにファイルを置いたり、共有することによって外部の共同研究者との実験データの授受に有効活用できる。
- ONION-objectでは、WEB管理画面からブラウザを通してバケット/オブジェクトの簡単な操作や設定が可能。
- ただし、ファイルへのアクセスログが十分ではないので、機密性の高いデータのやり取りには不向きである。
- 既存のSEDNAとの連携の可能性を検討する必要がある。
- アクセスキーを使って一般的なS3クライアントソフトを使って操作することも可能S3 API クライアント(その他フリーソフト)

レーザー研におけるシミュレーションのデータ利活用

- 大型計算機(スパコン)を用いたシミュレーション研究と実験データベースの活用は1980年代から運用実績があった。
- シミュレーション研究は個別研究でデータ共有の概念がほとんどなかった。
- データを共有する場合は、UNIXファイルの権限開放を利用。あるいは、ディスク、USBメモリ等の物理的な媒体でデータの受け渡しを行っていた。
(データ容量と通信環境の制約も大きかった)
- 最近でもデータを共通するユーザーが共通で利用しているサーバー内でのデータ受け渡しが常套化している。

高度なデータ共有システムが必要になっている

- 共同研究者とのデータ共有、オープン化
- マルチフィジックスシミュレーションの発展でシミュレーションの分業
- データ量の増大
- データ探索

極限状態を扱うレーザープラズマ物理では、シミュレーション結果と実験結果を矛盾なく結びつけるためには、双方で解析、データ校正、照合を繰り返す必要がある。

シミュレーションの問題点

モデル化可能な数値モデルを組み合わせ、離散化した時空間で数値解析を行う。

- ・数値モデルの不足、不確実性
- ・計算資源の限界

実験の問題点

高価な実験設備、計測器と熟練した計測技術でデータを取得している。

- ・時間、空間分解能の限界
- ・実験回数の制約
- ・実験条件の正確な把握

このようなことから、複雑現象の解析では、実験結果を矛盾なく説明するために条件、モデルを組み替えるなどしたシミュレーションを繰り返し、できる限り妥当な条件を定めてから結果を導き出す必要がある。

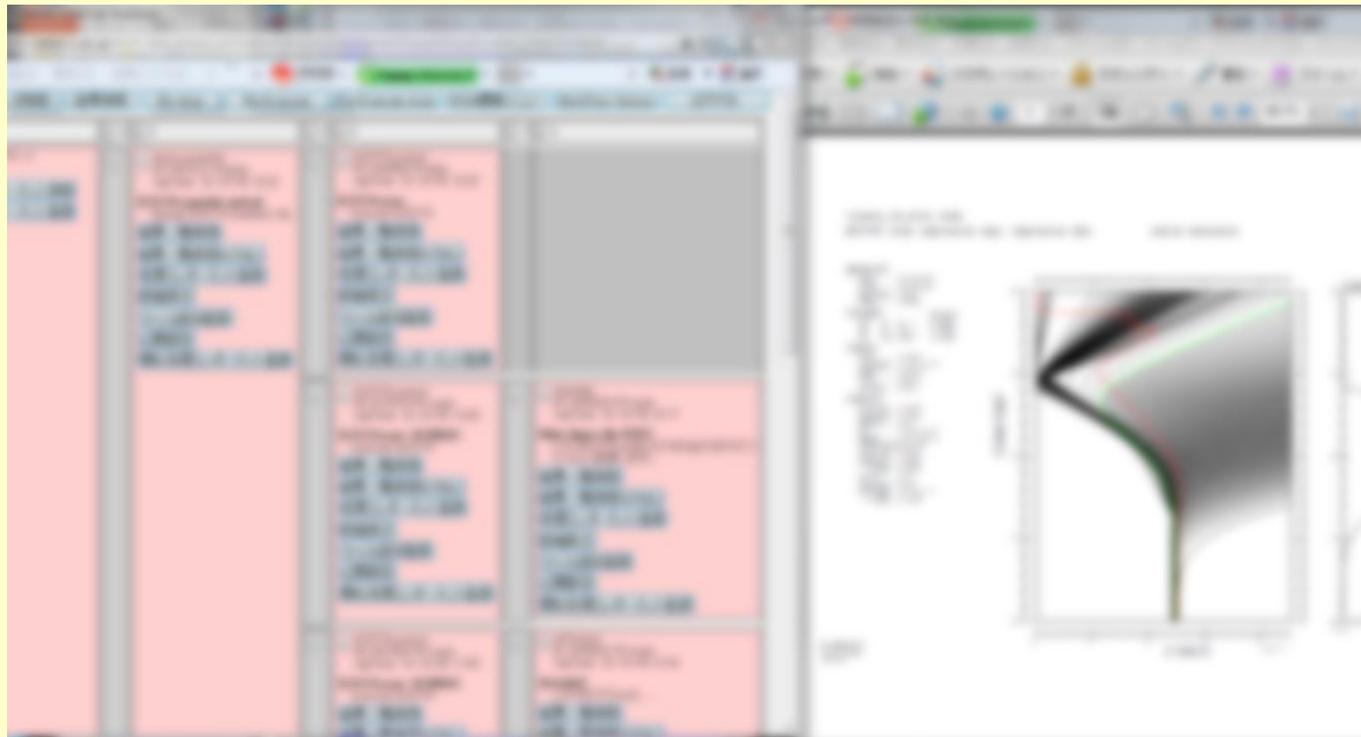
そのシミュレーションと実験の比較、連携、同定作業において、データベースは重要な役割を担う。
安定性、信頼性、利便性が求められる

過去にはSEDNAにシミュレーションコードを組み込む試みも行われた

実験データベースSEDNAに1次元のレーザープラズマシミュレーションコードを組み込んだ。
UI、あるいは直接namelistファイルをアップロードしてパラメータ入力を行い、ジョブ投入、ポストプロセス(データ可視化、PDFファイル作成)までを半自動化した。

SEDNAから利用できるようになります: ILESTA-1D

・シミュレーションの知識がなくても利用できるようになります。(現状は検索できませんが、本運用時には、実験データとセットで保存、検索可能になります。)



第2段階: STAR-1D、PINOCO-2Dの運用



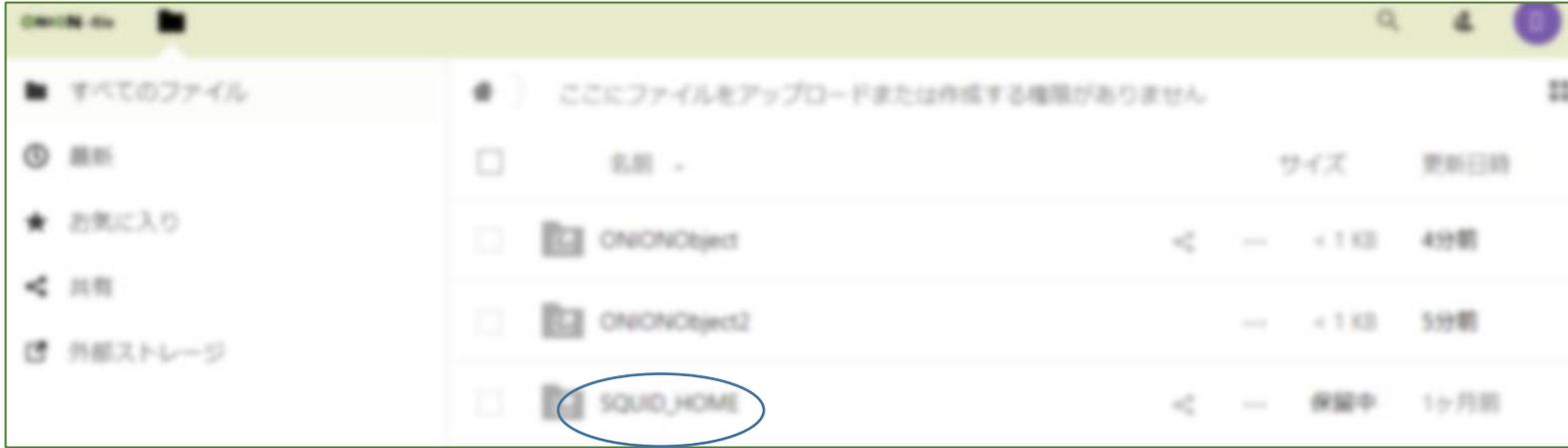
利用者数が伸び悩む

実験データベースと同じシステムでの
利用は普及しなかった。
個人ベースでのデータ共有のまま

問題点

- ・コード開発
- ・コード管理
- ・UIの拡充
- ・タグ付け
- ・データ量が膨大

- ・市販ソフトの活用 (ライトユーザー)
(UIなどユーザーフレンドリー)
- ・コンソールベース (ヘビーユーザー)

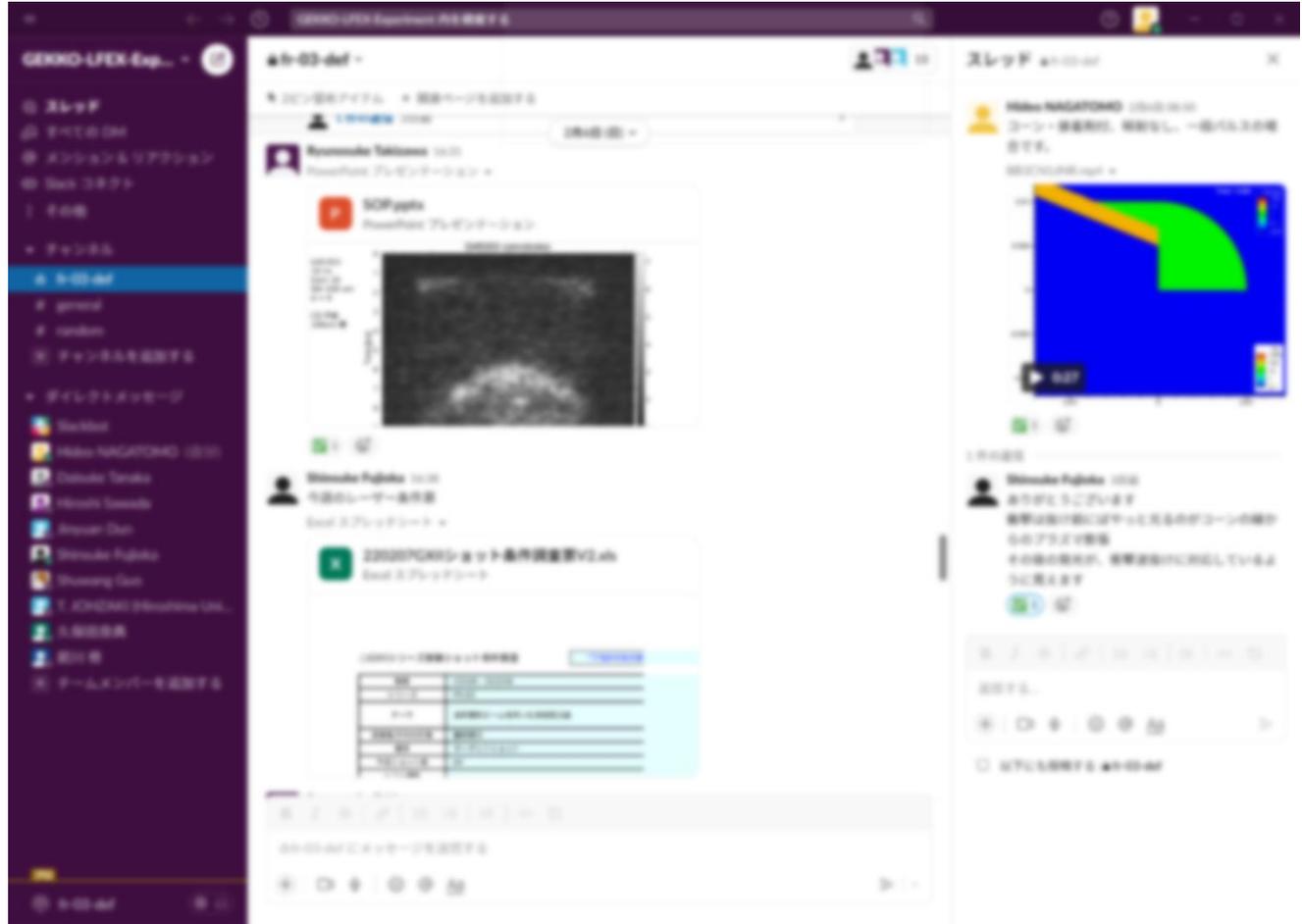


- データ所有者はSquidのアカウントが必要
- バケット毎にストレージ設定が必要
- Linuxで個別にアクセス権限を与えるより簡単で、使い勝手はとても良い。
- 安定性の向上に期待
- ポストプロセスを入れることはできないか。
(1次データを取得するより、必要な部分を必要な形式で取得できたらより便利)
- 限界は？



コミュニケーションツールとデータベースを融合できたら

- 実験期間中はSlackでコミュニケーションを取っている
 - 作業指示、記録
 - レーザーデータベース
 - シミュレーションデータベース
-
- この際の加工データ(画像、動画)を直接貼り付けているが、1次データやポストプロセスツールに直接アクセスできると実験の効率が向上する。



まとめ

- レーザー科学研究所(旧レーザー核融合研究センター、旧レーザーエネルギー学研究センター)では、大型装置の実験データとシミュレーションデータの利活用を進めてきた。
- 先端研究基盤共用促進事業(先端研究設備プラットフォームプログラム)「パワーレーザーDXプラットフォーム」の一環として、従来の実験データベースSEDNAに加え、きめ細かいアクセス制御ができるデータベースシステムが必要となる阪大CMCのONIONの利用を進めている。
- ONION-object: 従来のデータベースとスムーズな連携が可能であれば利用が広がる。今後の継続的、安定的な運用を期待する。
- ONION-file: シミュレーションデータの共有に有効である。
- 既存SEDNAとONIONの連携構築
(まだまだ知らない便利な機能があるようで使いこなせていません。)

本発表は、大阪大学レーザー科学研究所全国共同利用・共同研究データベースワーキンググループ、先端研究基盤共用促進事業(先端研究設備プラットフォームプログラム)「パワーレーザーDXプラットフォーム」、大阪大学データリテリフロンティア機構 学際共創プロジェクト における研究・開発成果を含みます。ONIONの利用においてはCMCの担当者から様々なアドバイスを頂きました。関係者の皆様に御礼申し上げます。