

核融合アーカイブ資料の 電子化とデータベースの構築

高山 有道@メタ階層ダイナミクスユニット／情報システム・セキュリティセンター／核融合アーカイブ室

村上 泉@プラズマ量子プロセスユニット／核融合アーカイブ室

2025年度 核融合科学研究所 成果報告会



レター

核融合アーカイブ資料の電子化とデータベースの構築

高山 有道^{1), 2), 3), 4)} , 村上 泉^{1), 3), 4)} 

1) 自然科学研究機構核融合科学研究所研究部, 2) 自然科学研究機構核融合科学研究所情報システム・セキュリティセンター, 3) 自然科学研究機構核融合科学研究所核融合アーカイブ室, 4) 総合研究大学院大学

Digitalization of Fusion Archive Materials and Development of their Database System

Arimichi Takayama^{1), 2), 3), 4)}, Izumi Murakami^{1), 3), 4)}

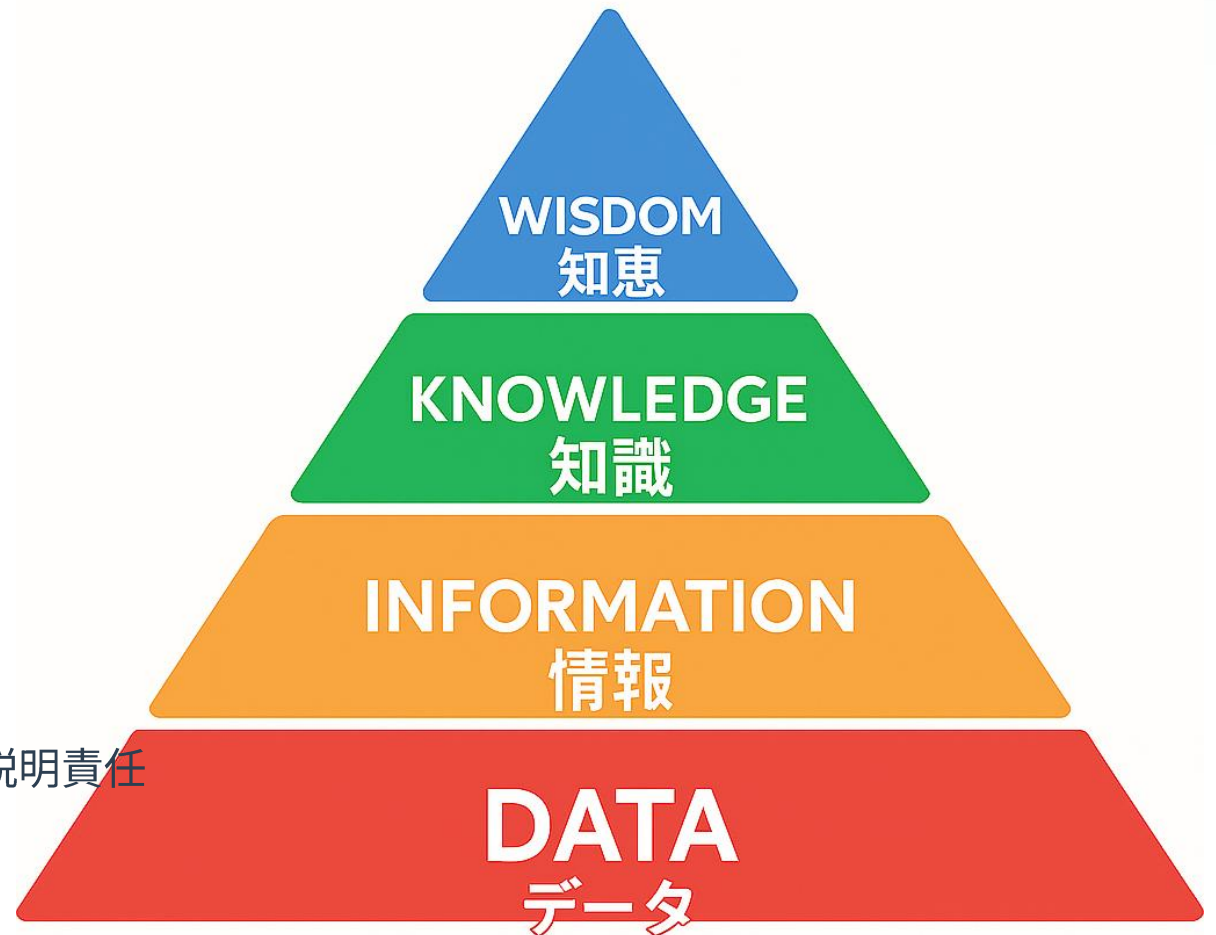
1) Department of Research, National Institute for Fusion Science, National Institutes of Natural Sciences, 2) Information Systems and Cyber Security Center, National Institute for Fusion Science, National Institutes of Natural Sciences, 3) Fusion Science Archives Office, National Institute for Fusion Science, National Institutes of Natural Sciences, 4) The Graduate University for Advanced Studies (SOKENDAI)

資料収集・公開の意義：DIKWモデル

保存された史料は、単なる「点」としての存在から、体系化されることで「知恵」へと昇華します。

- **Data (資料群):** 散逸を防ぎ収集された原資料
- **Information (整理):** 検索可能な目録・電子データ化
- **Knowledge (知識):** 研究史の文脈、成功・失敗の再発見
- **Wisdom (叡智):** 未来の核融合開発への示唆と社会への説明責任

DIKW Model



核融合アーカイブ室の歩みと課題



この図はMicrosoft 365 Copilotを用いて作成した

設立20周年の節目（2025年）

我が国の核融合研究の史料を恒常的に収集・整理。
現在、約1,000箱におよぶ膨大な資料を保有。

直面する課題

- ・ **劣化の進行:** 作成から50年を越える紙資料の劣化
- ・ **アクセスの制限:** 遠隔地からの参照が困難
- ・ **物理的限界:** 保管スペースの逼迫

成果のハイライト: オープンアクセス加速化事業 (2024年度)

27,504

電子化資料点数

135

総ページ数 (万)

700

対象ボックス数

文部科学省の事業採択を受け、短期間での大規模電子化とDBシステム構築を完遂。

電子化の実行戦略と技術仕様



保存と活用の折衷

400 dpi / 24 bit フルカラー。
不可逆圧縮PDFを採用し、判
読性とデータ軽量化を両立。



作業効率の追求

一部資料の断裁許可によりオ
ートフィーダを活用。
期間内完遂を実現。



多様な形態への対応

手書き/活字混在、図面、ロシ
ア語資料など、核融合研究特
有の多様なフォーマットを網
羅。

目録（メタデータ）の体系的整理

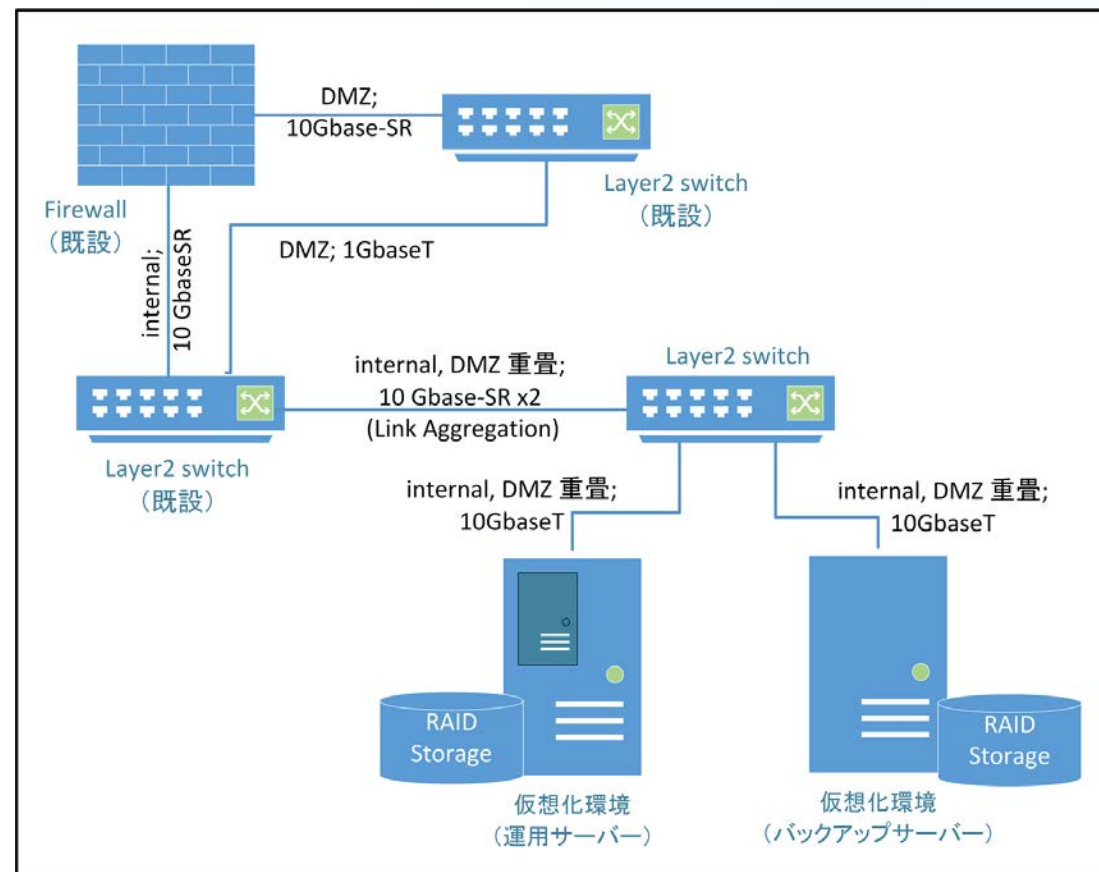
管理レイヤー	主要項目（EAD準拠）	新規管理データ
識別情報	識別番号(ID), 標題, 年代(西暦)	PDFバージョン
作成背景	出所・作成者名称, 機関情報, 研究分野	カラスペース, スキャナ設定
物理的属性	物的状態, 範囲と内容, 注記	スキャン業者・作業日
活用・公開	情報公開設定, 資料の性格, 参照ID	所蔵者, データサイズ

堅牢なシステムインフラ構成

2サーバー・クラスタ構成

- 運用サーバー: 16TB HDD x4 (zfs RAIDZ)
- バックアップ: 20TB HDD x8 (RAID5/140TB)

Proxmox VEによる仮想化環境。障害発生時にはバックアップ機へ即座に仮想マシンを移行可能。



技術スタックと開発環境



PHP / Laravel 11

堅牢なバックエンドフレームワーク。Vue.js 3を併用し、直感的なUI操作を実現。



Docker Container

nginx, MySQL, PHPをコンテナ化。開発・運用の可搬性とスケーラビリティを確保。



Google Auth API

セキュアなユーザー認証。管理者と一般利用者の柔軟な権限分離制御を実装。

洗練された検索・閲覧体験

- 高度な検索機能:
キーワードおよび複数条件の複合検索。
 - 高速ビューア:
ブラウザ上でのサムネイル表示とPDFプレビュー。
 - 申請フローの統合:
閲覧申込書の自動作成機能を搭載。
- ✓ 管理者が日々更新するメタデータの即時反映
 - ✓ 世代別バックアップによるデータ保全

核融合科学研究所 研究資料アーカイブ takayama.arimichi@nifs.ac.jp (アーカイブ管理者)

目録データ一覧 コレクション管理 ユーザー管理 ログダウンロード

閲覧・複写申請書が出力されていない閲覧済資料が存在します。閲覧・複写申請書は必ず提出をお願いします 閲覧・複写申請書をダウンロードする

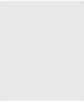
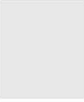
目録データ一覧

目録データ新規作成

コレクション番号 ボックス番号 ファイル番号 アイテム番号 サブアイテム番号 キーワード検索 詳細検索

キーワードを入力 検索する

公開 非公開 未確定

表紙	コレクション名(番号)	ボックス番号	タイトル	出所・作成者名称	範囲と内容	年代	識別記号 ↑
	関口忠氏資料 (002)	0145	大学等における核融合研究体制の変遷 核融合科学研究所設立の経緯 多極分散化から一本...	西川恭治 (広島大学理学部)	大学等における核融合研究体制の変遷 核融合科学研究所設立の経緯 多極分散化から一本化へ 平成9年3月 (1997.03...	1997.03	nifs-002-0145-015-003
	早川幸男氏資料 (003)	0301	第1回核融合専門部会 議事録	科学技術庁 原子力委員会 核融合専門部会	科学技術庁 原子力委員会 核融合専門部会 その1 第1回核融合専門部会議事録 (1958.05.19開催) 出席...	1958.05.19	nifs-003-0301-001-001
	早川幸男氏資料 (003)	0302	IAEA Conference on Plasma Physics and Controlled Nuclear Fusion Research	IAEA (International Atomic Energy Agency)	ザルツブルグ会議 (IAEA プラズマ核融合国際会議 (第1回)) その1 IAEA CONFERENCE ON PLASM...	1961.09.04-08	nifs-003-0302-021-001

遭遇した問題と現場での対応

⚠資料の状態不備

識別番号の重複やラベル剥離が発生。現物確認による再特定フローを構築し解決。

🕒期間の制約

1月末時点の遅延に対し、断裁判断基準の明確化とラインの増強でリカバリー。

今後の展望：AIによる価値の高度化



UI/UXの改善

利用者フィードバックに基づく検索性能の向上



AI技術の活用

自動OCR・要約・キーワード抽出の自動化



公開範囲の拡大

外部連携の強化



永久保存の議論

電子化後の現物廃棄基準の策定とデータ保全

まとめ

大量資料の電子化

約27,000点、135万ページの紙資料を短期間で電子化し、劣化リスクを大幅に低減した。

知識基盤（データベースシステム）の構築

EAD準拠のメタデータで資料の背景や文脈を含め検索参照可能な知識基盤を再構築した。

高可用性システムの構築

サーバークラスタ構成と使いやすいUI設計で、実運用に耐えるアーカイブシステムを構築した。

未来への展望

AIによるOCR精度向上や自動要約などで価値を創出し、公開範囲拡大も期待されている。