

令和7年度 成果報告書

基本情報(公開)

事業名 プログラム名		戦略的イノベーション創造プログラム(SIP)第3期 統合型ヘルスケアシステムの構築
研究開発課題名		E-1 大容量医療データの高速処理・高効率管理・高次解析基盤技術の開発
研究開発 担当者* 1	機関名	国立大学法人東京大学
	所属	生産技術研究所
	役職	教授
	氏名	合田 和生

実施期間*2	令和7年4月1日～令和8年3月31日
--------	--------------------

*1 委託研究開発契約書に定義

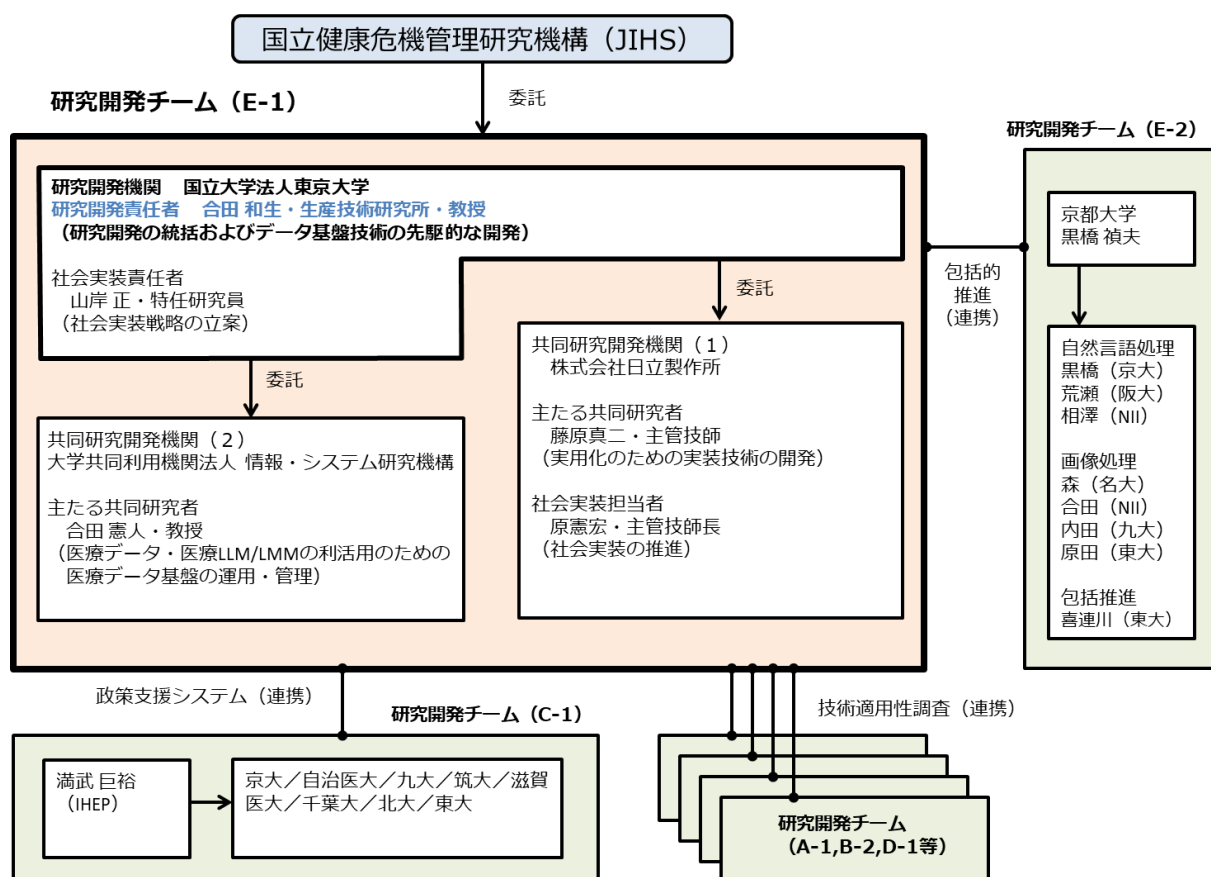
*2 年度の契約に基づき、本委託研究開発を行った期間又は中止までの期間

1. 研究開発テーマ概要(公開)

1.1 研究開発内容

本研究開発テーマでは、我が国のヘルスケアシステムで生成される国家規模の膨大な医療データを集約し、機動的にデジタルサービスを生成してヘルスケアシステムに提供可能とするデータ基盤技術確立することを目指す。SIP が構想する「医療デジタルツイン」の実現に向けた基礎を確保するものであり、具体的には、(1)一連の解析プロセスを飛躍的に高速化する技術、(2)複雑度の高い解析処理を飛躍的に高速化もしくは高精度化する技術、(3)システム管理者の意思決定を支援することによりその作業を飛躍的に高効率化する技術を確立することを目指す。また、当該技術をもとにソフトウェアを試作し、実験室環境に於いて当該技術の有効性を検証するほか、他の研究開発チームと連携の下でヘルスケアフィールドに於ける実証試験を行い応用の観点からの有効性を検証する。更に産業展開を想定した実装技術を確立する。

1.2 研究開発実施体制



1.3 研究推進のスケジュール

項目	令和5年度	令和6年度	令和7年度	令和8年度	令和9年度
(サブテーマ1) 大規模医療データを対象とする高速解析技術の開発	GPU/DPU活用 方式検討	PMEM活用方式検討・ ソフトウェア設計	ソフトウェア試作・ ベンチマーク試験 (規模限定)	ネットワーク化 複数サーバ方式検討	ソフトウェア高度化・ ベンチマーク試験 (大規模)
(サブテーマ2) 大規模医療データを対象とする高次解析技術の開発	近似問合せ 方式検討	類似検索方式検討・ ソフトウェア設計	ソフトウェア試作・ ベンチマーク試験 (規模限定)	近似処理と 類似検索の融合 方式検討	ソフトウェア高度化・ ベンチマーク試験 (大規模)
(サブテーマ3) 大規模医療データを対象とする高効率管理技術の開発	データ編成 自動化方式検討	システム構成 自動化方式検討・ ソフトウェア設計	ソフトウェア試作・ ベンチマーク試験 (規模限定)	クラウド環境への 拡張方式検討	ソフトウェア高度化・ ベンチマーク試験 (大規模)
(サブテーマ4) プロジェクトの総合的推進					
(サブテーマ5) 医療LLM/LMM開発基盤 の運用・管理			運用・管理		
(サブテーマ6) 医療データ・医療LLM/LMM の利活用のための医療データ 基盤の運用・管理			運用・管理・医療LLM/LMM開発支援		

2. 本年度の成果・進捗の概要(公開)

本年度は、サブテーマ(1)大規模医療データを対象とする高速解析技術の開発については、前年度に実施した設計に基づき、限定版ソフトウェアを試作し、中規模のベンチマーク試験を実施した。従来に比べて 3 倍程度の高速化効果を達成した。サブテーマ(2)大規模医療データを対象とする高次解析技術の開発については、前年度に実施した設計に基づき、限定版ソフトウェアを試作し、中規模のベンチマーク試験を実施した。従来に比べて 3 倍程度の高速化もしくは高精度化効果を達成した。サブテーマ(3)大規模医療データを対象とする高効率管理技術の開発については、前年度に実施した設計に基づき、限定的なソフトウェアを試作し、中規模のベンチマーク試験を実施した。従来に比べて 3 倍程度の効率化効果を達成した。サブテーマ(4)プロジェクトの統括と総合的推進に於いては、技術検討会を 12 回、個別の打ち合わせ(技術、対外発表等)を 14 回行った。更に人材の視点から、優秀な若手の研究員・技術者の積極的な配置を継続した。サブテーマ(5)医療 LLM/LMM 開発基盤の運用・管理に於いては、2024 年度「統合型ヘルスケアシステムの構築における生成 AI の活用」に於いて構築した「医療 LLM/LMM 開発基盤」について、2025 年度より、その一部の運用・管理を継続し、テーマ E-2 をはじめとする本 SIP 内の他テーマの研究開発および関連する研究開発に供した。サブテーマ(6)医療データ・医療 LLM/LMM の利活用のための医療データ基盤の運用・管理に於いては、

ては、2024 年度「統合型ヘルスケアシステムの構築における生成 AI の活用」に於いて構築した「医療データ基盤」について、2025 年度より、その一部の運用・管理を継続し、テーマ E-2 をはじめとする本 SIP 内の他テーマの研究開発および関連する研究開発に供した。

3. 成果物の公表(公開)

3.1 論文など(原著論文、学位論文、プロシーディングス、総説、解説、速報など)

論文数(総数)	(内国際誌)	(内国内誌)
6	6	0

1. Yuto Hayamizu, Masaru Kitsureawa, Kazuo Goda. Proactive Energy Management in Database Systems. ACM SIGENERGY Energy Informatics Review, Vol.4, No.5, pp.154-159, 2025.04.
<https://doi.org/10.1145/3727200.3727223>.
2. Norifumi Nishikawa, Akira Shimizu, Akira Ito, Shinji Fujiwara, Yuto Hayamizu, Masaru Kitsuregawa, Kazuo Goda. Dynamic Pruning for Recursive Joins. Companion of the 2025 International Conference on Management of Data (SIGMOD2025), pp.593-607, 2025.06.
<https://doi.org/10.1145/3722212.3724434>. (※1)
3. Yuto Hayamizu, Ryoji Kawamichi, Tsuyoshi Ozawa, Masaru Kitsuregawa, Kazuo Goda. anagodb: Offering Massive Parallelism for Database Engine. Companion of the 2025 International Conference on Management of Data (SIGMOD2025), pp.115-118, 2025.06.
<https://doi.org/10.1145/3722212.3725079>. (※1)
4. Jumpei Sato, Kazuo Goda, Hiroyasu Kiba, Taro Fujimoto, Shinji Fujiwara, Kohji Kimura, Norihiro Hara, Kenji Suzuki. Allowing Vector Attributes in Relational Database for Japanese Insurance Claims. Proceedings of the 20th World Congress on Medical and Health Informatics (MedInfo2025), Volume 329: pp.1602-1603, 2025.08.
DOI: 10.3233/SHTI251122. (※1)
5. Tarun Sreepada, A. C. Pogaku, U. K. Rage, Kazuo Goda. Accelerating Fuzzy Frequent Pattern Mining on GPUs with GPU Direct Storage. Proceedings of 2025 IEEE International Conference on Big Data (IEEE BigData 2025), pp.3045-3054, 2025.12. DOI: 10.1109/BigData66926.2025.11402218. (※1)
6. Mika Takata, Hiroki Yuasa, Kazuo Goda. Synopsis-Alloyed Index for Exact and Approximate Query Processing. IEEE Access, 14, pp.6797-6806, 2026.01. DOI: 10.1109/ACCESS.2025.3649020.

(※1)情報学分野では、査読を伴う論文を編纂した会議録が出版される学会発表は、原著論文と同等に扱われていることから、本項目に分類している。

3.2 学会発表など((国内・国際)学会口頭発表・ポスター発表、招待講演で成果を公表したもの)

1. anagodb の高速クエリ実行を実現する細粒度タスク管理実装技術, 川道亮治, 早水悠登, 小沢健史, 合田和生, DEIM2026, 2026/2/28, 国内, 口頭.
2. 分散インデックスへのシノプシス埋込みによる近似問合せ処理の検討, 高田実佳, 合田和生, DEIM2026, 2026/2/28, 国内, 口頭.
3. 不揮発メモリを含む階層型ストレージを対象とする効率的 R 木管理手法の検討, 吉岡弘隆, 合田和生, 喜連川優, DEIM2026, 2026/2/28, 国内, 口頭.
4. GPU 駆動 IO を用いたデータベース問合せ処理系の試作と初期評価, 小沢健史, 合田和生, DEIM2026, 2026/3/1, 国内, 口頭.
5. パーティションを意識した近似問合せ処理の安定性向上に関する検討, 李鐘豪, 合田和生, DEIM2026, 2026/3/1, 国内, 口頭.
6. 大規模ベクトルデータベース探索に於ける入出力効率の向上に向けた実験的考察, 呉果行, 合田和生, DEIM2026, 2026/3/2, 国内, 口頭.
7. 大規模ベクトルデータベースの検索性能と索引構築時間の実験的考察, 高尾大樹, 西川記史, 茂木和彦, 藤原真二, DEIM2026, 2026/3/2, 国内, 口頭.

3.3 診療ガイドライン、省令、基準、日本薬局方、添付文書改訂、国の技術文書(通知)等への反映

なし

3.4 研修プログラム、カリキュラム、シラバス、教材、e-learning 等の公表

なし

3.5「国民との科学・技術対話」に対する取り組み

1. データ管理ソフトウェア技術の動向や課題：黎明期から最近の生成 AI まで, 合田和生, International Disk Drive Equipment and Materials Association(IDEMA)日本 HDD 協会主催オンラインセミナー, 2025/5/27, 国内.
2. 実世界・ソーシャルビッグデータ高度処理・可視化基盤, 合田和生, 東大駒場リサーチキャンパス公開, 2025/5/30-31, 国内.
3. ビッグデータを活用するための基盤ソフトウェア技術, 合田和生, 東京大学柏キャンパス一般公開, 2025/10/24-25, 国内.
4. データプラットフォーム技術の潮流：約 150 年の歴史から眺める AI 時代, 合田和生, Storage Networking Industry Association(SNIA)日本支部主催ストレージトレンドセミナー, 2026/2/27, 国内.
5. いま、「データ&ナレッジ」が熱い！～AI 時代を支えるデータ&ナレッジマネジメントの最前

線～，葛西智基、西川記史，DEIM2026 ランチョンセミナー，2026/3/4，国内．

3.6 その他

1. 【共同発表】日立と東大、ビッグデータ検索を最大 135 倍高速化する「動的プルーニング技術」を開発——製造業のトレーサビリティや医療・金融分野のデータ利活用を支援し、社会課題解決に貢献——，<https://www.iis.u-tokyo.ac.jp/ja/news/4794/>，<https://www.hitachi.com/ja-jp/press/articles/2025/06/0619a/>，2025/6/19.

以上