別紙 3

厚生労働科学研究費補助金
政策科学総合研究事業 (臨床研究等 ICT 基盤構築・人工知能実装研究事業)

分担研究報告書
クラウド上の医療 AI 利用促進のためのネットワークセキュリティ構成類型化と実証及び施策の提言

研究代表者　岡村 浩司　国立成育医療研究センター 室長

研究要旨

人工知能技術の急速な進展は医療分野にも多大な可能性をもたらし、医療の質向上はもちろんのこと、効率化、地域格差の解消、医療従事者の負担軽減などを目指した取り組みが活発に展開されている。国立成育医療研究センターでは AI ホスピタル事業を契機として、職員を対象としたデータサイエンス教育、啓蒙活動を出発点に、グラム染色画像を入力とした感染症起因菌同定支援、尿沈渣顕微鏡写真からマルベリー小体を自動検出するファブリー病スクリーニングなど、実診療における補助手段としての医療 AI サービスの開発を進めてきた。これらのサービスは医療 AI プラットフォーム技術研究組合との連携のもとでコンテナ化され、クラウドからの公開と仮想デスクトップ基盤を介した安全な通信環境の構築が図られ、さらには運用コストの最適化を目的としたサーバレスアーキテクチャへの移行も行われた。一方、大規模言語モデルの技術革新、AI エージェントの登場によってアプリケーション開発は AI 支援型から駆動型へと移行しつつある。AI 駆動型開発による再構築を試み、開発効率、コード品質、セキュリティ実装の各面において従来手法との比較を行ったところ、定型的な作業における大幅な効率化の一方で、医療、情報、セキュリティの各専門知識を連携させた指示と最終的な品質確認の不可欠性を確認した。こうした変化と並行し、医療 AI サービスの普及を阻む要因として、情報管理部門が電子カルテネットワークへの接続に過度な恐怖感を抱き、新たな取り組みを忌避する事例が多いことも明らかとなった。医療 AI サービスは推論結果を返すのみで個人情報を保持しない構造であり、ランサムウェア攻撃の主要標的となった事例は確認されていない。責任回避を優先する組織的傾向がサービス普及の機会を逸失させている現状がある。さらに、生成 AI および AI エージェントの登場は、セキュリティ対策と人材育成の両面において従来の専門家依存型の体制からの脱却を可能にしつつあり、基本的な脆弱性への対応や未経験者による開発支援といった領域での有効性が示されている。加えて、クラウドプロバイダが提供するマイクロサービスアーキテクチャはゼロトラストの実装を容易にする特性を多く備えており、コスト削減とセキュリティ強化の同時実現という観点から、生成 AI 登場以前に設計された既存の SASE 等のセキュリティツールに代わる、新たなセキュリティ対策としての地位を固めつつある。本研究は、これら一連の技術的、組織的、制度的課題を横断的に検討し、電子カルテ端末から医療 AI サービスを安全かつ安心して利用できる環境の実現に向けた技術的、運用的指針を提示する。

A. 研究目的

人工知能技術は、ディープラーニングの登場から生成 AI さらには AI エージェントへと急速に進化を遂げ、医療分野におけるその影響はもはや研究開発の領域にとどまらず、サービスの開発、運用、セキュリティ対策のあり方そのものを根本から塗り替えつつある。こうした技術変革の潮流のなかで、皮膚がん診断をはじめとする数多くの有用性が実証されて以来、医療の質向上、効率化、地域格差の解消、医療従事者の負担軽減等を目指した研究開発が世界規模で活発に展開されてきた。我が国においても内閣府主導の戦略的イノベーション創造プログラムのもとで2018年よりAIホスピタルの取り組みが開始され、国立成育医療研究センター(NCCHD)は医療 AI プラットフォーム技術研究組合(HAIP)とともに採択を受け、感染症起因菌同定支援、腎細胞がん病理画像グレーディング支援、ファブリー病スクリーニングという三つの医療 AI サービスをコンテナとして実装し、限定的な公開ながら実運用可能な状態へと到達した。これらのサービスは Linux サーバ上のウェブアプリケーションとして開発され、コンテナ化と仮想デスクトップ基盤(VDI)を介したクラウド公開によって開発効率、運用コスト、セキュリティの各側面での改善が図られてきたが、電子カルテ端末からの安全なアクセスの実現、さらには患者および市民を含む一般ユーザへの開放という最終目標に向けては、依然として多くの技術的、組織的課題が残されていた。

こうした状況に対し、本研究が直面してきた課題の性質は、研究の進展とともに大きく変容してきた。当初はゼロトラストセキュリティモデルの実装とSASEアーキテクチャの試験的導入を中心とした技術的検討が主眼であったが、2024年末頃よりAIコーディングエージェントの実用化が急速に進み、開発のあり方そのものが根本から変わりつつある。大規模言語モデルと AI コーディングエージェントを組み合わせた AI 駆動型開発の実践において、ユーザが日本語の指示を与えるだけで、開発環境の整備からコーディング、テスト、デバッグ、ドキュメント作成に至るソフトウェア開発の全工程をエージェントが自律的に推進する状況が現実のものとなった。ファブリー病スクリーニングシステムの再構築を題材とした検討では、大学院生による未経験者開発においても物体検出 AI モデルおよびウェブアプリケーションの構築が実現され、AIエージェントが人材育成のツールとしても高い有効性を持つことが示された。また、仮想マシン上でのウェブサーバ運用からサーバレスアーキテクチャへの移行は、コスト効率の観点のみならず、サーバレスの各構成要素がアイデンティティベースのアクセス制御、すなわちゼロトラストの思想に沿って設計されているという特性から、セキュリティ強化の観点においても極めて合理的な選択であることが確認された。一方で、クラウドサービスへの依存度が高い設定管理の複雑さという課題も明らかとなったが、この領域こそ AI エージェントの活用が最も効果を発揮しうる領域でもある。

セキュリティ対策に関して、国家サイバー統括室によるアタックサーフェス管理(ASM)の継続的な調査対象となっている本研究環境では、従来の開発手法において OpenSSL やウェブサーバのバージョン問題、WordPress の残存ファイルといった指摘を受けてきた経緯がある。しかし AI 駆動型開発への移行後は、こうした指摘がなくなったことが確認されている。AIエージェントがSQLインジェクション対策や XSS への対処を含むセキュリティ実装を開発プロセスの中で

自律的に提案、適用する能力を持つことが、この変化の主要な要因と考えられる。従来、セキュリティ対策は専門家の知識と経験に依存する領域とされてきたが、AIエージェントの活用によってその民主化が現実のものとなりつつあることを、本研究の経験は具体的な形で示している。そして2026年になり、Anthropicが発表したClaude Code Securityの登場は、この状況をさらに決定的に変化させるものであった。コードベース全体のデータフローを横断的にトレースし、複数コンポーネントにまたがる複雑な脆弱性パターンを検出する能力、そして敵対的検証プロセスによる誤検知の大幅な削減は、従来の専門家集団による脆弱性探索という作業の前提を根底から覆すものであり、医療機関の情報管理者が抱いてきた過度な恐怖感や責任回避的な姿勢に対しても、客観的かつ実証的な安全性の根拠を提示できる可能性を秘めている。

本研究の目的は、以上の技術的および社会的変化を踏まえ、電子カルテ端末から医療AIサービスを安全かつ安心して利用できる環境の実現に向けた技術的、運用的指針を提示することにある。具体的には、AI駆動型開発がもたらす開発効率、コード品質、セキュリティ実装への影響を実証的に検討し、サーバレスアーキテクチャへの移行によるゼロトラスト実装の現実的な経路を示し、Claude Code Securityに代表される最新のAIネイティブなセキュリティツールの位置づけを医療情報セキュリティの文脈において整理する。また、これまで専門家依存、人材不足という構造的問題として語られてきたセキュリティ対策と開発体制の課題が、AIエージェントの積極的な活用によって解決可能な問題へと性質を変えつつあることを示し、医療、情報、セキュリティの各専門領域が従来とは異なる形で連携しながら、患者および市民の参画を促しつつ医療AI全体の発展へと繋げていくための新たな体制構築の方向性を論じることを、本研究の目的とする。

B. 研究方法

感染症起因菌同定支援およびファブリー病スクリーニングともに、顕微鏡画像のアノテーションにはMicrosoft VoTTを用いた。ラベルと位置情報をJSON形式で出力し、画像分類のための切り出しや、物体検出のためのYOLO形式への出力を独自Pythonスクリプトで行った。

感染症起因菌同定支援では、NCCHDの患者から得られた画像データを用い、腸内細菌科のバクテリア、エンテロコッカス属のバクテリア、バシラス属のバクテリア、レンサ球菌属のバクテリア、ヘモフィルス属のバクテリア、シュードモナス属のバクテリア、コアグラーゼ陰性ブドウ球菌、黄色ブドウ球菌、緑膿菌、肺炎レンサ球菌、化膿レンサ球菌、リステリア、コリネバクテリウム、淋菌などのグラム陰性球菌、さらに真菌であるカンジダを区別して検出するためのアノテーションを手作業で行なった。

まずTensorFlowを利用して画像の分類を試みた。データ拡張にはKerasを利用した。ImageNetで訓練されたInception V3の転移学習をHitachi SR24000/DL1を用いて実行した。物体検出については、Microsoft Azureコンピューティングインスタンスのサイズ Standard_NC12s_v3を利用して構築されたHAIPのAI開発基盤を利用した。アルゴリズムは、PyTorchを基盤とするYOLOv5を採用した。CentOS 7に設定したYOLOv5を用いて訓練を行い、物体検出モデルを作成した。

ファブリー病スクリーニングでは、NCCHDだけでなく、聖マリアンナ医科大学、上海の復旦大学の3機関から患者の尿検体を

収集した。各検体 10 mL を 3000 rpm で 10 分間遠心分離し、得られた沈渣を再懸濁させ顕微鏡化で観察、静止画及び動画の撮影を行なった。尿沈渣の構成成分は、マルベリー小体に加え、赤血球、白血球、結晶、細菌、精子、扁平上皮細胞、マルベリー小体様構造物、その他、合計 9 カテゴリに分類した。

Amazon EC2 のインスタンスタイプ g4dn.xlarge を利用し、Amazon Linux 2 に設定した YOLOv8 を用いて訓練を行い、物体検出モデルを作成した。

コンテナ作成は、CentOS 7 に設定した Docker にて行い、デプロイ確認は Minikube を利用した。ウェブアプリケーションとしての公開は Amazon ECS を利用し、また、HAIP サービス事業基盤からの公開は Azure Kubernetes Service を利用した。VDI クライアントは Microsoft Remote Desktop、サーバは Azure Virtual Desktop でクラウドの Windows にアクセスし、そのブラウザから HAIP ラボ基盤、およびサービス事業基盤にデプロイされているコンテナにアクセスさせた。ウェブカメラの制御は JavaScript のメディアストリーム API に含まれている getUserMedia を利用し、クラウドの Windows に接続されているデバイスを動作させた。

遺伝カウンセリング支援は、Amazon SageMaker から AWS SDK for Python を利用して Amazon Bedrock を利用する環境を構築し、オレゴンリージョンの Anthropic Claude 3 Sonnet および Claude 3.5 Sonnet を利用して比較を行った。サービス間のアクセス許可は AWS IAM のロール作成にて行った。RAG は最初は Amazon Kendra にて構築したが、後に Amazon Bedrock Knowledge Bases の利用に移行した。

サーバレスコンピューティングは、AWS Lambda に加え、Amazon API Gateway、Amazon S3、Amazon CloudFront、AWS WAF を用いて基本的なウェブサービスを構築し、必要に応じて Amazon Bedrock などを組み込んだ。サービス間のアクセス許可については AWS IAM を利用した。

AI コーディングエージェントについては、VS Code の拡張機能である Cline、また、Amazon Q Developer およびその後継である Kiro を試し、LLM としては主に Anthropic Claude Sonnet を用いた。

C. 研究結果

高度先進医療を提供する NCCHD は、免疫不全、臓器移植後に免疫抑制剤の投与を受けるなど、感染症に関してハイリスクな患者を多数抱えている。適切な抗生物質の選択など治療方針の決定は患者の予後に直結し、また不必要な抗生物質の使用は耐性菌の問題もあり、責任ある対応が求められている。そこで菌血症患者検体の顕微鏡写真から起因菌を迅速に同定する医療 AI の開発を行なった。

小児の菌血症患者の血液培養からグラム染色を行なって得られた顕微鏡写真に、生化学反応や質量分析によって決定された細菌や真菌の種類を正解ラベルとして教師データを作成した。まず、10 μm 四方のクロップに対してアノテーションを行い、グラム陽性

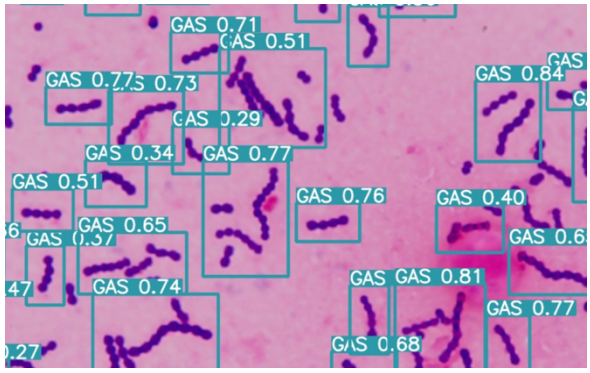

図 1 人食いバクテリア、化膿レンサ球菌の検出

桿菌、グラム陰性桿菌、グラム陽性球菌、グ

ラム陰性球菌、それから背景の 5 分類を試みた。Inception V3 の転移学習により訓練を行ったモデルに対し、ランダムに切り出した 10 μm 四方のクロップでテストを行った。多数のクロップに対する結果のうち、背景を除いた多数決を取ることで良好な結果が得られることを確認した。

感染症起因菌をより詳しく同定するため、15 細菌および 1 真菌を物体検出により区別するアノテーションを行なった。赤血球を加えた 17 ラベル、23,753 枚の写真から 347,234 クロップの切り出しで訓練を行なった。モデルは YOLOv5x を採用し、V100 を搭載する HAIP の AI 開発基盤で、150 エポック 34 時間をかけて独自モデルを完成させた(図 1)。

図 2 VDI でクラウドのウェブカメラを利用

本システムでは、互いに異なる培養条件に対応する複数の学習済みモデルを構築し、判別対象菌の培養条件に応じたモデルを選択して解析を行うことで、外観が類似した菌種であっても高精度な判別を可能としており、この点については特許を出願した。

扱う医療データを安全にやり取りする目的で、コンテナを VDI 環境で公開している。ユーザはリモートデスクトップクライアントを利用して、クラウドのブラウザにアクセスし、デスクトップ画像の通信で利用する形態である。顕微鏡写真はデータをアップロードすることもできるが、顕微鏡に接続されたコンピュータからの簡易的な利用を見据え、ディスプレイに表示された画像をウェブカメラで撮影して検出できるように設計されている(図 2)。その際には、ユーザが手元で操作するウェブカメラからのリアルタイム入力は、VDI によりクラウド側に存在するウェブカメラと直結している。

次にファブリー病であるが、この疾患は分解酵素の欠失や活性の低下により不必要な糖脂質が細胞内に蓄積し、年月を経て体全体に障害を与える先天性の代謝異常症である。10 歳頃になって手足の痛みなどの症状が現れるが、診断は難しく、原因が判明した頃には心臓や腎臓などの損傷が進み、40 歳頃に亡くなるケースが多々見られる。酵素補充療法などの治療が確立しており、早期に発見することがきわめて重要となる。

患者の尿には、糖脂質が蓄積した糸球体上皮細胞由来の特異的な形状の物体が観察されることが報告されていて、マルベリー小体と呼ばれている。NCCHD では独自にサンプルを収集し、また聖マリアンナ医科大学、上海の復旦大学小児医療センターの協力も得て、顕微鏡写真にマルベリー小体が写っているか否かを判別できる医療 AI モデルを開発し発表した。限られた数のサンプルから大量

の写真撮影を行い、尿に含まれるマルベリー小体、マルベリー小体様物質、赤血球、白血球、扁平上皮細胞、精子、細菌、結晶に加え、その他の計9種類を識別する物体検出モデルを作った (図 3)。

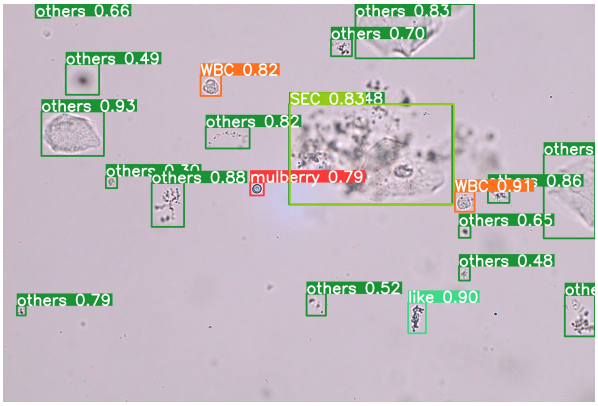

図 3 尿沈渣顕微鏡写真からのマルベリー小体検出

その後、AI コーディングエージェントが利用できるようになり、プログラミング未経験であった小児科医の大学院生にモデルの再構築とウェブアプリケーション作成に取り組ませ、課題等を調査した。通常、ウェブアプリケーション開発には複数のプログラミング言語やフレームワーク、インフラ知識の習得が前提とされるが、AI コーディングエージェントの活用により、そのような広範な事前学習なしにある程度機能するアプリケーションの構築が可能であることが示された。サーバレスアーキテクチャにおける個別サービスの詳細仕様についても、AI エージェントが代替して実装することで、未経験者の学習負担が大幅に軽減された。一方で、AI エージェントの出力が確率的であることに起因するリスクへの対応や、広範な権限を持つ AI への不安、そして最終的な品質確認における人的レビューの必要性など、未経験者が直面する固有の課題も明らかとなった。プログラミングスキルそのものの重要性は依然として認識されており、AI エージェントの活用が開発参入の障壁を下げる一方で、それを適切に活用、監督するための素養は引き続き求められることが示唆された。

経験者の立場からは、従来の開発手法と比較して大幅な効率向上が認められた。特にウェブサービスにおけるフォーム処理、ファイルのアップロード機能、および CRUD 操作においてコーディング負担の軽減が明らかであった。エラー対応が網羅的に実装され、一貫したコーディング規約が自動的に適用された。コンテナ化においては、環境構築の効

図 4 AI 駆動型開発によるウェブアプリケーション

率化とイメージサイズの軽量化が顕著であり、モニタリング等の設定も併せて構成された。サーバレスアーキテクチャの採用においては、個別サービスの詳細仕様を AI エージェントが代替して実装することによる学習負担の軽減が認められ、プログラミング未経験者にとって特筆すべき成果であった(図 4)。

AI エージェントによるセキュリティ実装として、コンテンツセキュリティの領域では XSS 対策、SQL インジェクション対策、ファイルアップロード制御が自動的に組み込まれた。セキュリティヘッダーの設定においては、他サイトからの不正アクセス制御、悪意あるコンテンツの読み込み防止、クリックジャッキング対策などの基本的防御機構が適用された。認証、認可の基本実装としては、安全なユーザー認証トークンの管理、セッション管理、およびアクセス制御が実装された。セキュリティパターンの統一、継続的な脆弱性評価、リアルタイムモニタリング、およびインシデントへの自動対応が標準的に組み込まれた。一方で、一つの指示によってあらゆるセキュリティ対策が網羅されるわけではなく、人的レビューの実施、定期的な更新、セキュリティポリシーの見直し、安全な MCP の利用、および固有要件への個別対応が別途必要であることが明らかとなった。また、今回のような医療 AI サービスの開発において利用可能な情報セキュリティチェックリストを Claude Sonnet 4.6 を用いて作成した(資料)。

大学院生からは、AI エージェントの活用により機能するアプリケーションの構築が可能であったとの評価が得られた。一方で、確率的な動作に起因するリスクへの対応、広範な権限を持つ AI への不安、プログラミングスキルの重要性の再認識、および人的な最終確認の必要性についての懸念も示された。AI エージェントを活用した標準化された開発手法への移行が効率的な人材育成につながる可能性が示唆されるとともに、従来型の人材戦略からの脱却という観点からも有効であるとの見解が述べられた。課題は残るものの、当初は悲観的に捉えていた状況に対して楽観的な展望を持って取り組んでいきたいとの感想が得られた。

日本人における頻度は欧米人と比べて高く、7000 人に 1 人と言われている。7000 検体に 1 検体、マルベリー小体が存在するか否かを医師や臨床検査技師が手作業で調べることはきわめて苦痛で困難であり、見落としも多くなると考えられる。マルベリー小体はあったとしても、多数確認されるわけでもない。この物体検出は動画にも対応しており、実際には、大量の写真あるいは動画を自動的に撮影し、自動的に検出するスクリーニングとしての開発を進めており、毎年各学年で行われる学校検尿と結びつけた社会実装を目指している。

最後に、遺伝カウンセリングを生成 AI で代替する試みであるが、まず、日本遺伝カウンセリング学会と日本人類遺伝学会が実施している認定遺伝カウンセラーの認定試験でどれほどの点数を取ることができるのか確認してみた。大規模言語モデルは Amazon Bedrock から API を介して Claude 3 Sonnet を利用し、2020 年度の基礎問題全 26 問の日本語原文をプロンプトとして入力したところ、正解率は 48%であった。選択肢を正解まで絞り切れていない回答が 11 問あり、これらには半分の点数を与えて計算しているが、合格レベルにはない状況である。その後、Claude 3.5 Sonnet が発表されたので改めて試したところ、正解率は 85%で合格レベルまで跳ね上がった。試験問題を解答するようシステムプロンプトとして指示を与えているだけでこ

のような結果が得られ、最近では基盤モデルとも呼ばれる大規模言語モデルは日進月歩の改良が進められていることがよく分かる。正解が得られなかった点については、関連情報を含むウェブページの HTML を集め、RAG(検索拡張生成)の構築により正解を出力できるよう改変することができた。

マルチモーダルであるため、家系図を理解することもでき、関連する論文を PDF のまま読み込んで患者向けに分かりやすい情報提供を行うこともできる。実際のカウンセリングにおいても、役に立つ情報を引き出せているが(図 5)、そもそも医師の指導のもとに行われなければならず、ハルシネーションやプライバシーの問題など解決しなければならない問題が多く残っている。

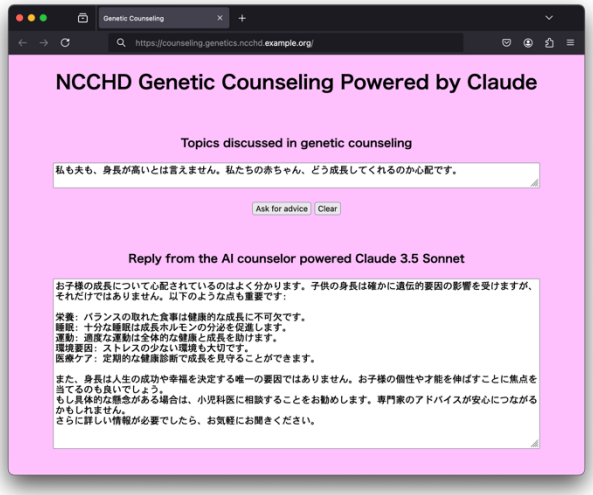

図 5 遺伝カウンセリングを生成 AI で代替

これら医療 AI サービスは開発段階ということもあるが、そうでなくとも使用頻度が高いとは言えず、クラウドの仮想マシンやコンテナとしての運用では、24 時間 365 日の連続稼働となり、ほとんど使われていないのに課金される状態が続くことになる。これに対し、サーバレスアーキテクチャを採用した場合は必要最小限のリソースで、API のコール数とデータ転送量に対してのみ課金が発生するため、コストを大幅に下げることが可能である。サーバーのセットアップ、メインテナンス、スケーリングといったインフラ管理からも解放される。各コンポーネントには自動的に復旧される仕組みも備わり、耐障害性の高い仕組みを作り上げることができ、作業を進めている(図 6)。

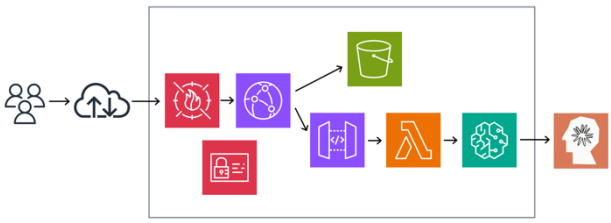

図 6 サーバレスのビルディングブロック例

D. 考察

従来のクラウドコンピューティングでは、1 つの仮想マシンに数多くのサービスや機能を載せ、運用を行っていたが、スケーラビリティ、独立性、柔軟性、コスト効率、開発速度などの点からクラウドプロバイダ自体がマイクロサービスアーキテクチャに移行し、クラウドユーザに対して各種多様なサービスを提供している。これらがサーバレスのビルディングブロックとなり、ユーザが用意する仮想マシンが不要となる。サーバレスを採用することは、セキュリティの観点においても多くの利点をもたらす。攻撃範囲の限定、迅速な脆弱性対応、動的な実行環境による攻撃対象からの回避、細粒度の最小権限付与によるリスク低減、監視ログの細分化による問題特定の容易化といった特性は、医療情報システムにおけるセキュリティ要件と親和性が高い。このような環境では、従来のネットワーク境界に依存したセキュリティ対策は成り立たず、ID ベースのアクセス制御が必要となるが、AWS IAM、Microsoft Entra ID、

Google Cloud IAM といったサービスがこれを担い、ゼロトラストセキュリティモデルの実現を見通し良く支援する。

サーバレスアーキテクチャへの移行は、その利点が明確である一方、各クラウドプロバイダが提供する個別サービスの詳細仕様の習得や、既存システムからの移行作業における複雑な設定変更など、開発者にとって相応の学習コストと実装負担を伴うものであった。しかしながら今回の取り組みを通じて、AI コーディングエージェントの活用によりこれらの障壁が大幅に低減されることが示された。AI エージェントは各クラウドサービスの詳細仕様を代替して実装するのみならず、コンテナ化における環境構築の効率化、イメージサイズの軽量化、モニタリング設定の自動構成といった移行作業全般を効率的かつ一貫した品質で遂行する。さらに、XSS 対策、SQL インジェクション対策、セキュリティヘッダーの設定、認証、認可の基本実装といったセキュリティ対策が標準的に組み込まれることから、サーバレスへの移行作業そのものをより安全に進められることが期待される。すなわち、AI コーディングエージェントの活用は、サーバレスアーキテクチャが本来持つセキュリティ上の利点を損なうことなく、その移行プロセスを加速させる手段として有効であると考えられる。

今回の結果は、AI コーディングエージェントが定型的な実装作業において顕著な効率化をもたらすことを示している。標準的な実装の自動化、一貫性のある対策の適用、セキュリティパターンの統一といった利点は、特に未経験者にとって開発への参入障壁を大きく下げるものである。しかしながら、本取り組みを通じて、むしろ経験者が AI エージェントを活用することによる恩恵がより多岐にわたることが示唆された。AI エージェントは定型的な作業を効率的に処理する一方、一つの指示によってあらゆるセキュリティ対策が網羅されるわけではなく、実運用に向けては医療、情報、セキュリティの各専門知識を連携させた立場からの具体的な指示と最終的な品質確認が不可欠である。

患者情報保護をはじめとする医療業界固有の要件やセキュリティ対策について、生成 AI が初期段階から網羅的な対応を行うわけではないことが明らかとなった。現状、電子カルテネットワークから医療 AI サービスへの直接アクセスは許容されておらず、インターネット側への接続は NTP サーバとの同期や MDM サーバの認証に限られているなど、医療情報システムには固有の制約が存在する。このような環境においては、利用者側が明確な方向性と具体的な指示を持って AI エージェントに臨む必要があり、医療分野と情報分野のこれまでとは異なる新しい形の連携が重要となる。セキュリティポリシーの継続的な更新、安全な MCP の利用、および固有要件への個別対応を組み合わせた、医療、情報、セキュリティの専門家による包括的な管理体制の構築が求められる。

医療機関の ICT セキュリティ分野では、専門家の慢性的な不足、人材確保の地域格差、不公平な処遇といった深刻な人材課題が存在する。AI エージェントの活用は、常時監視体制の実現、異常検知の高度化、即時対応、インシデント分析の効率化、最新セキュリティ情報の継続的更新、ベストプラクティスの自動適用、規制要件への適合性確認、脆弱性評価の自動実施、パッチ適用の自動評価とロールバック保証付き運用管理を通じて、これらの課題解決に寄与する可能性がある。さらに、教育の自動化、対応訓練の実施支援、マニュアル類の自動更新といった人材育成面での活用も期待される。AI エージェントを活

用した標準化された開発手法への移行は従来型の人材戦略からの脱却を促し、医療機関における一般的な業務運営の一部として AI エージェントが機能する未来、そして今後の若い世代がこれらの課題を独自に解決してゆく新たな道筋を拓くものと考えられる。今後の医療 AI 開発において AI エージェントの重要性は無視できないものとなっており、医療、情報、セキュリティの専門家が連携しつつ AI エージェントを適切に指示、監督できる体制の整備が急務である。

E. 結論

開発を行った医療 AI サービスに対し、運用コストの最適化を主な動機としてサーバレスアーキテクチャへの移行を進めた。その過程でマイクロサービスアーキテクチャがゼロトラスト実装を容易にする特性を多く備えていることを理解することができた。各構成要素がアイデンティティベースのアクセス制御を前提として設計されているサーバレス環境は、概念としては提唱されながらも具体性に乏しいと感じられてきたゼロトラストを、自然な形で実装へと落とし込める基盤となりうる。一方で、ゼロトラストを標榜する統合的な商用セキュリティサービスは依然として複雑で分かりにくく、必要性が不明瞭な高機能とそれに付随する高価格を前にすれば、医療機関が気安く導入できるものではない。また、生成 AI 登場以前に設計された SASE をはじめとする既存のセキュリティツールの多くは、サーバレス運用を前提としておらず、現在の開発、運用環境の実態と乖離しつつあることも、本研究を通じて明らかとなった課題の一つである。

こうした状況に対し、この一年間で起きた変化は、当初の想定をはるかに超えるものであった。AI コーディングエージェントの実用化により、ソフトウェア開発は AI 支援型から AI 駆動型へと本質的なシフトを遂げ、開発効率、コード品質、セキュリティ実装の各面において従来手法との明確な差異が生じている。国家サイバー統括室による ASM の継続的な調査において、従来の開発手法では指摘を受けていた OpenSSL のバージョン問題や WordPress の残存ファイルといった脆弱性が、AI 駆動型開発への移行後には一切検出されなくなったことは、その効果を客観的に示す具体的な成果である。SQL インジェクション対策や XSS 対策をはじめとするセキュリティ実装が AI エージェントによって開発プロセスに自律的に組み込まれるようになったことは、これまで専門家の知識と経験に依存してきたセキュリティ対策の民主化という観点から、本質的な意義を持つ。

そして 2026 年、Anthropic による Claude Code Security の発表は、医療情報セキュリティの文脈においても見過ごすことのできない画期的な転換点となった。コードベース全体のデータフローを横断的にトレースし、複数コンポーネントにまたがる複雑な脆弱性パターンを検出する推論ベースの能力と、AI が自らの検出結果に異議を唱える敵対的検証プロセスによる誤検知の大幅な削減は、従来の専門家集団による脆弱性探索という作業の前提を根底から変えるものである。重大度のみならず信頼度スコアを付与することで対処の優先順位付けを合理化し、検出から修正パッチの提案、適用までをシームレスなワークフローで完結させるこの仕組みは、セキュリティ専門人材の慢性的な不足という医療機関が長年抱えてきた構造的問題に対して、現実的な解を提示するものといえる。先日発表された Claude Mythos はそのレベルさらに引き上げていると報道されている。これまでランサムウェア被害の多くが引き起

こされてきたパスワード設定の不備などの基本的な脆弱性は、こうした AI ネイティブなセキュリティツールによって大幅に改善される可能性があり、医療 AI サービス自体が攻撃の主要標的となった事例が確認されていないという現実と合わせて考えれば、電子カルテネットワークへの接続に過度な恐怖感を抱き新たな取り組みを忌避してきた医療機関の情報管理者に対し、実証的な根拠をもって安全性を説明できる環境が整いつつある。

人材育成の観点においても、状況は大きく変わった。AI エージェントを活用することで、未経験の大学院生が物体検出 AI モデルとウェブアプリケーションの開発を独力で完遂できたことは、経験者や専門家への依存を前提とした従来の人材戦略からの脱却が現実のものとなったことを示している。AI エージェントは開発ツールであると同時に学習のツールでもあり、次世代の医療従事者や研究者が独自に課題を発見し解決していく可能性を、本研究の経験は具体的な形で示唆している。医療、情報、セキュリティという三つの専門領域がこれまでとは異なる形で連携しながら、患者および市民の参画を促しビッグデータに依存する医療 AI のさらなる発展へと繋げていくためには、こうした新しい開発、運用体制の構築を前向きに捉え、悲観的になりがちであった状況を積極的に転換していく姿勢が求められる。セキュリティ対策の実績を着実に積み重ねながら、管理者、患者、市民に対して分かりやすい説明を継続的に行うという地道な作業の重要性は変わらないが、その作業を支える技術的基盤は、この一年間で質的に異なるものへと変貌を遂げたといえる。

F. 研究発表

岡村 浩司, 山中 宏勇. AI 駆動型開発がもたらす医療 AI サービス構築の革新と課題. *第 45 回医療情報学連合大会*, **4-B-1-03** (2025)

Yamanaka H, So T, Sakamoto N, Aoto S, Li XK, Wang Y, Shen Q, Migita O, Kosuga M, Okamura K. Automated detection of mulberry bodies in urinary sediment for non-invasive Fabry disease screening. *Clin. Chem. Lab. Med.* [doi: 10.1515/cclm-2026-0345] (2026)

Hayashi H, Kashizuka E, Sakata K, Motomiya N, Asakawa Y, Hayasaki M, Yokoi T, Yoshida T, Nishina S, Okamura K. Detection of pediatric cataracts through non-invasive facial photography using deep convolutional neural networks for early diagnosis. *BMC Ophthalmol.* [doi: 10.1186/s12886-026-04795-9] (2026)

Yamanaka H, Okamura K. Exploring deep learning and data requirements through image classification of *Erigeron annuus* and *Erigeron philadelphicus*. *BMC Res. Notes* **19**, 127 (2026)

Fukami M, Okamura K, Sasaki S, Kagami M, Dateki S. Chromosomal and hormonal factors involved in human sexual dimorphism. *Endocr. J.* **73**, 175–181 (2026)

Miyai M, Tanaka-Yachi R, Aizawa K, Okamura K, Kusuhara H, Akutsu H, Nakamura K. Removing extracellular matrix from the cell surface prior to seeding enhances the adhesion of primary human hepatocytes to the culture vessel. *Genes Cells* **30**, e70059 (2025)

Taniguchi K, Hasegawa F, Okazaki Y, Hori A, Ogata-Kawata H, Aoto S, Migita O, Kawai T, Nakabayashi K, Okamura K, Fukui K, Wada S, Ozawa K, Ito Y, Sago H, Hata K. Approaches to evaluate whole exome sequencing data that incorporate genetic intolerance scores for congenital anomalies, including intronic regions adjacent to exons. *Mol. Genet. Genomic Med.* **13**, e70092 (2025)

Shibata M, Umezawa A, Aoto S, Okamura K, Nasu M, Mizuno R, Oya M, Yura K, Mikami

S. Applicability of the regression approach for histological multi-class grading in clear cell renal cell carcinoma. *Regen. Ther.* **28**, 431–437 (2025)

G. 知的財産権の出願

岡村 浩司, 仁科 幸子. 乳幼児の視覚異常スクリーニング方法、視覚異常スクリーニングシステム及び表示システム. 特願 2026-060427 (2026 年 4 月 1 日)

松井 俊大, 岡村 浩司. 菌種判別装置、菌種判別方法および菌種判別プログラム. 特願 2025-018598 (2025 年 2 月 6 日)

資料　医療 AI モデル開発とサービス公開における情報セキュリティチェックリスト

医療 AI モデル開発とサービス公開における情報セキュリティチェックリスト

対象： 医療 AI モデルの開発から公開・運用までの全フェーズ
目的： 安全な医療 AI サービスの構築と継続的な品質確保
注記： ☆印は LLM・AI エージェント活用時に特に注意が必要な項目

A. 開発環境・ツールのセキュリティ

A-1. AI エージェント・LLM ツールの選定と利用ポリシー

☐ 利用する LLM・AI エージェントのプライバシーポリシーを確認し、医療情報を含むプロンプトの入力禁止を徹底した ☆
☐ 組織としての利用ポリシーを策定し、開発チーム全員に周知した
☐ AI エージェントの出力が確率的であることを開発チーム全員が認識している ☆
☐ エンタープライズ向けプランまたはオンプレミス環境での利用を検討した

A-2. API キー・MCP 管理と権限付与

☐ API キーをソースコードに直接記述せず、秘密管理サービスを利用している（AI エージェント生成コードも含めて確認）☆
☐ MCP サーバの提供元が信頼できることを確認し、AI エージェントのアクセス範囲を最小限に制限した ☆
☐ AI エージェントによる自律的なコマンド実行に人的承認フローを設けた ☆

A-3. 開発・本番環境の分離

☐ 開発・ステージング・本番環境を明確に分離し、本番環境への自動デプロイに人的承認を組み込んだ
☐ 開発環境に実際の患者データ・個人情報を使用していない
☐ AI エージェントが生成した設定ファイルに本番環境の認証情報が含まれていないことを確認した ☆

B. AI モデル開発におけるデータセキュリティ

B-1. 学習データの管理

☐ 学習データの匿名化・非識別化が適切に行われている
☐ 学習データの取得・保存・転送時に暗号化が施されている
☐ 学習データへのアクセス制御が実装され、アクセスログが記録されている
☐ AI エージェントを用いたデータ処理スクリプトに患者情報が混入していないことを確認した ☆

B-2. モデルの管理

☐ 学習済みモデルのバージョン管理が行われている
☐ モデルの再学習・更新時にデータセキュリティの確認を再実施する手順が定められている
☐ モデルの出力に個人情報が含まれないことを確認した
☐ LLM を用いた特徴量生成・前処理において意図しないデータ漏洩がないことを確認した ☆

C. コード・アプリケーションセキュリティ

C-1. AI エージェント生成コードのレビュー

☐ AI エージェントが生成したコードを必ず人的レビューし、そのまま本番利用しない ☆
☐ 依存ライブラリの脆弱性を確認した（AI エージェントが古いバージョンを採用する場合がある）☆
☐ AI エージェントが一つの指示で全セキュリティ対策を網羅するわけではないことを認識し、個別に確認した ☆

C-2. 基本的なセキュリティ実装

☐ XSS 対策・SQL インジェクション対策が実装されている
☐ セキュリティヘッダーが適切に設定されている（他サイトからのアクセス制御、クリックジャッキング対策等）
☐ 安全なユーザー認証・セッション管理・アクセス制御が実装されている
☐ ファイルアップロードの検証・制御が実装されている

D. インフラ・クラウド・サーバレス構成

D-1. サーバレス移行とセキュリティ

☐ サーバレスアーキテクチャへの移行作業において AI エージェントを活用し、移行手順の一貫性と安全性を確保した ☆
☐ IAM による最小権限の原則が適用されている（AWS IAM・Microsoft Entra ID・Google Cloud IAM 等）
☐ サーバレス環境におけるデータの保存・転送時の暗号化が標準的に設定されている
☐ イベント駆動型サービス（AWS Lambda 等）における認証・リクエスト署名が適切に設定されている

D-2. コンテナ・監視設定

☐ コンテナイメージに不要なパッケージが含まれておらず、イメージサイズが最小化されている（AI エージェント生成の Dockerfile も確認）☆
☐ リアルタイムモニタリングとログ収集が設定されている
☐ 異常検知とインシデント自動対応の仕組みが構成されている
☐ パッチ適用の管理とロールバック手順が定められている

E. 医療固有の要件

E-1. 法的・倫理的要件

☐ 個人情報保護法・医療情報システムの安全管理に関するガイドラインへの適合を確認した
☐ 医療機器該当性の確認を行い、必要に応じて薬機法上の手続きを実施した
☐ 倫理審査・院内承認プロセスを経ている
☐ 患者情報保護に関する要件を AI エージェントへの指示に明示的に組み込んだ ☆

E-2. ネットワーク・アクセス制御

☐ 電子カルテネットワークと医療 AI サービス間の接続制限を確認・遵守している
☐ サービス側からの電子カルテネットワークへのアクセスが遮断されていることを確認した
☐ 許可された通信経路（NTP 同期・MDM 認証等）以外の外部接続が発生していないことを確認した
☐ AI エージェントが医療固有のネットワーク制約を認識した上で設定を生成していることを確認した ☆

F. AI サービス公開時のセキュリティ

F-1. 公開前確認

☐ 脆弱性評価・ペネトレーションテストを実施した（AI エージェントによる継続的な脆弱性評価も活用）☆
☐ 入力値の検証とサニタイズが全入力フォームに実装されている
☐ レート制限・DoS 対策が実装されている
☐ プロンプトインジェクション攻撃への対策を講じた（LLM を利用したサービスの場合）☆

F-2. 公開後の管理

☐ 利用者認証とアクセスログの記録・監視体制が整っている
☐ インシデント発生時の対応計画と連絡体制が整備されている
☐ 利用規約・プライバシーポリシーが整備・公開されている
☐ AI サービスの出力に関する免責事項と人的確認の必要性が明示されている ☆

G. 運用・保守・継続的管理

G-1. 定期的な評価と更新

☐ セキュリティポリシーの定期的な見直しと更新を実施している
☐ 利用する AI エージェント・LLM ツールのアップデートと変更内容を定期的に確認している ☆

☐ 依存ライブラリ・コンテナイメージの定期的な脆弱性スキャンを実施している
☐ AI モデルの再学習・更新時にセキュリティ確認を再実施する手順が定められている

G-2. 人的管理体制

☐ AI エージェントが生成・更新した設定・コード・ドキュメントの最終確認を人が行う体制が整っている ☆
☐ セキュリティに関する教育・訓練を定期的に実施している（AI エージェントによる支援も活用）☆
☐ インシデント対応訓練を定期的に実施している
☐ 担当者不在時のバックアップ体制が整っている

H. LLM・AI エージェント利用に関する固有リスク

H-1. AI エージェント利用時の固有リスク管理

☐ プロンプトに機密情報・患者情報・認証情報を含めていない ☆
☐ AI エージェントの自律的な判断に過度に依存せず、重要な意思決定には人的判断を介在させている ☆
☐ AI エージェントが生成したセキュリティ設定が、医療固有の要件を満たしているかを専門家が確認した ☆
☐ AI エージェントへの権限付与は必要最小限にとどめ、定期的に見直している ☆

H-2. AI エージェント活用の継続的改善

☐ AI エージェント活用による標準化された開発手法への移行が組織的に推進されている
☐ AI エージェントを活用した人材育成・教育の仕組みが整備されている
☐ AI エージェントの活用状況と成果を定期的に評価し、組織の開発プロセスに反映している
☐ 医療・情報・セキュリティの専門家が連携し、AI エージェントへの指示と最終確認を分担する体制が整っている ☆

改訂履歴 Ver. 1.0　Claude Sonnet 4.6 初版作成