

講座名 医療 AI 開発学講座／医療 AI デジタルツイン開発学講座

代表者	特任准教授	河添 悦昌
構成員	特任教授	美代 賢吾
構成員	特任助教	篠原 恵美子
構成員	特任助教	平林 真衣
構成員	特任研究員	嶋本 公德

【 講座概要 】

- 少子高齢化と人口減少が進む社会の質を維持し、さらに発展させるためには、ICT（情報通信技術）とAI（人工知能）技術とを基軸として、データを活用した合理的、効率的な社会運営が必要と考えます。当講座は医学部附属病院 企画情報運営部を協力講座として、新しい医療サービスの基盤となる ICT/AI 技術の開発を目指しております。また、本分野に携わる若い研究者への支援拠点の形成と、学部や大学の枠を超えた共同研究を展開し医療の発展に貢献することを目指して参ります。本講座は2023年7月末で第1期を終了し、8月より第2期として医療 AI デジタルツイン開発学講座に名称を変更し活動を続けて参ります。

【 研究内容 】

1. 主な研究活動

統合型ヘルスケアシステムの構築

- 本邦では異なる医療機関を横断した医療情報を統合し活用することがほとんど行われておりません。その理由の一つとして、医療情報の標準規格が十分に普及していないことがあげられます。大病院に導入される大きなシェアを占める電子カルテベンダー間でも、相互に情報を交換可能なシステムを開発する動きはみられず、また、同一ベンダーの電子カルテであっても病院ごとに各種マスタ類のコードが異なるため、病院を超えた情報の統合を行うことが困難です。そこで、第3期 戦略的イノベーション創造プログラム（SIP）「統合型ヘルスケアシステムの構築（PD：永井 良三 自治医科大学学長）」では、次世代の医療情報交換の標準規格である HL7 FHIR に準拠した医療情報システムの普及と、これらを基盤とする医療デジタルツインの開発を目指しております。本寄附講座では、このプロジェクトにおけるテーマB5「医療機器・材料のトレーサビリティデータ収集・分析システムの構築を通じた医療機器開発・改良支援、医療資源最適化、病院経営最適化支援（代表：美代 賢吾 特任教授、NCGM 医療情報基盤センター長）」とテーマD1「医療機関・ベンダー・システムの垣根を超えた医療データ基盤構築による組織横断的な医療情報収集の実現（代表：河添 悦昌 特任准教授）」に取り組むことで、統合型ヘルスケアシステムの構築に貢献することを目指しております。

カルテテキストからの情報抽出技術の開発

- 電子カルテに記録されるテキストから病名や症状、所見などの情報を抽出し、構造化データとすることができれば、ビッグデータとして利用可能になると考えられます。これまで、文部科学省基盤研究（B）「遺伝性疾患のスクリーニングに向けた診療記録からの表現型の抽出と臨床応用評価」の支援を受けて、約350例の希少・難治性疾患からなる症例報告に、詳細

な注釈情報を付与したテキストコーパスを開発し、診療テキストから高い性能で情報を抽出する機械学習モデルの開発を進めてきました。また、テキストの著作権の許可を得て、研究者が利用できるようにコーパスを公開してきました。さらに、抽出した病名や症状、所見などを既存の用語集に対応付けることで統計解析可能なデータするための研究を、文部科学省基盤研究（B）「医療用語のエンティティリンキングに向けた実践的医療用語辞書の開発（代表：河添 悦昌 特任准教授）」の支援を受けて進めております。

リアルワールドテキスト処理の深化によるデータ駆動型探薬

- 膨大なカルテ記録に含まれる医薬品や有害事象の情報を自然言語処理によって抽出し、統計解析することで、新しい薬効を持つ医薬品を発見することが期待されます。本寄附講座では、科学技術振興機構 戦略的創造研究推進事業（CREST）「データ駆動・AI 駆動を中心としたデジタルトランスフォーメーションによる生命科学研究の革新「リアルワールドテキスト処理の深化によるデータ駆動型探薬（主たる共同研究者：河添 悦昌 特任准教授）」の支援を受けて、これを可能とするためのプラットフォームを開発しており、これをがん治療の分野に適用し新しい薬効を持つ医薬品の発見を進めております。

2. 教育活動

- 医学系研究科社会医学専攻の博士課程と公共健康医学専攻専門職修士課程（SPH）の大学院生の研究指導を担当するとともに、医学部4年次とSPHにおいて「AIと医療」の講義、社会人を対象とする医療RWD人材育成事業（東京大学）、においては、医療画像を対象とする機械学習と自然言語処理の講義を行ってきました。また、高度医療経営人材養成プログラム（東京大学）においては担当教員を務め、大学病院を始めとする公益性の高い病院を対象とする経営幹部の養成に貢献してきました。

【今後の展望】

- 統合型ヘルスケアシステムの構築においては、医療情報の標準規格の普及と各種情報システムベンダー製品への社会実装、ならびに医療デジタルツインの開発を進め、医療系の入門者に向けたデータサイエンス技術と人工知能技術の講義を通して、医療ビッグデータを利活用可能な人材の裾野を広げることを目指します。また、医療テキストからのより高度な情報抽出技術の開発を進め、これを基盤として新しい医学・医療知識の発見を目指します。

【令和5年度活動実績】

<論文・著書>

■ 和文論文

- 草竹大暉, 平林（宮部）真衣, 吉野孝. Twitter上で多数の訂正情報が発信される流言の特徴分析, 情報処理学会論文誌, Vol. 64, No. 1, pp. 24-32 (2023).

■ 英文論文

- Yoneda K, Seki T, Kawazoe Y, Ohe K, Takahashi N; Neonatal Research Network of Japan. Immediate postnatal prediction of death or bronchopulmonary dysplasia among very

preterm and very low birth weight infants based on gradient boosting decision trees algorithm: A nationwide database study in Japan. PLoS One. 2024 Mar 27;19(3):e0300817.

- Seki T, [Kawazoe Y](#), Ohe K. Graph Representation Learning-Based Fixed-Length Clinical Feature Vector Generation from Heterogeneous Medical Records. Stud Health Technol Inform. 2024 Jan 25;310:715–719. doi: 10.3233/SHTI231058.
- [Shibata D](#), [Shinohara E](#), [Shimamoto K](#), [Kawazoe Y](#). Towards Structuring Clinical Texts: Joint Entity and Relation Extraction from Japanese Case Report Corpus. Stud Health Technol Inform. 2024 Jan 25;310:559–563. doi: 10.3233/SHTI231027.
- Seki T, [Kawazoe Y](#), Ohe K. Clinical Feature Vector Generation using Unsupervised Graph Representation Learning from Heterogeneous Medical Records. AMIA Annu Symp Proc. 2024 Jan 11;2023:618–623.
- [Hirabayashi M](#), [Shibata D](#), [Shinohara E](#), [Kawazoe Y](#). Influence of Tweets Indicating False Rumors on COVID-19 Vaccination: Case Study. JMIR Form Res. 2023 Sep 5;7:e45867. doi: 10.2196/45867.

<学会・講演会発表>

■ 国内学会

- [河添 悦昌](#), [永島 里美](#), [大江 和彦](#). アレルギー情報の標準化を目指すJ-FAGYアレルゲン用語集. 第43回医療情報学連合大会.
- [榎原 芽美](#), [柴田 大作](#), [篠原 恵美子](#), [河添 悦昌](#), [大江 和彦](#). UMLSからの同義語を追加した形態素解析辞書を使用したPhenotypingの性能評価. 第27回日本医療情報学春季学術大会, 2023年7月1日.
- [関 倫久](#), [河添 悦昌](#), [大江 和彦](#). 国内医療機関ウェブサイトにおけるアクセス解析サービスの利用状況の調査研究. 第27回日本医療情報学春季学術大会, 2023年6月30日.
- [西村涼太](#), [平林 \(宮部\) 真衣](#), [吉野孝](#). 流言に対する耐性獲得のための定期的な注意喚起手法の提案, 情報処理学会グループウェアとネットワークサービス研究発表会, 2023年1月23日.
- [下津拓未](#), [平林 \(宮部\) 真衣](#), [吉野孝](#). マイクロブログにおけるアンチシズルワードの分析: 「おいしくない」ことを意図する具体的表現の分類, 情報処理学会グループウェアとネットワークサービス研究発表会, 2023年1月23日.

■ 国際学会

- Tomohisa Seki, Yoshimasa Kawazoe, Kazuhiko Ohe. Clinical Feature Vector Generation using Unsupervised Graph Representation Learning from Heterogeneous Medical Records. AMIA 2023 Annual Symposium in New Orleans (15 Nov 2023). (再掲)
- Daisaku Shibata, Emiko Shinohara, Kiminori Shimamoto and Yoshimasa Kawazoe. Towards structuring clinical texts: Joint entity and relation extraction from Japanese case report corpus. MedInfo 2023 in Sydney (12 July 2023). (再掲)
- Tomohisa Seki, Yoshimasa Kawazoe and Kazuhiko Ohe. Graph representation learning-based

fixed-length clinical feature vector generation from heterogeneous medical records.
MedInfo 2023 in Sydney (11 July 2023). (再掲)

<講演会・研究集会>

- 河添 悦昌. 医療機関・ベンダー・システムの垣根を超えた医療データ基盤構築による組織横断的な医療情報収集の実現. SIP 第3期「統合型ヘルスケアシステムの構築」2023年度公開シンポジウム. 2024年3月22日(金)
- 河添 悦昌. 内閣府 SIP: 統合型ヘルスケアシステムの構築に向けた組織横断的な医療情報収集の実現. 日本医療情報学会 NeXEHRs 研究会. 2023年12月13日(水)
- 河添 悦昌. アレルギーのコード JFAGY の策定. 日本医療情報学会 NeXEHRs 研究会. 「医療 DX における標準化: 次世代医療情報標準 FHIR の新しい展開」- 3文書6情報から看護・栄養と歯科口腔ケア分野へ-. 2023年11月28日(火)
- 河添 悦昌. 統計的言語処理と大規模言語モデルならびに医学研究への応用. 精神神経科セミナー. 2023年10月5日.
- 河添 悦昌. ChatGPT が医療に与えるインパクト. 日本医療情報学会 NeXEHRs 研究会 緊急シンポジウム ChatGPT は医療情報研究と医療を変えるか? (資料)
- 河添 悦昌. 医療デジタルツインを加速する自然言語処理. 第31回日本医学会総会 U40 委員会企画 AI は医師を置き換えるか? ~医療 AI の未来予想図~, 2023年4月22日. (招待講演)

■ 競争的研究資金

1. 内閣府, 第3期戦略的イノベーション創造プログラム (SIP) 統合型ヘルスケアシステムの構築, テーマ D1 「医療機関・ベンダー・システムの垣根を超えた医療データ基盤構築による組織横断的な医療情報収集の実現」(代表: 河添 悦昌 | R5-R9)
2. 内閣府, 第3期戦略的イノベーション創造プログラム (SIP) 統合型ヘルスケアシステムの構築, テーマ D2 「統合型の医学概念・知識連結データベースの構築及び医療文書の自動分析基盤の整備」(主たる共同研究者: 河添 悦昌, 代表: 奈良先端科学技術大学院大学 荒牧 英治 教授 | R5-R9)
3. 文部科学省, 科学研究費助成事業 基盤研究 (B) 「医療用語のエンティティリンキングに向けた実践的医療用語辞書の開発」(代表: 河添 悦昌, 分担: 篠原 恵美子 | R5-R7)
4. 厚生労働省, 厚生労働行政推進調査事業 「次世代の医療情報の標準規格拡充等に資する研究」(代表: 河添 悦昌, 分担: 篠原 恵美子 | R5-R6)
5. 厚生労働省, 厚生労働科学特別研究事業 「海外における標準化を有した高品質医療リアルワールドデータ基盤整備の調査研究」(代表: 山下 貴範, 分担: 河添 悦昌 | R5-R5)
6. 科学技術振興機構, 戦略的創造研究推進事業 (CREST) データ駆動・AI 駆動を中心としたデジタルトランスフォーメーションによる生命科学研究の革新 「リアルワールドテキスト処理の深化によるデータ駆動型探査」(主たる共同研究者: 河添 悦昌, 代表: 奈良先端科学技術大学院大学 荒牧 英治 教授 | R4-R9)
7. 文部科学省, 科学研究費助成事業 基盤研究 (C) 「日本の食の魅力を伝達する異言語間シズ

ル情報提供手法に関する研究」(代表: 平林真衣 | R3-R5)

8. 文部科学省, 科学研究費助成事業 挑戦的研究(萌芽)「COVID-19 関連制限と外国人居住者の健康福祉: ビッグデータ解析による自然実験」(分担: 河添悦昌, 篠原恵美子, 平林真衣, 代表: 東京大学 神馬征峰 教授 | R3-R5)

■ 社会活動

1. 河添悦昌. ICD10対応標準病名マスター検討委員会作業班 委員
2. 河添悦昌. 厚生労働省 患者情報登録WG/造血器腫瘍サブWG 委員
3. 河添悦昌. 日本医療情報学会 次世代健康医療記録システム共通プラットフォーム課題研究会 HL7-FHIR 日本実装検討WG 委員