

Progress report 2024

Artificial Intelligence and Digital Twin in Healthcare

Project Professor: Kengo Miyo, Ph.D.

Project Associate Professor: Yoshimasa Kawazoe, M.D., Ph.D.

Project Assistant Professor: Emiko Shinohara, Ph.D.

Kiminori Shimamoto, M.S.

Course Outline

- To maintain and further develop social quality in the face of declining birthrates and aging population, we believe the necessitation of the rational and efficient societal management of using data, with information and communication technology (ICT) and artificial intelligence (AI) as the pillars. This course aims to develop ICT/AI technologies to underpin new medical services, in collaboration with the Planning and Information Operations Department of the School of Medicine Hospital. Additionally, we seek to establish a support hub for young researchers in this field and promote collaborative research across faculties and universities to advance healthcare. The inaugural phase of this course concluded at the end of July 2023, and continued its activities since August under the name “Medical AI Digital Twin Development Course” for the second phase.
- Extracting information such as disease names, symptoms, and findings from the text recorded in electronic medical records and structuring it into data could enable its use as big data. Supported by the Japan Society for the Promotion of Science (JSPS) Grant-in-Aid for Scientific Research (B) program “Extraction and Clinical Assessment of Phenotypic Information from Clinical Records for the Screening of Genetic Disorders”, we have developed a text corpus comprising approximately 350 cases of rare and refractory diseases, annotated with detailed information. We have also made significant progress in developing machine learning models capable of extracting information from clinical texts with high accuracy. Furthermore, with copyright permission, we have made the corpus publicly available for researchers. Moreover, supported by the Japan Society for the Promotion of Science (JSPS) Grant-in-Aid for Scientific Research (B) program, under the project titled “Development of a Practical Medical Terminology Dictionary for Entity Linking of Medical Terms” (led by Representative Yoshimasa Kawazoe, Project Associate Professor), we aim to associate extracted disease names, symptoms, findings, etc., with existing terminologies to generate data suitable for statistical analysis.

Research Activities

Building an integrated healthcare system

- If information such as disease names, symptoms, and diagnoses can be extracted from the text of electronic medical records and transformed into structured data, then such data could be statistically analyzed. We have developed a text corpus of case reports consisting of approximately 350 cases of rare and intractable diseases with detailed annotation information, and have been developing high-performance machine learning models to extract information from medical texts. We have also made our corpus available to researchers after obtaining copyright permission for reproduction of the texts.

Development of information extraction techniques from medical records text

Data-driven drug discovery through the deepening of real-world text processing

- Utilizing natural language processing to extract information about pharmaceuticals and adverse events from extensive medical records hold promise for discovering new drugs with novel therapeutic effects through statistical analysis. Supported by the Japan Science and Technology Agency’s Core Research for Evolutional Science

and Technology “Innovation in Life Sciences Research through Data-Driven and AI-Driven Digital Transformation: Data-Driven Drug Discovery through the Deepening of Real-World Text Processing” (led by Principal Investigator Yoshimasa Kawazoe, Project Associate Professor), our endowed lecture series is developing a platform to facilitate this process. We are applying this platform in cancer treatment to advance the discovery of new pharmaceuticals with novel therapeutic effects.

Educational Activities

- We oversee research supervision research for doctoral students in the Department of Social Medicine within the Graduate School of Medicine, as well as for the Master’s Program in Public Health (SPH) specializing in professional training. Additionally, we deliver lectures on “AI and Healthcare” for fourth-year medical and SPH students. Furthermore, we provide lectures on machine learning and natural language processing focusing on medical imaging for the Healthcare Real-World Data (RWD) Talent Development Program aimed at working professionals (sponsored by the University of Tokyo). also serve as a faculty member in the Advanced Medical Management Personnel Development Program at the University of Tokyo, contributing to the training of management executives, particularly those from nonprofit hospitals such as university hospitals.

Future Prospects

- In advancing an integrated healthcare system, we aim to promote the dissemination of standard protocols for medical information, integrate various information system vendor products into society, and advance the development of medical digital twins. Through introductory lectures on data science and artificial intelligence in the medical field, we seek to broaden the base of professionals capable of effectively utilizing healthcare big data. Additionally, we are progressing in developing more sophisticated

information extraction techniques from medical texts to serve as a foundation for discovering new medical and healthcare knowledge.

Journal Publication

1. [Kawazoe Y](#), Tsuchiya M, [Shimamoto K](#), Seki T, [Shinohara E](#), Yada S, Wakamiya S, Imai S, Aramaki E, Hori S. Natural language processing of electronic medical records identifies cardioprotective agents for anthracycline induced cardiotoxicity. *Sci Rep*. 2025 Feb 24;15(1):6678. doi: 10.1038/s41598-025-91187-6.
2. Seki T, [Kawazoe Y](#), Ito H, Akagi Y, Takiguchi T, Ohe K. Assessing the performance of zero-shot visual question answering in multimodal large language models for 12-lead ECG image interpretation. *Front Cardiovasc Med*. 2025 Feb 6;12:1458289. doi: 10.3389/fcvm.2025.1458289.
3. Seki T, [Kawazoe Y](#), Takiguchi T, Akagi Y, Ito H, Kubota K, Miyake K, Okada M, Ohe K. Sex Differences in Post-Noncardiac Surgery Risks Assessed Using the Revised Cardiac Risk Index - A Nationwide Retrospective Cohort Study. *Circ J*. 2025 Feb 15. doi: 10.1253/circj.CJ-24-0846.
4. [Kawazoe Y](#), Nagashima S, Yokota S, Ohe K. Development of a code system for allergens and its integration into the HL7 FHIR AllergyIntolerance resource. *Int J Med Inform*. 2024 Nov 30;195:105739. doi: 10.1016/j.ijmedinf.2024.105739.
5. [Kawazoe Y](#), [Shimamoto K](#), Seki T, Tsuchiya M, [Shinohara E](#), Yada S, Wakamiya S, Imai S, Hori S, Aramaki E. Post-marketing surveillance of anti-cancer drugs using natural language processing of electronic medical records. *NPJ Digit Med*. 2024 Nov 9;7(1):315. doi: 10.1038/s41746-024-01323-1.
6. Seki T, Takiguchi T, Akagi Y, Ito H, Kubota K, Miyake K, Okada M, [Kawazoe Y](#). Iterative random forest-based identification of a novel population with high risk of complications post non-cardiac surgery. *Sci Rep*. 2024 Nov 5;14(1):26741. doi: 10.1038/s41598-024-78482-4.
7. [Emiko Shinohara](#), [Kiminori Shimamoto](#), [Yoshimasa Kawazoe](#). Uncovering condition information loss in medical text extraction: The challenge of non-contiguous spans. *Next Research*.

2024 Dec;1(2).

8. Yada S, Nishiyama T, Wakamiya S, Kawazoe Y, Imai S, Hori S, Aramaki E. Utility analysis and demonstration of real-world clinical texts: A case study on Japanese cancer-related EHRs. *PLoS One*. 2024 Sep 11;19(9):e0310432. doi: 10.1371/journal.pone.0310432.
9. Seki T, Kawazoe Y, Ohe K. A Comparative Study of Access Analysis Service Utilization on Japanese Medical Institutions' Websites with GDPR-Compliant Cases. *Stud Health Technol Inform*. 2024 Aug 22;316:1238-1242. doi: 10.3233/SHTI240635.

Research fundings

- Cabinet Office, Third Phase of the Cross-ministerial Strategic Innovation Promotion Program (SIP): "Development of an Integrated Healthcare System," Theme D1: "Realization of Cross-Organizational Medical Information Collection Through the Construction of a Medical Data Infrastructure that Overcomes Barriers Between Medical Institutions, Vendors, and Systems" (Representative: Yoshimasa Kawazoe | FY2023–FY2027).
- Cabinet Office, Third Phase of the Cross-ministerial Strategic Innovation Promotion Program (SIP): "Development of an Integrated Healthcare System," Theme D2: "Construction of an Integrated Medical Concept and Knowledge-Linked Database and Development of an Automated Analysis Platform for Medical Documents" (Main Co-Researcher: Yoshimasa Kawazoe; Representative: Professor Eiji Aramaki, Nara Institute of Science and Technology | FY2023–FY2027).
- Ministry of Education, Culture, Sports, Science and Technology (MEXT), Grants-in-Aid for Scientific Research (KAKENHI), Basic Research (B): "Development of a Practical Medical Terminology Dictionary for Entity Linking in Medical Terms" (Representative: Yoshimasa Kawazoe; Collaborator: Emiko Shinohara | FY2023–FY2025).
- Ministry of Health, Labour and Welfare, Research Project for the Promotion of Health and Welfare Administration: "Research Contributing to the Expansion of Next-Generation Standards for Medical Information" (Representative: Yoshimasa Kawazoe; Collaborator: Emiko Shinohara | FY2023–FY2024).
- Japan Science and Technology Agency (JST), Core Research for Evolutional Science and Technology (CREST): "Revolutionizing Life Science Research Through Digital Transformation Centered on Data-Driven and AI-Driven Approaches—Advancing Real-World Text Processing for Data-Driven Drug Discovery" (Main Co-Researcher: Yoshimasa Kawazoe; Representative: Professor Eiji Aramaki, Nara Institute of Science and Technology | FY2022–FY2027).

Social activities

- We participated in the HL7-FHIR (Fast Healthcare Interoperability Resources) Implementation Study Working Group, which was established under the Next Generation Healthcare Recording System Common Platform Study Group of the Japan Association for Medical Information, and we proceeded with the implementation of HL7-FHIR in Japan.
- As a part of the Working Group for Core Center for Cancer Genome Medicine of the Ministry of Health, Labor and Welfare, we were in charge of infrastructural development for implementing genomic medicine for cancer at the University of Tokyo Hospital.
- As a part of the Working Group for the ICD10-compliant Standard Disease Name Master Review Committee, we developed the standard disease name master.