



NARA INSTITUTE of SCIENCE and TECHNOLOGY Social Computing Laboratory
奈良先端科学技術大学院大学 ソーシャル・コンピューティング研究室

YEARBOOK 2020

奈良先端科学技術大学院大学
ソーシャル・コンピューティング研究室

2020 第5年報

(2020/04/01-2021/03/31)



Introduction

ソーシャル・コンピューティング研究室の 5 年目の年報です。本年度から研究室が情報科学領域に正式設置され、5 月末に約 5 年間過ごした学際融合領域研究棟 4 階から情報科学領域棟 6 階に引越しました。研究室設置にあたり、教員の職位もあがり、荒牧特任准教授は教授に、若宮特任助教は准教授に、矢田博士研究員は助教となりました。引越しでは、若宮准教授が、子細に什器を選定し、クールかつ居心地のよい研究環境が実現できました。コロナ禍でオンラインでの活動が多い中での引越しは大変な作業でしたが、引越作業を終えて、晴れやかな新年度を迎えることができました。



研究室の構成は、新たに 2 名の博士後期課程学生、8 名の博士前期課程学生と 1 名の特別研究学生、さらに、研究員の Kong 先生、博士研究員の工藤先生が加わり、研究スタッフ 6 名、博士後期課程学生 7 名、博士前期課程学生 11 名、特別研究学生 1 名の中規模の構成となりました。

コロナ禍でオンラインが中心となり、これまでと異なる環境にスタッフも学生も苦労しましたが、その中でもコンスタントに論文出版や国際会議での発表ができ、学生の受賞もありました。何よりも最大の成果として、本研究室の 1 期生である磯君、柴田君が研究室初の博士号を取得し

ました。それぞれ、米国 IT 企業と大学病院という好対照な新天地を選び、研究を継続しています。今後のより一層の活躍を研究室一同祈念しています。

プロジェクトに関しては、AIP-PRISM 創薬と SIP-AI ホスピタルといった大きなプロジェクトを主軸に、日独仏共同研究 KEEPHA プロジェクトが開始しました。ドイツ人工知能研究センター（DFKI）とフランス国立科学研究センター（CNRS）との国際共同研究であり、邦人学生の海外滞在研究や留学生の活躍の場も広がると期待しています。

企業共同研究に関しても、富士ゼロックス、Yahoo! Japan 研究所、アマダ AI イノベーション研究所、株式会社エス・エム・エス、キャンサー・ソリューションズ株式会社（CANSOL）といった企業と実施しました。また、新たに NEC 研究所とエスエス製薬株式会社への学術指導を調整中です。

これらの多くのプロジェクトの結果を過去 5 年間回し続けた結果、スタッフが複数のプロジェクトを掛け持ちしつつ自主的に骨子となりながら推進し、学生がリサーチアシスタントとしてピンポイントで参加し、論文化を狙うというスタイルが確立してきたように思います。このスタイルは、プロの研究スタッフの安定性と学生の瞬発力の高さというよいところを合わせることができる新しい運営の方法論だと思っています。

もう一つの大きな動向が自然言語処理の隆盛です。10 年前には想像もしていませんでしたが、AI を支える基盤技術の 1 つであり、専門性が高い技術として、大きな注目を浴びています。この結果、自然言語処理はかつてない規模で進展しつつあり、フォローするのにもチーム力が必要になっています。今後、どのように自然言語処理を研究するのかという大きな選択を迫られていると考えています。

Research & Project

現在、約 10 の研究プロジェクトが動いており、1～3 名のスタッフが中心となって推進しています。国プロとして、AIP-PRISM 創薬、SIP-AI ホスピタル、日独仏共同研究 KEEPHA に加え、企業との共同研究としては、Yahoo! Japan 研究所（検索クエリ、SNS の応用）、アマダ AI イノベーション研究所（板金レポートの NLP）、株式会社エス・エム・エス（介護記録の NLP）、NEC 研究所（看護記録の NLP）、ASrid（希少疾患患者の語りの NLP）、CANSOL（がん患者の語りの NLP）など、NPO との共同研究も含む多くの研究を推進しています。これらは材料となるテキストは異なりますが、基本的には言語処理をベースにした研究であり、相互にシナジー効果を持ちながら、同時並行で推進できました。また、医療言語処理研究だけでなく、異分野融合

国際ワークショップ「パンデミック時代のソーシャル・コンピューティング」の主催などを通して、ソーシャル・コンピューティングに関する国際的な研究交流も積極的に推進しています。

Education

全体ミーティングである論文紹介 MTG に加え、

- 医療系の研究である MED-MTG（荒牧先生，矢田先生担当）
- SNS 系の研究である DATA-MTG（若宮先生担当）
- 留学生が中心となっている INTERNATIONAL-MTG（Kong 先生担当）

という3つの個別 MTG を行っており、学生はいずれか1つ以上に参加しています。なお、教員は、個別ミーティングにおいても全教員が参加し、情報の共有を図っています。

上記に加え、10名を超すスタッフ一同が集まるスタッフ MTG、3名の教員+秘書が集まるコアスタッフなどにより、研究室運営、学生へのサポートが検討されています。

2020 年度修士論文：

- Mutinda Faith Wavinya 「Detecting Redundancy in Electronic Medical Records using BERT」

2020 年度博士論文：

- Hayate Iso 「Teaching Machines to Write from Data」
- Daisaku Shibata 「A Study on Clinical Natural Language Processing Covering Electronic Medical Records and Patient Narratives」

Social Activity

アウトリーチに関しては、年間10回を超える招待講演はそのままに、教員がそれぞれの学会で若手の中心的存在として活動しています。ASrid（希少疾患患者の語りの NLP）、CANSOL（がん患者の語りの NLP）など、NPO といった企業でないパートナーとの共同研究も行っているのが特徴です。また、大阪国際がんセンター、関西医科大学や京都大学乳腺外科との共同研究など多くの医療施設との研究も継続しています。

最後に、今年度新たに始めた試みとして外部の参加者も含めたソーシャル・コンピューティング・ジャーナルクラブがあります。コロナ禍で研究会の開催が困難な状態ですが、オンラインならば逆に気軽に遠隔地を繋ぐこともできます。現在、毎月一度、東京大学医学部附属病院の河添

研究室や筑波大学などいくつかの研究室を接続し、カジュアルな話題提供ベースの討論会を開催しています。コロナ禍が収束した後も継続したいと考えています。

Member

- リサーチ・スタッフ
 - 荒牧英治（教授）
 - 若宮翔子（准教授）
 - 矢田竣太郎（助教）（2020/11-）
 - Kongmeng Liew（研究員）（2020/10-）
 - 工藤紀子（博士研究員）（2020/04-）
 - 眞鍋雅恵（研究員）
 - Dawar Khan（博士研究員）（2020/06-08）
- リサーチサポート・スタッフ
 - 金子雅美（技術補佐員） 研究室事務
- データ・スタッフ（非常勤）
 - 中江睦美（技術補佐員）（-2020/07, 2021/03-）
 - 川上庶子（技術補佐員）（-2020/07, 2021/03-）
 - 友廣公子（研究員）
- 学生
 - 磯 颯（D3）（2021/03 博士課程修了）
 - 柴田 大作（D3）（2021/03 博士課程修了）
 - 村山 太一（D2）
 - Nigo Sumaila（D1）
 - 勘場 大（D1）
 - Lean Franzl Lim Yao（D1・2020/10-）（コロナのため 2021/01 来日）
 - Faith Wavinya Mutinda（D1・2020/10-）（2020/09 修士課程修了）
 - 氏家 翔吾（M2・2020/10-）
 - Gamar Ivan Azuaje（M2・2020/10-）
 - 高 志偉（M2・2020/10-）
 - 安西 崇（M1）

- 伊藤 英里 (M1)
- 古賀 貴士 (M1) (2020/12-)
- Danay Tasew Tazebe (M1)
- 久田 祥平 (M1)
- 廣田 一輝 (M1)
- Vipul Mishra (M1)
- 前田 拓哉 (M1・2020/10-)
- 特別研究学生
 - 中村 優太 (東京大学医学部附属病院) (2020/04-09 研究室に滞在)

Publication

業績概要

- 原著論文 7 本
 - 英文原著 4 本
 - 和文原著 3 本
- 国際会議 (査読有)
 - 4 本

書籍

1. 荒牧英治, 若宮翔子, 矢田竣太郎: 第 2 章患者発信のカルテに見えない患者の声をリアルワールドエビデンスへ利活用, 日経 BP 医療・健康データビジネス/テクノロジー総覧, 2020. (2020/12/11)

論文

1. 柿本大輔, 平林 (宮部) 真衣, 荒牧英治, 吉野孝: RumorFinder: 情報真偽確認促進システムの一般公開とその利用評価, 情報処理学会論文誌, 62(1), 193-203, 2021 (2021/1/15)
2. Shogo Ujiie, Shuntaro Yada, Shoko Wakamiya, and Eiji Aramaki: Identification of Adverse Drug Event-Related Japanese Articles: Natural Language Processing Analysis, JMIR Medical Informatics, 8(11), e22661, 2020. (2020/11/27)

3. Shohei Hisada, Taichi Murayama, Kota Tsubouchi, Sumio Fujita, Shuntaro Yada, Shoko Wakamiya, and Eiji Aramaki: Surveillance of early stage COVID-19 clusters using search query logs and mobile device-based location information, Scientific Reports, 10, 18680 (2020). (2020/10/29)
4. 小沢彩歌, 平和也, 村山太一, 藤田澄男, 伊藤美樹子, 荒牧英治: WEB 上 Q&A サービスへの投稿ログデータにおける認知症者の介護で家族が抱える困難と悩み, ヘルスコミュニケーション雑誌, 11(2), 2020. (2020/10/2)
5. 荒牧英治, 若宮翔子, 河添悦昌: カルテ情報の自動構造化システムと疾患数理モデルの逐次的構築, 及び, 自動構造化機能を有した入力機構の開発, 医療情報学 (特集 厚生労働科学研究成果報告書 第3回), 40(1), 12-13, 2020. (2020/9/4)
6. Panote Siriara, Yuanyuan Wang, Yihong Zhang, Shoko Wakamiya, Péter Jeszenszky, Yukiko Kawai, Adam Jatowt: Beyond the Shortest Route: A Survey on Quality-Aware Route Navigation for Pedestrians, IEEE Access, vol. 8, pp. 135569-135590, 2020. (2020/7/27)
7. Taichi Murayama, Nobuyuki Shimizu, Sumio Fujita, Shoko Wakamiya, Eiji Aramaki: Robust two-stage influenza prediction model considering regular and irregular trends, PLOS ONE, 15(5):e0233126, 2020. (2020/5/21)

国際会議

1. Zhiwei Gao, Shuntaro Yada, Shoko Wakamiya, and Eiji Aramaki: Offensive Language Detection on Video Live Streaming Chat, In Proceedings of the 28th International Conference on Computational Linguistics (COLING 2020), pp. 1936–1940, 2020. (2020/12/8, Online)
2. Kongmeng Liew, Yukiko Uchida, Nao Maeura, and Eiji Aramaki: Classification of Nostalgic Music Through LDA Topic Modeling and Sentiment Analysis of YouTube Comments in Japanese Songs, In Proceedings of the First Workshop on NLP for Music and Audio (NLP4MusA2020), 2020. (2020/10/16, Online)
3. Taichi Murayama, Shoko Wakamiya, and Eiji Aramaki: Fake News Detection using Temporal Features Extracted via Point Process, In Proceedings of the International Workshop on Cyber Social Threats (CySoc2020), 2020. (2020/6/5, Online)
4. Shuntaro Yada, Ayami Joh, Ribeka Tanaka, Fei Cheng, Eiji Aramaki and Sadao Kurohashi: Towards a Versatile Medical-Annotation Guideline Feasible Without Heavy Medical

Knowledge: Starting From Critical Lung Diseases, In Proceedings of the 12th International Conference on Language Resources and Evaluation (LREC 2020) (Poster), 4565-4572, 2020.

国内会議

1. Gamar Azuaje, Kongmeng Liew, Shuntaro Yada, Shoko Wakamiya, Eiji Aramaki, Dawar Khan: Birdscribe: A Semantic Writing Assistant Employing Text-based Image Generation and Modification, NLP 2021 CAI+CAI first workshop, 2021 (2021/3/19, Online)
2. 村山太一, 若宮翔子, 荒牧英治: フェイクニュース検出データセットにおける通時的バイアス, 言語処理学会第 27 回年次大会 (NLP2021), 2021 (2021/3/17, オンライン)
3. 中村優太, 花岡昇平, 野村行弘, 林直人, 阿部修, 矢田竣太郎, 若宮翔子, 荒牧英治: 診療記録で事前学習した言語モデルからの学習データ中の人名漏洩リスクの推定, 言語処理学会第 27 回年次大会 (NLP2021), 2021 (2021/3/17, オンライン)
4. 氏家翔吾, 磯颯, 荒牧英治: 文脈化埋め込み表現を用いた対照学習による病名正規化, 言語処理学会第 27 回年次大会 (NLP2021), 2021 (2021/3/16, オンライン)
5. 村山太一, 若宮翔子, 荒牧英治, 小林亮太: Twitter におけるフェイクニュース拡散モデルの提案, 第 13 回データ工学と情報マネジメントに関するフォーラム (DEIM2021), 2021 (2021/3/3, オンライン) 学生プレゼンテーション賞
6. 村山太一, 若宮翔子, 荒牧英治: 検索クエリを活用した複数国におけるインフルエンザ流行予測, 第 13 回データ工学と情報マネジメントに関するフォーラム (DEIM2021), 2021 (2021/3/3, オンライン)
7. 安西崇, 村山太一, 矢田竣太郎, 若宮翔子, 荒牧英治: ならまちハッピーマップ: クラウドソーシングと深層学習による街の雰囲気に基づく経路探索手法の提案, 第 13 回データ工学と情報マネジメントに関するフォーラム (DEIM2021), 2021 (2021/3/3, オンライン) 学生プレゼンテーション賞
8. Zhiwei Gao, Kongmeng Liew, Taichi Murayama, Shuntaro Yada, Shoko Wakamiya, Eiji Aramaki, Nobuyuki Shimizu, Sumio Fujita: Public Concern About COVID-19 Through Search Queries, 第 13 回データ工学と情報マネジメントに関するフォーラム (DEIM2021), 2021 (2021/3/2, オンライン)
9. 廣田一輝, 村山太一, 矢田竣太郎, KongmengLiew, 若宮翔子, 荒牧英治: 長期 Twitter ユーザの投稿内容の経時的分析, 第 13 回データ工学と情報マネジメントに関するフォー

- ーラム (DEIM2021), 2021 (2021/3/1, オンライン)
10. 村山太一, 若宮翔子, 荒牧英治: オルタナティブメディアに着目した日本のツイッターにおけるニュース拡散の分析, 第5回計算社会科学ワークショップ (CSSJ2021), 2021. (2021/2/28, オンライン)
 11. 勘場大, 眞鍋雅恵, 若宮翔子, 荒牧英治: 自然言語処理を用いた WebQA サイトからのがん患者の医療ニーズ抽出, 第40回医療情報学連合大会, 2020. (2020/11/21, 静岡, アクトシティ浜松)
 12. 中村優太, 花岡昇平, 野村行弘, 林直人, 阿部修, 矢田竣太郎, 若宮翔子, 荒牧英治: 診療記録で事前学習した汎用言語モデルからの個人情報流出リスクの定量的評価, 第40回医療情報学連合大会, 2020. (2020/11/21, 静岡, アクトシティ浜松) 研究奨励賞
 13. 氏家翔吾, 矢田竣太郎, 若宮翔子, 荒牧英治: 日本語病名正規化システムの開発, 第40回医療情報学連合大会, 2020. (2020/11/19, 静岡, アクトシティ浜松) 研究奨励賞
 14. 伊藤英里, 若宮翔子, 荒牧英治: 家族介護者の悩み抽出のための用語単位およびエピソード単位の分類手法の検討, 第12回日本ヘルスコミュニケーション学会学術集会, 2020. (2020/9/26, オンライン)
 15. 氏家翔吾, 矢田竣太郎, 若宮翔子, 荒牧英治: スパン表現を用いた病名の抽出・正規化の同時学習, NLP 若手の会 (YANS) 第15回シンポジウム, 2020. (2020/9/23, オンライン)
 16. 中村優太, 花岡昇平, 野村行弘, 林直人, 阿部修, 矢田竣太郎, 若宮翔子, 荒牧英治: 読影レポートを対象とした共参照解析データセット作成に関する検討, NLP 若手の会 (YANS) 第15回シンポジウム, 2020. (2020/9/23, オンライン)
 17. Zhiwei Gao, Shuntaro Yada, Shoko Wakamiya, Eiji Aramaki: A Preliminary Analysis of Offensive Language Transferability from Social Media to Video Live Streaming, The 34th Annual Conference of the Japanese Society for Artificial Intelligence, 1K3-ES-2-04, 2020. (2020/6/9, オンライン)
 18. 柴田大作, 河添悦昌, 嶋本公德, 篠原恵美子, 荒牧英治: 診療記録で事前学習した BERT による疼痛表現の抽出, 第24回日本医療情報学会春季学術大会, 2020. (2020/6/6, オンライン)

外部獲得資金（本年度予算 68,240 千円）

- 荒牧
 - AI ホスピタル 28,000 千円
 - AIP-PRISM 18,500 千円
 - 新学術（総括班） 4,000 千円
 - その他，分担（基盤 A，基盤 B など）多数
 - 若宮
 - 若手 1,800 千円
 - その他，分担（基盤 B，厚労科研など）
 - 矢田
 - 若手 1,200 千円
-
1. 2020, 医薬品安全性監視のための言語を越えた知識強化情報抽出（AIP ネットワークラボ「日独仏 AI 研究」）研究分担者：荒牧英治（研究代表者：松本裕治）（230 千円）
 2. 2020, 健康ソーシャルデータの高信頼化とナビゲーションに関する研究（科学研究費補助金 若手研究）研究代表者：若宮翔子（1,800 千円）
 3. 2020, デジタル環境における偶然の情報遭遇を促すユーザインタフェースの設計と検証（科学研究費補助金 若手研究）研究代表者：矢田竣太郎（1,200 千円）
 4. 2020, モノ作り技術及びその活用現場に関するオントロジー構築の研究（株式会社アマダ AI イノベーション研究所）（4,000 千円）
 5. 2020, カイポケテキストデータの分類に関する研究（株式会社エス・エム・エス）（2,700 千円）
 6. 2020, （進行）乳がん患者のナラティブな語り研究～言葉から見える患者の特性～（キャンサー・ソリューションズ株式会社）（1,000 千円）
 7. 2020, がん患者のナラティブな語りに関する研究～言葉から見える治療と就労の両立～（一般社団法人 CSR プロジェクト）（500 千円）
 8. 2020, ソーシャルメディアによる疾患予測に関する研究（エスエス製薬株式会社）（760 千円）
 9. 2020, 学際領域におけるビッグデータとクラウドソーシング活用の研究（ヤフー株式会社）（1,100 千円）

10. 2019, がん診療連携拠点病院等における医療提供体制の均てん化のための評価に既存資料を活用する (厚生労働科学研究費), 研究分担者: 荒牧英治 (研究代表者: 宮代勲) (300 千円)
11. 2019, がん診療連携拠点病院等における医療提供体制の均てん化のための評価に既存資料を活用する (厚生労働科学研究費), 研究分担者: 若宮翔子 (研究代表者: 宮代勲) (300 千円)
12. 2019, 災害時における状況を考慮したユーザセンリックな行動促進情報基盤 (科学研究費補助金 基盤研究 (B)), 研究分担者: 荒牧英治 (研究代表者: 灘本明代) (400 千円)
13. 2019, 災害時における状況を考慮したユーザセンリックな行動促進情報基盤 (科学研究費補助金 基盤研究 (B)), 研究分担者: 若宮翔子 (研究代表者: 灘本明代) (400 千円)
14. 2019, 常識的知識を活用した言語理解・推論に基づく議論マイニングの新展開 (科学研究費補助金 基盤研究 (A)), 研究分担者: 荒牧英治 (研究代表者: 岡崎直観) (2,400 千円)
15. 2018, マスギャザリング時や新興・再興感染症の発生に備えた感染症サーベイランスの強化とリスクアセスメントに関する研究 (厚生労働行政推進調査事業費), 研究分担者: 荒牧英治 (研究代表者: 島田智恵) (300 千円)
16. 2018, マスギャザリング時や新興・再興感染症の発生に備えた感染症サーベイランスの強化とリスクアセスメントに関する研究 (厚生労働行政推進調査事業費), 研究分担者: 若宮翔子 (研究代表者: 島田智恵) (300 千円)
17. 2018, 医療テキスト構造化のための言語・知識処理基盤の構築 (AIP-PRISM (創薬分野)), 研究分担者: 荒牧英治 (研究代表者: 黒橋禎夫) (18,500 千円)
18. 2018, レジリエントな手術チームのシステムダイナミクスの解明 (科学研究費補助金 基盤研究 (B)), 研究分担者: 荒牧英治 (研究代表者: 中島和江) (50 千円)
19. 2018, 【戦略的イノベーション創造プログラム AI (人工知能) ホスpitalによる高度診断・治療システム】音声入力とコミュニケーション支援のための医療辞書データベース作成 (ヒュービットジェノミクス株式会社) (28,000 千円)
20. 2016, 思春期からの主体価値の発展過程解明 (新学術領域研究) 研究分担者: 荒牧英治 (研究代表者: 笠井清登) (4,000 千円)