

NAIST ソーシャル・コンピューティング研究室 2018

The 3rd Year (2018/01/01-2019/03/31)

概要

ソーシャル・コンピューティング研究室も設置後 3 年を迎えました。博士課程から修士過程の学生までが揃い、また、若宮博士研究員が特任助教に昇任することで、教員 2 名体制となり、小規模ながら研究室の体制が整いました。2018 年度は、4 名の 2 期生 (Camille Marie Ruiz さん、竹内 瞭くん、村山 太一くん、山本 英弥くん) が修了し、このうち 1 名 (村山くん) が博士課程に進学し、残り 3 名が卒業しました。昨年修了した 1 期生 2 名 (磯 颯くん、柴田 大作くん) は両名とも博士課程に進学したので、彼・彼女らが初めての OB・OG となります。

このように人数としては小規模な研究室ですが、その活動はサイズ感に似合わぬ大きなものだと思います。権威のある論文誌への採択、AIP-PRISM や SIP-AI ホスピタルといった所謂<国プロ>への参加、この他、常に政府や病院、企業からの共同研究依頼や調査依頼があり、10 以上のプロジェクトが走っている状態になりました。これを教員 2 名、研究員 3 名と学生で動かすわけです。社会に求められている実感を嬉しいと感じる反面、いつもオーバーフロー状態で、今後、スタッフの増員とプロジェクトの取捨選択が課題となっています。

● 研究

研究面については、英語原著論文 (医学研究、情報研究) と国際会議 (主に情報研究) をターゲットに活動しています。

英語原著論文に関しては、博士課程学生の柴田くんの論文の著名な一般ジャーナルである PLoS One への採択をはじめ、若宮特任助教が Web 系医療情報のトップジャーナルである JMIR 誌 2 本の採択を得るなど、大きな貢献がありました。一方、国際会議においても、荒牧の MEDINFO への採録など一定の成果がありました。ただし、自然言語処理や Web のトップカンファレンスへの採録はなく、今後の投稿体制の確立が今後の課題です。

最後に、学術上のインパクトは小さいですが、「小児頻用医薬品に関する医薬品添付文書における記載状況の調査」など、社会的に重要な研究結果を出版できたことも重要だと考えています。多忙な医療現場ではなかなか実施できない体系立った言語資料調査を行うことも、日本で数少ない医療言語処理を扱う本研究室の重要なミッションだと思っています。

● 運営

予算を初めとした運営については、これまで中心であった文科省科研費や厚労省科研費などに加えて、本年度から AIP-PRISM や SIP-AI ホスピタルといった大きなプロジェクトに参画しました。研究内容は以下のとおりです。多くは本研究室でないと実施できない内容であり、数少ない医療言語処理の拠点としての責任を果たすべく、真摯に取り組んでいます。また、予算額も 2 億円規模となり、研究室の規模を考えると十分すぎる予算となりました。

1. 国立感染症研究所感染症疫学センター共同研究：オリンピック・パラリンピック時のソーシャルメディアにおける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)
2. 国立成育医療センター共同研究：薬剤添付文章における有害事象の表記ゆれ調査(厚生労働 ICT 基盤構築研究事業)
3. 東京大学医学部附属病院共同研究：電子カルテ入力支援システムの高度化 (厚生労働 ICT 基盤構築研究事業)
4. 慶応大学薬学部共同研究：調剤薬局薬歴報告から患者副作用表現の抽出 (AMED 事業)
5. 国立医薬基盤研究所，国立がんセンター，京都大学共同研究：肺疾患読影所見解析のための言語リソースの構築 (AIP-PRISM)
6. ヒュービットジェノミクス株式会社共同研究：大規模医療用語辞書の構築 (SIP-AI ホスピタル)
7. 大阪大学医学部附属病院共同研究：外科手術チームの会話分析
8. Yahoo! 研究所共同研究：ソーシャルメディアを用いた感染症予測

他多数

● 教育

教育面については，博士課程学生の磯くんの産総研インターン，柴田くんの企業インターンなどがあったこと，修士課程学生も各々の研究に取り組んでいたことから，週1回の少人数ミーティングを中心とし，全体でのミーティングは必要に応じて実施するという少人数ならではのフレキシブルな体制で運営を行いました。研究対象の分散の大きさを考慮すると，論文紹介など全体ベースの議論が効率的でないとの学生の意見を反映したのですが，持続可能な体制ではないため，来年度以降の再デザインが必要です。また，今年度は4名のインターンシップ学生を受け入れ，主に SNS 関連のテーマに取り組んでもらいました。短期間での成果を国際会議論文としてまとめて発表を行うなど，双方にとって実りのある機会となったと考えています。

2017 年度修士論文：

- 磯 颯 「Forecasting Word Model: Twitter-based Influenza Surveillance and Prediction」 (2017 年 9 月)
- 柴田 大作 「自然言語処理による認知症スクリーニング」 (2018 年 3 月)

2018 年度修士論文：

- Camille Marie Ruiz 「Melancholic Ramblings on Twitter: Detecting Depression from Social Media Activity」 (2018 年 9 月)
- 竹内 瞭 「京都市内のランドマークを対象とした Web 情報を用いた印象音の推定」 (2019 年 3 月)
- 村山 太一 「リソース・モデル・エリアを考慮したインフルエンザ流行予測」 (2019 年 3 月)
- 山本 英弥 「複合語の構成素情報を考慮した医療用語難易度の推定」 (2019 年 3 月)

● 社会活動

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，荒牧の言語処理学会理事，若宮の日本データベース学会理事など，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。

構成員 (2018 年度)

- リサーチ・スタッフ
 - 荒牧英治 (特任准教授)
 - 若宮翔子 (特任助教) テニューア・トラック制度 (-2018/11), AIP-PRISM による雇用 (2018/12-)
 - 伊藤薫 (博士研究員) AMED ゲノムによる雇用 (-2019/03)
 - 城綾実 (博士研究員) AIP-PRISM, SIP-AI ホスピタルによる雇用 (2019/02-)
 - 岩尾 友秀 (研究員) AMED 総合診療医プロジェクトによる雇用 (-2019/03)

- リサーチサポート・スタッフ
 - 本田ちひろ (研究技術員) (-2018/06)
 - 磯崎 聖子 (研究技術員) (2018/07-2019/03)
 - 大家 聡樹 (研究技術員) (2018/07-2019/03)

- 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, アルバイト)
 - 中江睦美 (技術補佐員)
 - 友広庶子 (技術補佐員)
 - 友廣公子 (技術補佐員) 「カルテ標準化」厚労科研・医療データ整形

- 学生
 - 磯颯 (D2)
 - 柴田大作 (D1)
 - Camille Marie Ruiz (M2) (2018/09 修了・就職)
 - 竹内瞭 (M2) (2019/03 修了・就職)
 - 田口勝弥 (M2) (2019/03 退学)
 - 村山太一 (M2) (2019/03 修了・博士課程進学 (当研究室))
 - 山本英弥 (M2) (2019/03 修了・就職)
 - Nigo Sumaila (M2) (2019/09 修了予定)

- インターン
 - Paolo Casani (University College London) (2017/10-2018/06)
 - Wannita Takerngsaksiri (Chulalongkorn University, Thailand) (2018/05-2018/08)
 - Mir'atul Khusna Mufida (State Polytechnic of Batam, Indonesia) (2018/11-2018/12)
 - Lean Franzl L. Yao (Ateneo de Manila University, Philippines) (2019/1)

業績概要

- 原著論文 13 本
 - 英文原著 7 本
 - 和文原著 6 本
- 国際会議（査読有）
 - 10 本
- 国内発表
 - 14 件
- 受賞
 - 1 件

外部獲得資金（本年度予算 66,300 千円）

- 荒牧
 - AI ホスピタル 15,000 千円（一部来年度繰越予定）
 - AIP-PRISM 15,000 千円
 - 厚労科研（荒牧班） 8,000 千円
 - 新学術（総括班） 4,000 千円
 - その他，分担（基盤 A，挑戦的萌芽など）多数
- 若宮
 - 若手 B 1,100 千円
 - その他，分担（基盤 A など）

業績

論文

1. Eiji Aramaki, Mai Miyabe, Chihiro Honda, Seiko Isozaki, Shoko Wakamiya, Akira Sato, Isao Miyashiro: KOTOBAKARI Study: Using Natural Language Processing of Patient Short Narratives to Detect Cancer Related Cognitive Impairment, MEDINFO 2019: eHealth-enabled Health, 2019 (to appear).
2. Tomoyuki Kabutoya, Hisahiko Sato, Eiji Aramaki, Kazuomi Kario, Ryoza Nagai: The clinical characteristics of heart failure from case reports presented at the Regional Meeting of the Japanese Society of Internal Medicine, Internal Medicine, 2583 (18), 2019.
3. Shoko Wakamiya, Mizuki Morita, Yoshinobu Kano, Tomoko Ohkuma and Eiji Aramaki. Tweet Classification Toward Twitter-Based Disease Surveillance: New Data, Methods, and Evaluations, J Med Internet Res 2019;21(2):e12783. (<http://dx.doi.org/10.2196/12783>)
4. Shoko Wakamiya, Shoji Matsune, Kimihiro Okubo and Eiji Aramaki. Causal Relationships Among Pollen Counts, Tweet Numbers, and Patient Numbers for Seasonal Allergic Rhinitis Surveillance: Retrospective Analysis, J Med Internet Res 2019;21(2):e10450. (<http://dx.doi.org/10.2196/10450>)
5. 荒牧英治，若宮翔子，岩尾友秀，川上庶子，中江睦美，松本妙子，友廣公子，栗山猛: 小児頻用医薬品に関する医薬品添付文書における記載状況の調査，医療情報学（原著-研究速報），Volume 38 (Number 6), 2018.
6. Daisaku Shibata, Kaoru Ito, Hiroyuki Nagai, Taro Okahisa, Ayae Kinoshita, Eiji Aramaki: Idea density in Japanese for the early detection of dementia based on narrative, PLoS One, Volume 13 (Number 12), pp. e0208418, 2018.

(<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0208418>)

7. Misa Usui, Eiji Aramaki, Tomohide Iwao, Shoko Wakamiya, Tohru Sakamoto, Mayumi Mochizuki: Extraction and Standardization of Patient Complaints from Electronic Medication Histories for Pharmacovigilance: Natural Language Processing Analysis in Japanese, *JMIR Public Health and Surveillance*, Volume 6 (Number 3), pp. e11021, 2018. (<https://medinform.jmir.org/2018/3/e11021/>)
8. Shoko Wakamiya, Yukiko Kawai, Eiji Aramaki: Twitter-Based Influenza Detection After Flu Peak via Tweets With Indirect Information: Text Mining Study, *JMIR Public Health and Surveillance*, Volume 4 (Number 39), pp. e65, 2018. (<https://publichealth.jmir.org/2018/3/e65/>)
9. 荒牧英治, 岩尾友秀, 若宮翔子, 伊藤薫, 矢野憲, 大江和彦: 症例検索システムの試行運用に基づいた利用状況に関する基礎的検討, *医療情報学*, Volume 38 (Number 4), pp. 245-258, 2018
10. 柿本大輔, 宮部真衣, 荒牧英治, 吉野孝: 流言拡散防止のための情報確認行動促進システムの構築, *ヒューマンインタフェース学会誌*, Volume 20 (Number 1), pp. 1-12, 2018.
11. 宮部真衣, 四方朱子, 久保圭, 荒牧英治: 音声認識を用いた言語能力自動測定システム“言秤”の構築, *自然言語処理「言語処理の応用システム」特集号*, Volume 25 (Number 1), pp. 33-56, 2018.
12. 荒牧英治, 若宮翔子, 矢野憲, 永井宥之, 岡久太郎, 伊藤薫: 病名アノテーションが付与された医療テキスト・コーパスの構築 (技術資料), *自然言語処理「言語処理の応用システム」特集号*, Volume 25 (Number 1), pp. 119-152, 2018.
13. 柴田大作, 若宮翔子, 木下彩栄, 荒牧英治: 音声発話による認知症スクリーニング技術の開発, *医療情報学*, Volume 37 (Number 6), pp. 303-312, 2018.

国際会議

1. Eiji Aramaki, Chihiro Honda, Shoko Wakamiya, Akira Sato, Isao Miyashiro: Quick Cognitive Impairment Test for Cancer Patients using Emotional Stroop Effect, *MedInfo 2019 (Poster)* (France, Lion), 2019 (2019/08/27).
2. Wannita Takerngsaksiri, Shoko Wakamiya and Eiji Aramaki, City Link: Finding Similar Areas in Two Cities using Twitter Data, In *Proc. of International Symposium on Web and Wireless Geographical Information Systems (W2GIS 2019)* (Kyoto, Japan), Vol. 11474, pp. 13-27, 2019 (2019/05/16).
3. Shoko Wakamiya, Panote Siriaraya, Yihong Zhang, Yukiko Kawai, Eiji Aramaki, Adam Jatowt: Pleasant Route Suggestion based on Color and Object Rates, In *Proceedings of the 12th ACM International Conference on Web Search and Data Mining* (Australia, Melbourne), pp. 786-789, 2019 (2019/02/13).
4. Hayate Iso, Kaoru Ito, Hiroyuki Nagai, Taro Okahisa and Eiji Aramaki: Parsing Japanese Tweets into Universal Dependencies (non-archival submission), *Universal Dependencies Workshop 2018 (UDW 2018)* (Belgium, Brussels), 2018 (2018/08/28).
5. Paolo Casani, Hayate Iso, Shoko Wakamiya, Eiji Aramaki: Wisdom in adversity: a twitter study of the Japanese Tsunami, In *Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)* (Spain, Barcelona), 2018 (2018/08/28).
6. Kaoru Ito, Hiroyuki Nagai, Taro Okahisa, Shoko Wakamiya, Tomohide Iwao, Eiji Aramaki: J-MeDic: A Japanese Disease Name Dictionary based on Real Clinical Usage,

- International Conference on Language Resources and Evaluation (LREC) (Japan, Miyazaki), pp. 2365-2369, 2018 (2018/05/11). (<http://aclweb.org/anthology/L18-1375>)
7. Naomi Yamashita, Hideaki Kuzuoka, Keiji Hirata, Takashi Kudo, Eiji Aramaki, Kazuki Hattori: How Information Sharing about Care Recipients by Family Caregivers Impacts Family Communication, In Proceedings of the Conference on Human Factors in Computing Systems (CHI) (Canada, Montreal, Quebec), No. 222, pp. 1-13, 2018 (2018/04/22).
 8. Hideya Yamamoto, Kaoru Ito, Chihiro Honda, Eiji Aramaki: Does Digital Dementia Exist?, AAAI 2018 Spring Symposium (US, Palo alto, California), pp. 310-311, 2018 (2018/03/26).
 9. Katsuya Taguchi, Eiji Aramaki: Novel Location De-identification for Machine and Humans, IUI 2018 UISTDA Workshop (Tokyo, NII), 2018 (2018/03/11).
 10. Daisaku Shibata, Shoko Wakamiya, Kaoru Ito, Mai Miyabe, Ayae Kinoshita, Eiji Aramaki: VocabChecker: Measuring Language Abilities for Detecting Early Stage Dementia, In Proceedings of the 23rd International Conference on Intelligent User Interfaces Companion (IUI Companion) (Tokyo, NII), 2018 (2018/03/10).

国内会議

1. 山本英弥, 伊藤薫, 荒牧英治: 複合語の構成素情報を考慮した病名難易度の推定, 言語処理学会第 25 回年次大会 (NLP2019) 発表論文集 (名古屋, 名古屋大学), pp. 1495-1498, 2019 (2019/3/16).
2. 磯颯, 上原由衣, 石垣達也, 能地宏, 荒牧英治, 小林一郎, 宮尾祐介, 岡崎直観, 高村大也: Data-to-Text における主題遷移のモデル化, 言語処理学会第 25 回年次大会 (NLP2019) 発表論文集 (名古屋, 名古屋大学), pp. 727-730, 2019 (2019/3/15).
3. Sumaila Nigo, 若宮翔子, 荒牧英治: User's Content-Oriented Social-Bot Detection on Twitter, 第 11 回データ工学と情報マネジメントに関するフォーラム (長崎, ホテルオークラ JR ハウステンボス), 2019 (2019/03/04).
4. 村山太一, 清水伸幸, 藤田澄男, 若宮翔子, 荒牧英治: 位置関係を考慮した地域ごとのインフルエンザ流行予測, 第 11 回データ工学と情報マネジメントに関するフォーラム (長崎, ホテルオークラ JR ハウステンボス), 2019 (2019/03/04).
5. 村山太一, 若宮翔子, 荒牧英治: 患者情報の季節性と検索クエリに基づくトレンドを用いたインフルエンザ患者数予測, 第 35 回情報論的学習理論と機械学習 (IBISML) 研究会 (北海道, かでる 2.7), 2018 (2018/11/05).
6. 柴田大作, 若宮翔子, 伊藤薫, 荒川豊, 吉江智秀, 荒牧英治: 電子カルテテキストを自動臨床データベース化する要約システムの開発, 日本医療情報学会「医用知能情報学研究会」人工知能学会「医用人工知能研究会」(SIG-AIMED) 合同研究会 (神奈川, 慶応義塾大学 矢上キャンパス), 2018 (2018/11/21).
7. 森田真季, 荒牧英治, 灘本明代, 宮部真衣: 嗜好とリアルタイム性を考慮した飲食店検索システムの構築, 情報処理学会関西支部 支部大会 (大阪, 大阪大学中之島キャンパス), 2018 (2018/09/30).
8. 長田颯斗, 荒牧英治, 宮部真衣: 言語能力測定システムにおける負荷軽減手法の提案, 情報処理学会関西支部 支部大会 (大阪, 大阪大学中之島キャンパス), 2018 (2018/09/30).
9. 柿本大輔, 宮部真衣, 荒牧英治, 吉野孝: 流言情報の真偽確認促進システムの評価, 情報処理学会関西支部 支部大会 (大阪, 大阪大学中之島キャンパス), 2018 (2018/09/30).
10. 矢野憲, 岩尾友秀, 荒牧英治: MedInput: 病名の自動予測補完による医療テキスト入力支援ツールの構築, 言語処理学会 第 24 回年次大会 (岡山, 岡山コンベンションセンター), 2018 (2018/03/15).
11. 村山太一, 若宮翔子, 荒牧英治: WORD GINI: 使用頻度の偏りを捉える指標の提案とその応用, 言語処理学会 第 24 回年次大会 (岡山, 岡山コンベンションセンター), 2018

(2018/03/14) .

12. 岡久太郎, 柴田大作, 伊藤薫, 荒牧英治: 聞き手の推定した語彙量と話し手の TTR との乖離に関する検討, 言語処理学会 第 24 回年次大会 (岡山, 岡山コンベンションセンター), 2018 (2018/03/13) .
13. 宮部真衣, 荒牧英治: 流言はどのような言葉で指摘されるか? Twitter における流言訂正表現の分析, ライフインテリジェンスとオフィス情報システム研究会 (LOIS) (沖縄 那覇), 2018 (2018/03/01).
14. 塩野雅孝, 荒牧英治, 若宮翔子, 伊藤薫, 後藤考市, 河合優子, 渡昭博, 猿木信裕: 胃腸炎をモデルとした群馬県における感染症早期探知システムの構築, 全国地方衛生研究所公衆衛生情報研究協議会研修会 (東京), 2018 (2018/01/25) 最優秀賞

受賞

1. 2018, 全国地方衛生研究所公衆衛生情報研究協議会研修会 (東京), 最優秀賞

解説／総説

1. 荒牧英治: 人工知能による医薬品副作用情報の自動抽出に関する動向, 医薬品情報学, Volume 20 (Number 1), 2018.

外部資金 (新規獲得分)

1. 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: 松井珠乃) (1844 千円)
2. 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) (22500 千円)
3. セキュリティの高い医療情報データベースの構築とそれらを利用した医療有用情報の抽出、解析技術等の開発プロジェクト (SIP-AI ホスピタルによる高度診断・治療システム), 研究分担者: 荒牧英治 (研究代表者: 本多周一) (30500 千円)
4. 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) (1000 千円)

June 20, 2019

[http:// sociocom.jp](http://sociocom.jp)

socialcomputing-office@is.naist.jp