



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 2019

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  
2019 第4年報  
(2019/04/01-2020/03/31)



## Introduction

ソーシャル・コンピューティング研究室の4年目の年報です。本年度は、修士学生の Faith Wavinya Mutinda さん、氏家 翔吾くん、Gamar Ivan Azuaje くん、高 志偉くん、研究生の Danay Tasew Tazebe くん、研究員の矢田 竣太郎さん、眞鍋 雅恵さんが加わりました。この結果、研究スタッフ5名、博士学生3名、修士5名、研究生1名の中規模の構成となりました。本年度修士課程を修了する Nigo Sumaila くんは博士課程に進学するため、学生が減ることなく、拡大しつづけた1年でした。

研究に関しては、論文7報、国際会議6報と例年よりも減少しましたが、AIP-PRISM 創薬、SIP-AI ホスピタルといった大きなプロジェクトで中心的な役割を果たしました。また、富士ゼロックス、LINE 株式会社、Yahoo! Japan 研究所、アマダといった企業との共同研究も多く、社会に近い領域で研究開発を行えました。さらに、世間を騒がしている新型コロナ肺炎に関するデータ解析に協力するなど、本邦におけるソーシャルメディアと医療情報において有数の拠点に相応しい活動を行えたと思います。

昨今、大学の情報学研究は ICT 企業の研究とよく似てきています。これは企業と大学との人材の流動性が高まるといったよい面もあるのですが、研究室の規模で企業と同じような研究を行うと、企業よりも遥かに小さなデータや開発の規模で、現実味のない研究を行ってしまう危険もあります。企業にできる研究は企業にまかせて、大学の研究室は大学でしかできない研究を行うことも必要ではないでしょうか？ 大学でしかできない研究の一つが医療研究であり、営利企業ではできない医療研究を実施可能です。本研究室が実施する共同研究が多いのは、如何に本研究室が本邦において貴重な存在であるかを示していると思います。

今後も、ソーシャルメディアテキストの情報学的応用と医療言語処理において日本のリーダーとなる人材を育成することをミッションとし、研究・教育ともに日本の情報学研究のモデルとなることを目指します。

## Research

共同研究に連動した研究にて、多くの成果がありました。また、氏家くんが中心となって進めた富士ゼロックス共同研究による副作用判定支援のための文書分類システムのプレスリリースなど実用化に近いレイヤーでの研究もあり、応用を特徴としている本研究室らしい活動が行えました。一方で、学生の主著による論文は2報にとどまり、プロジェクトに追われ、新しい研究の立ち上げに十分な時間を割けない実態も浮き彫りになりました。今後の学生の増加を考え

ると、定例ミーティング体制、メンターなどシステムティックな研究教育体制が必要だと考えています。

## Project

AIP-PRISM や SIP-AI ホスピタルという2つの内閣府プロジェクトを軸に、他にも多くの研究を推進しました。

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（厚生労働省事業）
  3. 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）
  4. 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）
  5. 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
  6. 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
  7. 富士ゼロックス共同研究：医学論文からの医薬品有害事象の自動抽出
  8. 大阪国際がんセンター共同研究（厚生労働省事業）
  9. エスエス製薬学術指導
  10. SMS 学術指導
  11. 帝國製薬学術指導
- など

## Education

今年度前半は、博士課程学生の磯くんの海外インターン、柴田くんの東大病院でのインターンなどがあったこと、残りの学生も各々の研究に取り組んでいたことから、昨年度同様週1回のData ミーティングを中心とし、フレキシブルな体制で運営を行いました。今年度後半は、10月入学の修士課程学生・研究生の増員により、Data ミーティング（週1回）に加え、医療関連

の研究に関する Med ミーティング（週1回）と論文紹介ミーティングであるペーパーリーディング・ミーティング（週1回）を開催しました。また、Data ミーティングにおいても全員が毎週進捗報告するというシステムを隔週にするなど、フレキシブルから少し体制だった教育システムを模索しつつあります。

研究内容は、医療文書における文の意味的類似度判定（Faithさん）や副作用文書分類（氏家くん）といった医療言語処理に関する研究に加え、言語教育（Gamarくん）やWeb上の暴言の特定（高くん）といった医療と距離をおいた研究も開始し、研究のスコープを大きく広げつつあります。また、今年度は2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（ツイートからの医療施設に関する評判抽出と分析（清水くん）、新型コロナウイルスに関するツイートの分析（土屋くん））に取り組んでもらいました。さらに、入学予定学生に研究技術員やアルバイトを募り、共同研究に関わるタスクなどに従事してもらいました。他分野からの入学生が多く、何を勉強すべきかという問い合わせが多いのですが、入学前の研究技術員やアルバイトを導入することで、研究室活動に事前に参加して実際のタスクベースで必要な技術をその都度勉強して進めてもらうことができるため、学生にとっても在籍メンバーにとっても大きなメリットをもたらしていると考えています。

2019 年度修士論文：

Nigo Sumaila 「Web Search-based Surveillance of Multiple Diseases in Multiple Countries」

## Social Responsibility

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。

## Member

- 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） SIP-AI ホスピタル、AIP-PRSIM による雇用

# YEARBOOK 2019

- 城綾実（博士研究員）（2019/02-2020/03）
- 矢田竣太郎（研究員）（2019/09-）
- 眞鍋雅恵（研究員）（2019/04-）
- 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
  - 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）
  - 氏家 翔吾（研究技術員）（2019/08-09）
  - 久田祥平（研究技術員）（2019/12-2020/03）（来年度新入学生）
  - 古賀 貴士（研究技術員）（2020/02-2020/03）（来年度新入学生）
  - 安西 崇（アルバイト）（2020/01-2020/03）（来年度新入学生）
  - 伊藤 英里（アルバイト）（2020/03）（来年度新入学生）
- 学生
  - 磯颯（D3）
  - 柴田大作（D2）
  - 村山太一（D1）
  - Nigo Sumaila（M2）（2020/03 修了・博士課程進学（当研究室））
  - Faith Wavinya Mutinda（M1・2019/04-）
  - 氏家 翔吾（M1・2019/10-）
  - Gamar Ivan Azuaje（M1・2019/10-）
  - 高 志偉（M1・2019/10-）
- 研究生
  - Danay Tasew Tazebe（2019/10-2020/03）（来年度新入学生）
- インターン
  - 清水 風（神戸高専）（2019/09）
  - 土屋 諒太（松江高専）（2020/03）

## 業績概要

- 原著論文 12 本

- 英文原著 6 本
- 和文原著 6 本
- 国際会議（査読有）
  - 10 本

外部獲得資金（本年度予算 62,500 千円）

- 荒牧
  - AI ホスピタル 22,390 千円
  - AIP-PRISM 13,800 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，厚労科研など）多数
- 若宮
  - 若手研究 900 千円
  - その他，分担（厚労科研，基盤 A，基盤 B など）
- 城
  - 分担（基盤 B）

## Publication

### 論文

1. Nigo Sumaila, Shoko Wakamiya, Eiji Aramaki: Making Twitter Safer: Uncovering Social-Bot on Twitter through User's Content Approach, DBSJ Journal, 18(1), 1-7, 2020. (2020/3/1)
2. Kazuya Taira, Taichi Murayama, Sumio Fujita, Mikiko Ito, Kei Kamide, Eiji Aramaki: Comparing the medical term usage patterns of professionals with those of search engine/community question answering service users in Japan: Log Analysis, J Med Internet Res, 0(0), e0, 2020. (2020/2/4) (impact factor=4.9)
3. 中島和江, 吉岡大輔, 田中晃司, 増田真一, 荒牧英治, 中島伸: 「医療安全—患者と医師が信頼しあえる外科医療を目指して」手術はどのようにうまく行われているのか: 手術チームのコミュニケーション解析 (定期学術集会特別企画記録), 日外会誌, 121(1), 117-119, 2020. (2020/2/3)
4. Tomohide Iwao, Genta Kato, Isao Ito, Eiji Aramaki, Tomohiro Kuroda: A survey of clarithromycin monotherapy and side effect of ethambutol for patients with MAC Lung Disease in Japan: a

- retrospective cohort study using the database of health insurance claims, Pharmacoepidemiology and Drug Safety, 1,6, 2019. (2019/12/25) (impact factor=2.8)
5. 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. (2019/8/25)
  6. 山下直美, 葛岡英明, 山本英弥, 荒牧英治, 平田圭二, 工藤喬, 服部一樹: うつ病患者を介護する家族間の情報共有が家庭内コミュニケーションに及ぼす影響, 情報処理学会論文誌, 60(9) 1580-1593, 2019. (2019/5/9)
  7. 柴田大作, 伊藤薫, 若宮翔子, 荒牧英治: 対照群付き高齢者コーパスの構築とそれを用いた認知症予備軍スクリーニング技術の開発, 人工知能学会論文誌, 第 34 巻 4 号, 2019. (2019/4/8)

#### 国際会議

1. 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, the 12th International Conference on Language Resources and Evaluation (LREC 2020), 2020. (2020/5/14, Marseille, France)
2. Faith Mutinda, Sumaila Nigo, Daisaku Shibata, Shoko Wakamiya, Eiji Aramaki: NAIST: Three approaches for Clinical Semantic Similarity, 2019 n2c2/OHNL Shared Task and Workshop, 2019. (2019/11/15, Washington D.C., US)
3. Mir'atul Khusna Mufida, Wenang Anurogo, Muhammad Zainuddin Lubis, Shoko Wakamiya, Eiji Aramaki: Disease Outbreaks related to Natural Disaster Analysis based on Twitter Data, ICAE 2019 & ICAESS 2019 (Poster), 2019. (2019/10/01, Batam, Indonesia) **Best Poster**
4. Eiji Aramaki, Chihiro Honda, Shoko Wakamiya, Akira Sato, Isao Miyashiro: Quick Cognitive Impairment Test for Cancer Patients using Emotional Stroop Effect, MEDINFO 2019: The 17th World Congress on Medical and Health Informatics (Poster), 2019. (2019/8/25, Lyon, France)
5. Hayate Iso, Yui Uehara, Tatsuya Ishigaki, Hiroshi Noji, Eiji Aramaki, Ichiro Kobayashi, Yusuke Miyao, Naoaki Okazaki, Hiroya Takamura: Learning to Select, Track, and Generate for Data-to-Text, In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (ACL), 2019. (2019/7/28, Florence, Italy)

6. Wannita Takerngsaksiri, Shoko Wakamiya, Eiji Aramaki: City Link: Finding Similar Areas in Two Cities using Twitter Data, In Proceedings of International Symposium on Web and Wireless Geographical Information Systems (W2GIS 2019), Vol. 11474, pp. 13-27, 2019. (2019/5/16, Kyoto, Japan)

## 国内会議

1. Mutinda Faith Wavinya, Sumaila Nigo, 柴田大作, 矢田峻太郎, 若宮翔子, 荒牧英治: Detecting Redundancy in Electronic Medical Records Using Clinical BERT, 言語処理学会第 26 回年次大会 (NLP2020), 613-616, 2020. (2020/03/18, オンライン)
2. 氏家翔吾, 岡守男, 島谷浩司, 井阪航, 有本茜, 矢田峻太郎, 若宮翔子, 荒牧英治: SIGNAL CATCHER: 医学論文を対象とした医薬品有害事象自動判定システムの構築, 言語処理学会第 26 回年次大会 (NLP2020), 609-612, 2020. (2020/03/18, オンライン)
3. 田川裕輝, 西埜徹, 谷口元樹, 谷口友紀, 大熊智子, 若宮翔子, 荒牧英治: 生成された読影所見の自動評価に向けた固有表現認識とモダリティ推定, 言語処理学会第 26 回年次大会 (NLP2020), 5-8, 2020. (2020/03/17, オンライン)
4. 家根和希, 原口和貴, 佐藤亮, 荒牧英治, 宮代勲, 灘本明代: チャット型漫才アプリケーションの提案, 第 12 回データ工学と情報マネジメントに関するフォーラム, A3-3, 2020. (2020/03/03, オンライン)
5. 村山太一, 若宮翔子, 荒牧英治: 訂正投稿を用いたフェイクニュース収集システムの開発, 第 12 回データ工学と情報マネジメントに関するフォーラム, P2-19, 2020. (2020/03/03, オンライン)
6. 村山太一, 若宮翔子, 荒牧英治: 点過程を用いたフェイクニュース検出, 第 12 回データ工学と情報マネジメントに関するフォーラム, A6-4, 2020. (2020/03/03, オンライン)
7. Sumaila Nigo, Shoko Wakamiya, Eiji Aramaki: Web Search-based Surveillance of Multiple Diseases in Multiple Countries, 第 12 回データ工学と情報マネジメントに関するフォーラム, C3-1, 2020. (2020/03/03, オンライン)
8. 村山太一, 若宮翔子, 荒牧英治: SNS を用いた選挙の争点に対する支持政党ごとの傾向抽出: 2019 年大阪ダブル選挙を例に, 第 4 回計算社会科学ワークショップ (CSSJ2020), 2020. (2020/03/01, 神戸大学, 兵庫)
9. 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 サービスへの投稿質問を用いた認知症介護に関する介護者の情報ニーズの探索的分析, ヘルスコミュ



ニケーション学会, 2019. (2019/09/21, 東京大学, 東京)

10. 原口和貴, 大塚真吾, 荒牧英治, 若宮翔子, 灘本明代: 複数ニュースを用いた漫才台本自動生成, 電子情報通信学会 データ工学研究会 (DE), 2019. (2019/09/10, 工学院大学, 東京) **優秀学生ポスター発表賞**
11. 家根和希, 原口和貴, 佐藤亮, 荒牧英治, 宮代勲, 灘本明代: がん患者を対象とした漫才ロボットによる笑いの実証実験, 電子情報通信学会 データ工学研究会 (DE), 2019. (2019/09/10, 工学院大学, 東京) **学生奨励賞**

## 外部資金

1. 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・若宮翔子 (研究代表者: 宮代勲) (200 千円・600 千円)
2. 災害時における状況を考慮したユーザセント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・若宮翔子 (研究代表者: 灘本明代) (300 千円・300 千円)
3. 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) (1400 千円)
4. 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・若宮翔子 (研究代表者: 松井珠乃) (300 千円・300 千円)
5. 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) (24500 千円)
6. 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)
7. 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・城綾実 (研究代表者: 中島和江) (1000 千円・50 千円)
8. 家庭内コミュニケーションの活性化により精神疾患患者の症状改善を促す情報環境の構築 (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: 山下直美) (3420 千円)



9. 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) (4000 千円)
10. 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) (900 千円)
11. 戦略的社会サービスのためのリアルタイム型サイバーフィジカル時空間分析に関する研究 (科学研究費補助金 基盤研究 (A)), 研究分担者: 若宮翔子 (研究代表者: 石川佳治) (100 千円)
12. 小児領域の医薬品の適正使用推進のための人工知能を用いた医療情報データベースの利活用に関する研究 (厚生労働科学研究費補助金 (臨床研究等 ICT 基盤構築研究事業)), 研究分担者: 荒牧英治 (研究代表者: 栗山猛) (研究代表者一括計上)