

戦略的創造研究推進事業  
さきがけ(個人型研究)  
追跡評価用資料

研究領域  
「知の創生と情報社会」  
(2008年度～2015年度)

研究総括: 中島 秀之

2022年3月



## 目次

要旨 .....	1
第 1 章 研究領域概要.....	3
1.1 戦略目標.....	3
1.2 研究領域の目的.....	3
1.3 研究総括.....	3
1.4 領域アドバイザー.....	3
1.5 研究課題および研究者.....	4
第 2 章 追跡調査 .....	9
2.1 追跡調査について.....	9
2.1.1 調査の目的.....	9
2.1.2 調査の対象.....	9
2.1.3 調査方法 .....	11
2.2 追跡調査概要.....	12
2.2.1 研究助成金.....	12
2.2.2 論文 .....	21
2.2.3 特許 .....	24
2.2.4 受賞 .....	26
2.2.5 招待講演 .....	30
2.2.6 報道 .....	30
2.2.7 共同研究や企業との連携.....	30
2.2.8 実用化・製品化.....	31
2.2.9 ベンチャー.....	31
2.3 研究成果から生み出された科学技術や社会・経済への波及効果.....	41
2.3.1 研究領域の展開状況(展開図).....	41
2.3.2 研究成果の科学技術の進歩への貢献.....	44
2.3.3 研究成果の社会・経済への貢献.....	45
2.3.4 その他の特記すべき事項(新たな展開や分野間融合).....	46
第 3 章 各研究課題の主な研究成果 2008 年度採択研究課題 .....	47
3.1.1 大規模グラフ系列からの知識体系化と理解支援手法の開発 (猪口明博) ..	47
3.1.2 ロボットの視覚・触覚を用いた環境情報獲得手法の開発 (大野和則) ....	48
3.1.3 仮説世界と物理世界の相互浸透モデリングによる知の創生 (大羽成征) ..	49
3.1.4 大規模並列化によるハイパフォーマンス人工知能技術 (岸本章宏).....	50
3.1.5 大規模画像データの潜在情報抽出に基づく画像生成 (島野美保子).....	51
3.1.6 擬似コード変換と統計解析による文書画像からの知識抽出 (寺沢憲吾) ..	52

3.1.7 健康被害を監視するための多言語ウェブサーベイランスシステム (Nigel H. COLLIER) .....	53
3.1.8 時空間解析に基づくインターネット異常トラフィックの検出とそのデータベ ース化 (福田健介).....	54
3.1.9 マルチソースデータ高度利用のための統計的データ融合 (星野崇宏)....	55
3.1.10 ネットワーク理論と機械学習を用いたウェブ情報の構造化・知識化 (松尾 豊) .....	56
3.2 2009 年度採択研究課題 .....	57
3.2.1 物語構造に基づく情報編纂基盤技術 (赤石美奈).....	57
3.2.2 マルチエージェントの交渉と協調に基づく集合的コラボレーション支援シ ステムの開発 (伊藤孝行).....	58
3.2.3 次世代データ同化: 自動モデル化と情報フロー抽出技術開発 (上野玄太)	59
3.2.4 大規模データに対する高速類似性解析手法の構築 (宇野毅明).....	60
3.2.5 映像分析による知識の抽出と、その利用による新たな映像合成 (岡部誠)	61
3.2.6 ベイジアンテレビ: 取材・配信・編集を自動化した緊急情報メディア (北本 朝展) .....	62
3.2.7 圧縮データ索引に基づく巨大文書集合からの関連性マイニング (坂本比呂 志) .....	63
3.2.8 実社会情報ネットワークからのプライバシー保護データマイニング (佐久間 淳) .....	64
3.2.9 密度比推定による大規模・高次元データの知的処理技術の創生 (杉山将)	65
3.2.10 インフルエンザ感染伝播のデータ同化モデルによる解析・予測技術 (鈴木 秀幸) .....	66
3.2.11 金融市場における相転移の時空間構造の自動抽出と予測 (高田輝子)...	67
3.2.12 空間的な情報システムの設計開発支援システム (中西泰人).....	68
3.3 2010 年度採択研究課題 .....	69
3.3.1 問題構造の解析に基づく組合せ最適化アルゴリズムの自動構成 (梅谷俊治) .....	69
3.3.2 大規模会話データに基づく個別適合型認知活動支援 (大武美保子).....	70
3.3.3 高精度でスケーラブルな多項関係予測の実現 (鹿島久嗣).....	71
3.3.4 組合せ論的計算に基づく超高次元データからの知識発見 (河原吉伸)....	72
3.3.5 自然言語テキストの高精度で頑強な意味解析とその応用 (Alastair Butler) .....	73
3.3.6 計算論的メディア操作の形式化 (浜中雅俊).....	74
3.3.7 高性能ストリーム・コンピューティング環境の構築 (山際伸一).....	75
3.3.8 能動センシングによる日用柔軟物の情報知識化とその応用 (山崎公俊)..	76



## 要旨

本報告書は、戦略的創造研究推進事業のさきがけ(個人型研究)の研究領域「知の創生と情報社会」(2008年度～2015年度)において、研究終了後一定期間を経過した後、副次的効果を含めて研究成果の発展状況や活用状況等を明らかにし、国立研究開発法人科学技術振興機構(JST)事業及び事業運営の改善等に資するために、追跡調査を実施した結果をまとめたものである。

この領域では、「多様で大規模な情報から『知識』を生産・活用するための基盤技術の創出」を戦略目標として、現在の計算機環境を活用した人工知能の創出や、現在のデータ爆発環境下での情報圧縮やリアルタイムセンシングという情報学的技術の適用、データ爆発環境下における統計的手法の社会適用等の研究が行われた。またこの研究を通じて、大規模なデータを獲得、処理する新しい仕組みを考えることで、社会が抱える問題を解決すると共に社会全体の機能を効率化し、人間の知的作業の質や量の向上を目指した。

追跡調査の結果、本領域の研究成果は終了後も発展、深化しているだけでなく、新たなデータマイニング技術やデータ処理のための基盤技術の開発等に展開していることがわかった。また応用的な技術としてデータを用いた解析によって人間の意思決定等を支援する技術が開発された。

特に注目すべき研究成果としては、クラスタリングと呼ばれるデータ処理技術により既存手法の問題点を改善した宇野毅明の研究、暗号化データに対して様々な統計量を高速計算できる平文拡張演算を開発した佐久間淳の研究、圧縮検索による大規模データベースからの剽窃検出に成功した坂本比呂志の研究、GPSと壁や障害物を観測したセンサ情報を組み合わせた自律的情報収集技術を開発した大野和則の研究、共想法の発展で高齢者の認知機能改善や声掛け見守りロボットを開発した大武美保子の研究等が挙げられる。

さきがけの研究に関連する学術論文は、成果論文が290報、発展論文が302報であった。特許出願は、研究期間中に、国内49件、海外6件、研究終了後は、国内80件、海外38件であった。そのうち、研究期間中の出願から、国内35件、海外6件が、研究終了後に出願から、国内28件、海外12件が登録されている。また研究終了後の受賞は93件あり、招待講演については497件であった。

研究期間中から研究終了後にかけて、本研究領域の研究者は、JSTのCRESTに7件が採択され、研究の発展だけでなく、ディープラーニングやロボティクス等の分野で、我が国の研究開発を先導する役割を果たしつつある。

上記のような研究終了後の進展を以下のような目次に沿って、本報告書をまとめた。

第1章は、研究領域の戦略目標、研究領域の目的、研究総括、領域アドバイザー、研究課題と研究代表者を記載した。

第2章は、追跡調査の実施の概要と、調査結果の概要、そして領域内の特徴について記載した。具体的には、2.1では、追跡調査の目的、調査の対象、調査方法を記載した。2.2では、各研究代表者が獲得した研究助成金、発表した論文、出願・登録した特許等の成果概要を記載した。2.3では、2.2の成果や各研究課題の展開状況から見た科学技術的およ

び社会・経済的アウトカムの概要を記載した。また、研究領域の展開状況を、展開図としてまとめた。

第3章は、さきがけの研究領域終了後の各研究課題の研究の継続と発展状況について、科学技術の進歩の貢献および社会・経済的な波及効果の観点から詳述した。具体的には、研究者によるさきがけ研究期間中の成果を端的に纏めるとともに、それを踏まえて研究終了後に発展した内容について、代表的な事例を個別に記載した。ベンチャー設立など、特筆した成果のあるものは、特記事項として整理した。

## 第 1 章 研究領域概要

### 1.1 戦略目標

「多様で大規模な情報から『知識』を生産・活用するための基盤技術の創出」

本戦略目標では、学術、医療、金融、防災、サービス分野等の多様なニーズに応じて、当該分野を高度化、効率化するための知的情報基盤の確立をめざし、様々な分野で生成・蓄積された多様で大規模な情報から『知識』を生産・活用するための基盤技術を創出する。

なお、ここで言う『知識』とは、社会における人間の活動目的に応じて必要とされる有用な情報のことであり、計算機を使用した情報処理技術等により創出される。

### 1.2 研究領域の目的

本研究領域「知の創生と情報社会」（2008 年度発足）は、多様もしくは大規模なデータから、有用な情報である「知識」を生産し、社会で活用するための基盤的技術となる研究を対象としている。

具体的には、大規模データを処理するための革新的な技術、統計数理科学を応用した分析・モデル化技術、あるいは実社会から得られる多様なデータを構造化・分析して知識を抽出する技術、センサによる情報取得やシミュレーション結果等の複数のリソースから新たな知識を創出する技術などの基盤技術に加えて、獲得した知識を実社会に適用するために必要とされる、シミュレーション、データの可視化、新しい情報社会の仕組みを支える応用技術などに関する研究が含まれる。

### 1.3 研究総括

中島秀之

（領域発足時）

公立ほこだて未来大学 理事長・学長

（領域終了時）

公立ほこだて未来大学 理事長・学長

（追跡調査時）

札幌市立大学 理事長・学長

### 1.4 領域アドバイザー

領域アドバイザー選定にあたっては、以下を基本にして、それぞれの分野から下記の方々をお願いした。

- 戦略目標に沿って、研究総括と共に研究を評価・指導できること
- 高度な専門知識のみならず科学技術や社会動向など幅広い見識を持っていること
- 若手研究者の育成に重きを置いてアドバイスが行えること



- 領域の幅広い技術範囲をカバーするために、専門分野やテーマのバランスを考慮し、工学、理学の幅広い分野での第一人者を選ぶこと。アドバイザーと専門分野は下記の通り
  - データ解析、マイニング、計算理論分野：麻生英樹、有村博紀
  - 人工知能分野：高野明彦、堀浩一
  - 人文科学分野：林晋、安田雪
  - 統計学、大量データ解析分野：樋口知之、鷺尾隆
  - ネットワーク理論分野：林幸雄

表 1-1 に領域アドバイザーを示す。

表 1-1 領域アドバイザー

領域 アドバイザー	所属	役職	任期
麻生 英樹	産業技術総合研究所知能システム 研究部門	主任研究員	2008年5月～2014年3月
有村 博紀	北海道大学大学院情報科学研究科	教授	2008年5月～2014年3月
高野 明彦	国立情報学研究所連想情報学研究 開発センター	センター長・教 授	2008年5月～2014年3月
林 晋	京都大学大学院文学研究科	教授	2008年5月～2014年3月
林 幸雄	北陸先端科学技術大学院大学知識 科学研究科	准教授	2008年5月～2014年3月
樋口 知之	統計数理研究所	所長	2008年5月～2014年3月
堀 浩一	東京大学大学院工学系研究科	教授	2008年5月～2014年3月
安田 雪	関西大学社会学部	教授	2008年5月～2014年3月
鷺尾 隆	第一三共RDノバーレ(株)/大阪大学 産業科学研究所	教授	2008年5月～2014年3月

(注)所属と役職はさきがけ終了時点に記載

## 1.5 研究課題および研究者

研究者として、第1期10名、第2期12名、第3期8名の計30名を採択した。第2期の伊藤は、内閣府の「最先端・次世代研究開発支援プログラム(NEXT)」に採択されたため、途中で研究を終了している。

表 1-2 研究課題と研究者(第1期、第2期、第3期)

期 (研究期間)	研究課題	研究者	採択時の 所属・役職	終了時の 所属・役職	追跡調査時の 所属・役職
第1期 (2008年10 月～2012年3 月)	大規模グラフ系 列からの知識体 系化と理解支援 手法の開発	猪口 明博	大阪大学産業科 学研究所 助教	大阪大学産業 科学研究所 助教	関西学院大学 理工学部 教 授

期 (研究期間)	研究課題	研究者	採択時の 所属・役職	終了時の 所属・役職	追跡調査時の 所属・役職
第1期 (2008年10月～2012年3月)	ロボットの視覚・触覚を用いた環境情報獲得手法の開発	大野 和則	東北大学大学院情報科学研究所助教	科学技術振興機構 さきがけ研究者	東北大学未来科学技術共同研究センター准教授
第1期 (2008年10月～2014年3月)	仮説世界と物理世界の相互浸透モデリングによる知の創生	大羽 成征	京都大学大学院情報科学研究科講師	京都大学大学院情報科学研究科 講師	ミイダス(株)HRサイエンス研究所シニアリサーチャー
第1期 (2008年10月～2012年3月)	大規模並列化によるハイパフォーマンス人工知能技術	岸本 章宏	公立ほこだて未来大学システム情報科学部情報アーキテクチャ学科 助教	東京工業大学大学院情報理工学研究科助教	IBM東京基礎研究所 リサーチ・スタッフ・メンバー
第1期 (2008年10月～2014年3月)	大規模画像データの潜在情報抽出に基づく画像生成	島野 美保子	東京大学生産技術研究所 研究員	東京大学生産技術研究所 研究員	情報・システム研究機構国立情報学研究所コンテンツ科学研究系 特任助教/日本学術振興会 特別研究員 RPD
第1期 (2008年10月～2012年3月)	擬似コード変換と統計解析による文書画像からの知識抽出	寺沢 憲吾	北海道大学大学院情報科学研究科知識メディアラボラトリー(VBL) PD 研究員	公立ほこだて未来大学システム情報科学部 助教	公立ほこだて未来大学システム情報科学部情報アーキテクチャ学科 准教授
第1期 (2008年10月～2012年3月)	健康被害を監視するための多言語ウェブサーベイランスシステム	Nigel Collier	国立情報学研究所情報学プリンシプル研究系 准教授	国立情報学研究所情報学プリンシプル研究系 准教授	the University of Cambridge Professor of Natural Language Processing
第1期 (2008年10月～2012年3月)	時空間解析に基づくインターネット異常トラフィックの検出とそのデータベース化	福田 健介	国立情報学研究所情報学アーキテクチャ科学研究系 准教授	国立情報学研究所情報学アーキテクチャ科学研究系 准教授	情報・システム研究機構国立情報学研究所アーキテクチャ科学研究系 准教授
第1期 (2008年10月～2012年3月)	マルチソースデータ高度利用のための統計的データ融合	星野 崇宏	名古屋大学大学院経済学研究科 准教授	名古屋大学大学院経済学研究科 准教授	慶應義塾大学経済学部/大学院経済学研究科 教授
第1期 (2008年10月～2012年3月)	ネットワーク理論と機械学習を用いたウェブ情報の構造化・知識化	松尾 豊	東京大学大学院工学系研究科 准教授	東京大学大学院工学系研究科 准教授	東京大学大学院工学系研究科 教授

期 (研究期間)	研究課題	研究者	採択時の 所属・役職	終了時の 所属・役職	追跡調査時の 所属・役職
第2期 (2009年10月～2013年3月)	物語構造に基づく情報編纂基盤技術	赤石 美奈	東京大学大学院工学系研究科准教授	法政大学情報科学部コンピュータ科学科教授	法政大学情報科学部大学院工学系研究科教授
第2期 (2009年10月～2011年3月)	マルチエージェントの交渉と協調に基づく集合的コラボレーション支援システムの開発	伊藤 孝行 <sup>1</sup>	名古屋工業大学大学院工学研究科 准教授	名古屋工業大学大学院工学研究科 准教授	京都大学大学院情報学研究科 教授
第2期 (2009年10月～2013年3月)	次世代データ同化：自動モデル化と情報フロー抽出技術開発	上野 玄太	統計数理研究所モデリング研究系 助教	統計数理研究所モデリング研究系 准教授	情報・システム研究機構データサイエンス共同利用基盤施設データ同化研究支援センター 教授・センター長/統計数理研究所モデリング研究系 教授/統計科学技術センター センター長
第2期 (2009年10月～2013年3月)	大規模データに対する高速類似性解析手法の構築	宇野 毅明	国立情報学研究所プリンシプル研究系 准教授	国立情報学研究所プリンシプル研究系 准教授	情報・システム研究機構国立情報学研究所情報学プリンシプル研究系 教授・主幹
第2期 (2009年10月～2013年3月)	映像分析による知識の抽出と、その利用による新たな映像合成	岡部 誠	マックスプランク情報科学研究所 博士研究員	電気通信大学情報理工学部総合情報学科 助教	静岡大学大学院総合科学技術研究科工学専攻 准教授
第2期 (2009年10月～2013年3月)	ベイジアンテレビ：取材・配信・編集を自動化した緊急情報メディア	北本 朝展	国立情報学研究所コンテンツ科学研究系 准教授	国立情報学研究所コンテンツ科学研究系 准教授	ROIS-DS人文学オープンデータ共同利用センター センター長/情報・システム研究機構国立情報学研究所コンテンツ科学研究系 教授
第2期 (2009年10月～2013年3月)	圧縮データ索引に基づく巨本文書集合からの関連性マイニング	坂本 比呂志	九州工業大学大学院情報工学研究院 准教授	九州工業大学大学院情報工学研究院 准教授	九州工業大学大学院情報工学研究院 教授

<sup>1</sup> 最先端・次世代研究開発支援プログラムに採択されたため、2010年度末でさきかけ研究を中止。

期 (研究期間)	研究課題	研究者	採択時の 所属・役職	終了時の 所属・役職	追跡調査時の 所属・役職
第2期 (2009年10月～2013年3月)	実社会情報ネットワークからのプライバシー保護データマイニング	佐久間 淳	筑波大学大学院システム情報工学研究科 准教授	筑波大学大学院システム情報工学研究科 准教授	筑波大学大学院システム情報工学研究科 教授
第2期 (2009年10月～2013年3月)	密度比推定による大規模・高次元データの知的処理技術の創生	杉山 将	東京工業大学大学院情報理工学研究科 准教授	東京工業大学大学院情報理工学研究科 准教授	理化学研究所革新知能統合研究センター長/東京大学大学院新領域創成科学研究科 教授
第2期 (2009年10月～2013年3月)	インフルエンザ感染伝播のデータ同化モデルによる解析・予測技術	鈴木 秀幸	東京大学生産技術研究所 准教授	東京大学生産技術研究所 准教授	大阪大学大学院情報科学研究科 教授
第2期 (2009年10月～2013年3月)	金融市場における相転移の時空間構造の自動抽出と予測	高田 輝子	大阪市立大学大学院経営学研究科 准教授	大阪市立大学大学院経営学研究科 准教授	大阪市立大学大学院経営学研究科 准教授
第2期 (2009年9月～2013年3月)	空間的な情報システムの設計開発支援システム	中西 泰人	慶應義塾大学環境情報学部 准教授	慶應義塾大学環境情報学部 准教授	慶應義塾大学環境情報学部 教授
第3期 (2010年10月～2014年3月)	問題構造の解析に基づく組合せ最適化アルゴリズムの自動構成	梅谷 俊治	大阪大学大学院情報科学研究科 准教授	大阪大学大学院情報科学研究科 准教授	大阪大学大学院情報科学研究科 寄附講座教授/理化学研究所革新知能統合研究センター 客員教員
第3期 (2010年10月～2016年7月)	大規模会話データに基づく個別適合型認知活動支援	大武 美保子	東京大学人工物工学研究センター 准教授	千葉大学大学院工学研究科 准教授	理化学研究所革新知能統合研究センター 認知行動支援技術チームリーダー
第3期 (2010年10月～2014年3月)	高精度でスケラブルな多項関係予測の実現	鹿島 久嗣	東京大学大学院情報理工学系研究科 准教授	東京大学大学院情報理工学系研究科 准教授	京都大学大学院情報学研究科 知能情報学専攻 教授

期 (研究期間)	研究課題	研究者	採択時の 所属・役職	終了時の 所属・役職	追跡調査時の 所属・役職
第3期 (2010年10 月～2014年3 月)	組合せ論的計算 に基づく超高次 元データからの 知識発見	河原 吉伸	大阪大学産業科 学研究所 助教	大阪大学産業 科学研究所 准教授	九州大学 マ ス・フォア・イ ンダストリ研 究所 教授/理 化学研究所革 新知能統合研 究センター構 造的学習チー ム チームリ ーダー
第3期 (2010年10 月～2014年3 月)	自然言語テキス トの高精度で頑 強な意味解析と その応用	Alastair Butler	東北大学高等教 育開発推進セン ター 研究員	東北大学高等 教育開発推進 センター 研 究員	弘前大学人文 社会学部 准 教授
第3期 (2010年10 月～2014年3 月)	計算論的メディ ア操作の形式化	浜中 雅俊	筑波大学大学院 システム情報工 学研究科 講師	筑波大学大学 院システム情 報工学研究科 講師	理化学研究所 革新知能統合 研究センター 音楽情報知能 チーム チー ムリーダー
第3期 (2010年10 月～2014年3 月)	高性能ストリー ム・コンピュー ティング環境の 構築	山際 伸一	高知工科大学情 報学群 准教授	筑波大学シス テム情報系 准教授	筑波大学シス テム情報系 准教授
第3期 (2010年10 月～2014年3 月)	能動センシング による日用柔軟 物の情報知識化 とその応用	山崎 公俊	東京大学 IRT 研 究機構 研究員	信州大学工学 部 助教	信州大学学術 研究院工学系 准教授

## 第 2 章 追跡調査

### 2.1 追跡調査について

#### 2.1.1 調査の目的

追跡調査は研究領域終了後、一定期間を経過した後、副次的効果を含めて研究成果の発展状況や活用状況を明らかにし、JST の事業および事業運営の改善に資するために行うもので、研究終了後の研究者の研究課題の発展状況等を調査した。

#### 2.1.2 調査の対象

本追跡調査は、さきがけ研究領域「知の創生と情報社会(2008 年度～2015 年度)」の研究者全員を対象とする。採択研究者は、2008 年度採択 10 名、2009 年度採択 12 名、2010 年度採択 8 名である。表 2-1 に調査対象と調査対象期間を示す。

表 2-1 調査対象と調査対象期間

採択年	研究者	さきがけ研究期間	さきがけ終了後の調査対象期間
第 1 期 (2008 年)	猪口 明博	2008 年 10 月～2012 年 3 月	2013 年 1 月～調査終了月
	大野 和則		
	大羽 成征	2008 年 10 月～2014 年 3 月	2015 年 1 月～調査終了月
	岸本 章宏	2008 年 10 月～2012 年 3 月	2013 年 1 月～調査終了月
	島野 美保子	2008 年 10 月～2014 年 3 月	2015 年 1 月～調査終了月
	寺沢 憲吾	2008 年 10 月～2014 年 3 月	2015 年 1 月～調査終了月
	Nigel Collier		
	福田 健介		
	星野 崇宏		
	松尾 豊		
第 2 期 (2009 年)	赤石 美奈	2009 年 10 月～2013 年 3 月	2014 年 1 月～調査終了月
	伊藤 孝行 <sup>2</sup>	2009 年 10 月～2011 年 3 月	2012 年 1 月～調査終了月
	上野 玄太	2009 年 10 月～2013 年 3 月	2014 年 1 月～調査終了月
	宇野 毅明		
	岡部 誠		
	北本 朝展		
	坂本 比呂志		
	佐久間 淳		
	杉山 将		

<sup>2</sup> 1 名、最先端・次世代研究開発支援プログラムに採択されたため、2010 年度末でさきがけ研究を中止。

	鈴木 秀幸		
	高田 輝子		
	中西 泰人		
第3期 (2010年)	梅谷 俊治	2010年11月～2014年3月	2015年1月～調査終了月
	大武 美保子	2010年11月～2016年3月	2017年1月～調査終了月
	鹿島 久嗣	2010年11月～2014年3月	2015年1月～調査終了月
	河原 吉伸		
	Alastair Butler		
	浜中 雅俊		
	山際 伸一		
	山崎 公俊		

### 2.1.3 調査方法

調査は、2020年5月～7月にかけて実施した研究者アンケート、2020年8月～10月にかけて実施した研究総括及び一部の研究者とのインタビュー、領域事後評価用資料等の文献、エビデンス情報収集のための各種データベース、取りまとめ後の情報に関する研究者への事実確認を基に実施した。具体的な調査方法は以下の通りである。

#### (1) 研究助成金

調査対象期間は、本研究領域の期間中(2008年10月～2016年3月)を含めて調査対象月とし、本研究領域の研究者が研究の代表を務める研究助成金を調査した。その中から、原則、研究助成金の総額が1千万円/件以上のものを抽出した。

ただし、各研究課題の開始後に研究助成を受け、当該研究課題が終了する前に、その助成期間が終了してしまう事案および当該研究課題終了と同年度に助成期間が終了する事案に関しては対象外とした。

研究助成資金の獲得状況の調査については、主に以下のWEBサイトを用い、2020年6月に検索した。

- 競争的研究資金の機関データベース(科学研究費助成事業データベース、厚生労働科学研究成果データベース)
- 公益財団法人助成財団センター([http://www.jfc.or.jp/grant-search/ap\\_search.php5](http://www.jfc.or.jp/grant-search/ap_search.php5))
- 日本の研究.com(<https://research-er.jp/>)

#### (2) 論文

論文の抽出は、文献データベースとしてScopus(データソース:2021年1月時点)を用い、文献タイプはArticle, Review, Conference Paperを対象とし、2021年1月に抽出した。研究期間中および研究終了後について研究者が著者になっている論文を、Author IDに紐づけて出力し、研究終了後に発表された論文リストを作成した。リストの論文については、各論文における研究者の所属情報や謝辞等の情報、また、研究者アンケートへの記載の有無を基に、①さきがけの成果と認められるもの、②さきがけの発展と認められるもの、③さきがけと無関係と考えられるもの、に分類した。また、研究終了報告書に記載のある論文で、上記の検索方法で抽出されなかった論文については、さきがけの成果と認められるものとして、リストに加えた。各分類における論文リストは、研究者への事実確認を通じて、確定させた。

#### (3) 特許

研究者が発明者になっているもので、研究期間中の特許出願および登録の状況と、研究終了以降の特許出願および登録の状況について調査した。特許データベースULTRA Patentを用いて2020年10月に検索した。当該データベースでは、研究者名(漢字名及びアルファベット名)で検索することにより、上記の必要な情報を一覧として得ることができる。



また、研究終了報告書や研究者アンケートに記載のある特許で、上記の検索方法で抽出されなかった特許についても、さきがけの成果と認められるものとして、リストに加えた。

#### (4) 受賞、招待講演、報道、共同研究や企業との連携等

受賞については、研究終了以降から現在に至るまでの受賞について、2020年8月にウェブ検索を実施し、各研究者の研究室ホームページ、科学研究費補助金(科研費)ホームページなどを参考にし、リストを作成した。また、主要な受賞については、受賞者リストから研究者が受賞者になっている賞を抜粋した。さらに、研究者アンケートに記載のある内容を追加した。

招待講演については、研究終了以降から現在に至るまでの受賞について、2020年8月にウェブ検索を実施し、各研究者の研究室ホームページ、科研費ホームページなどを参考にし、国内外の主要な会議についてリストを作成した。さらに、研究者アンケートに記載のある内容を追加した。

報道については、日経テレコンを用いて、研究者名+所属機関(研究開始時点/研究終了時点/現時点)で2020年6月に検索を行った後、研究者の成果かどうかを目視で確認することにより、絞り込みを行った。

共同研究や企業との連携等については、2020年5月にウェブ検索を実施し、各研究者の研究室ホームページ、科研費ホームページなどを参考にし、情報収集を行った。さらに、研究者アンケートに記載のある内容を追加した。

なお、追跡調査にあたっては、各研究者に依頼して、これらすべてのリストと各研究者の主な研究成果の草稿の確認を可能な限りご協力頂いた。

## 2.2 追跡調査概要

### 2.2.1 研究助成金

研究発展状況を把握するために、研究終了後にどれだけ外部資金を獲得しているかを把握することは非常に重要である。当該領域の代表研究者の外部資金獲得状況を表 2-2 に示す。

ほぼすべての研究者が研究終了後も科研費を中心に競争的研究資金を獲得して、研究開発を継続的に行っている。科研費の獲得件数では、星野崇宏が4件、福田健介、坂本比呂志、佐久間淳、杉山将、鹿島久嗣、山際伸一が3件を獲得している。大型資金の獲得については、CRESTでは、鈴木秀幸が2件、大羽成征、伊藤孝行、宇野毅明、佐久間淳、河原吉伸が1件と合計7件を獲得している。また、大武美保子がAIPを1件、大野和則がSIPを1件獲得している。

表 2-2 研究助成金獲得状況



研究者	研究期間 (年度)	研究種目	研究課題	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	金額 (百万円)
猪口 明博	2008～ 2011	さきがけ	大規模グラフ系列からの知識体系化と理解支援手法の開発	■	■	■															40.0
	2010～ 2013	科研費 若手研究 (A)	表構造の異なる複数の時間履歴データからの時系列分析多次元データベースの構築手法				■	■	■												
大野 和則	2008～ 2011	さきがけ	ロボットの視覚・触覚を用いた環境情報獲得手法の開発	■	■	■															40.0
	2013～ 2015	科研費 若手研究 (A)	レスキューロボドッグ：ロボットテクノロジーを用いた災害救助犬の建物内の探査の高度化						■	■											25.2
	2014～ 2018	科学技術 政策 SIP インフラ 維持管 理・更 新・マネ ジメント 技術 ロボ ット技術 の研究開 発 維持管 理ロボッ ト・災害 対応ロボ ットの開 発 維持管 理ロボッ ト	橋梁の打音検査ならびに近接目視を代替する飛行ロボットシステムの研究開発							■	■	■	■								500.0
	2018～ 2020	科研費 基盤研究 (B)	ヒトとイヌのコミュニケーション方法の解明・応用												■	■					

研究者	研究期間 (年度)	研究種目	研究課題	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	金額 (百万円)
	2018～ 2022	「次世代人工知能・ロボットの中核となるインテグレート技術開発」に係る公募について	ロボット技術と人工知能を活用した地方中小建設現場の土砂運搬の自動化に関する研究開発																		485.6
大羽 成征	2008～ 2013	さきがけ	仮説世界と物理世界の相互浸透モデリングによる知の創生																		50.0
	2012～ 2014	AMED	非正常環境におけるロボast適応 BCI																		15.0
	2015～ 2019	次世代人工知能・ロボット中核技術開発	人間と相互理解できる次世代人工知能技術の研究開発(辻井潤一)																		57.9
	2017～ 2019	CREST	自由行動下での神経情報操作・解読技術の開発と応用(渡辺大)																		26.0
岸本 章宏	2008～ 2011	さきがけ	大規模並列化によるハイパフォーマンス人工知能技術																		40.0
島野 美保子	2008～ 2013	さきがけ	大規模画像データの潜在情報抽出に基づく画像生成																		40.0
寺沢 憲吾	2008～ 2011	さきがけ	擬似コード変換と統計解析による文書画像からの知識抽出																		40.0
	2011～ 2014	科研費 若手研究 (A)	高次元特徴量ベクトルの最近傍探索を行う改良型LSHアルゴリズムの研究																		19.6
Nigel Collier	2008～ 2011	さきがけ	健康被害を監視するための多言語ウェブサーベイランスシステム																		40.0
福田 健介	2008～ 2011	さきがけ	時空間解析に基づくインターネット異常トラフィックの検出とそのデータベース化																		40.0

研究者	研究期間 (年度)	研究種目	研究課題	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	金額 (百万円)
	2011～ 2013	科研費 若手研究 (A)	大規模 DNS トラフィック の統計的解析と異常検出				■	■	■												16.8
	2015～ 2017	科研費 基盤研究 (B)	アンビエント DNS センサ ーに関する研究								■	■	■								18.3
	2018～ 2020	科研費 基盤研究 (B)	DNS バックスキャッター による IPv6 ネットワーク での大規模スキャン検 出												■	■					17.2
星野 崇宏	2008～ 2011	さきがけ	マルチソースデータ高度 利用のための統計的デー タ融合	■	■	■															40.0
	2011～ 2013	科研費 若手研究 (A)	ソーシャルメディア発信 情報のバイアス補正法の 研究				■	■	■												13.3
	2014～ 2017	科研費 基盤研究 (B)	非実験研究での介入効果 推定法の総合的研究と実 用化							■	■	■									14.4
	2018～ 2021	科研費 基盤研究 (B)	異質性を考慮した因果効 果の推定法の開発とその 応用											■	■	■					16.8
	2020～ 2022	科研費 国際共同 研究加速 基金(国際 共同研究 強化(A))	異質性を考慮した因果効 果推定の行動経済学分野 での活用:応用と手法開 発の往還														■	■	■		15.2
松尾 豊	2008～ 2011	さきがけ	ネットワーク理論と機械 学習を用いたウェブ情報 の構造化・知識化	■	■	■															40.0
	2013～ 2016	科研費 若手研究 (A)	ウェブ上のソーシャルメ ディア分析によるゲーミ フィケーションの理論構 築						■	■	■										24.8
	2016～ 2020	科研費 新学術領 域研究(研 究領域提 案型)	ディープラーニングと記 号処理の融合による予測 性の向上に関する研究									■	■	■	■						76.2
赤石 美奈	2009～ 2012	さきがけ	物語構造に基づく情報編 纂基盤技術		■	■	■														40.0
伊藤 孝行	2009～ 2010	さきがけ	マルチエージェントの交 渉と協調に基づく集会的 コラボレーション支援シ ステムの開発		■	■															20.0

研究者	研究期間 (年度)	研究種目	研究課題	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	金額 (百万円)	
	2010～ 2013	JSPS NEXT	環境社会最適化シミュレーションを可能にする社会最適化アルゴリズム創出とその応用																		152.1	
	2014～ 2017	委託研究 ・ 高度通 信・放送 研究開発 委託研究	マルチエージェント未来交通予測による渋滞緩和																			80.0
	2015～ 2017	科研費 基盤研究 (A)	マルチエージェント自動交渉理論とその評価に関する研究																			43.6
	2015～ 2020	CREST	エージェント技術に基づく大規模合意形成支援システムの創成																			150 ～ 498
上野 玄太	2009～ 2012	さきがけ	次世代データ同化:自動モデル化と情報フロー抽出技術開発																			40.0
	2012～ 2016	分野別の 研究開発 環境分野 気候変動 リスク情 報創生プ ログラム SOUSEI	気候変動予測データの統計学的解析手法の開発																			68.7
	2013～ 2015	科研費 基盤研究 (B)	データ同化による日射量・風力の確率分布予測																			17.7
	2017～ 2021	科研費 基盤研究 (A)	結合データ同化システム開発の方法と応用																			40.0
宇野 毅明	2009～ 2012	さきがけ	大規模データに対する高速類似性解析手法の構築																			40.0
	2014～ 2019	CREST	データ粒子化による高速高精度な次世代マイニング技術の創出																			150 ～ 498
	2019～ 2023	科研費 基盤研究 (A)	構造抽出による自然言語ビッグデータへの高次高精度なデータマイニング技術の開発																			44.3
岡部 誠	2009～ 2012	さきがけ	映像分析による知識の抽出と、その利用による新たな映像合成																			40.0
北本 朝展	2009～ 2012	さきがけ	ペイジアンテレビ:取材・配信・編集を自動化した緊急情報メディア																			40.0
	2015～ 2017	SICORP	災害対応・復旧のための人間中心型状況認識プラットフォーム																			30.0

研究者	研究期間(年度)	研究種目	研究課題	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	金額(百万円)	
	2016～2018	科研費 基盤研究(B)	デジタル史料批判:エビデンスベース人文情報学のための連結指向型研究基盤																		17.9	
	2019～2021	科研費 基盤研究(A)	歴史ビッグデータ研究基盤による過去世界のデータ駆動型復元と統合解析																			43.7
坂本 比呂志	2009～2012	さきがけ	圧縮データ索引に基づく巨大文書集合からの関連性マイニング																		40.0	
	2011～2014	科研費 若手研究(A)	圧縮マイニング:超大規模テキストに埋もれている知識の顕在化																			19.6
	2014～2016	科研費 基盤研究(B)	圧縮情報処理によるストリームデータからの知識発見																			12.2
	2017～2021	科研費 基盤研究(B)	ストリームデータを知識化する圧縮情報処理基盤の開発																			17.4
佐久間 淳	2009～2012	さきがけ	実社会情報ネットワークからのプライバシー保護データマイニング																		40.0	
	2012～2016	科研費 若手研究(A)	大規模非構造型秘密情報のためのアウトソース型プライバシー保護データマイニング基盤																			26.3
	2013～2018	CREST	自己情報コントロール機構を持つプライバシー保護データ収集・解析基盤の構築と個別化医療・ゲノム疫学への展開																			150～498
	2016～2018	科研費 基盤研究(B)	機械学習における統計的安全性の理論																			16.0
	2019～2022	科研費 基盤研究(B)	機械と人間の認識ギャップを考慮した深層学習セキュリティ・プライバシーに関する研究																			17.3
杉山 将	2009～2012	さきがけ	密度比推定による大規模・高次元データの知的処理技術の創生																			40.0
	2011～2015	科研費 新学術領域研究(研究領域提案型)	予測と意思決定のための機械学習理論の構築とその神経回路での実現																			96.6
	2013～2016	科研費 若手研究(A)	情報量に基づく新しい機械学習理論とその応用																			25.4



研究者	研究期間 (年度)	研究種目	研究課題	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	金額 (百万円)
	2017～ 2021	科研費 基盤研究 (A)	統計的強化学習の深化と 応用																		45.0
鈴木 秀幸	2009～ 2012	さきがけ	インフルエンザ感染伝播 のデータ同化モデルによる 解析・予測技術																		21.0
	2012～ 2014	CREST	再生可能エネルギーの大量 導入を考慮した電力シス テムの複雑ネットワーク 動力学モデル構築とその 最適化理論の創成																		41.5
	2018～ 2023	CREST	光ニューラルネットワーク の時空間ダイナミクス に基づく計算基盤技術																		260.0
高田 輝子	2009～ 2012	さきがけ	金融市場における相転移 の時空間構造の自動抽出 と予測																		40.0
	2011～ 2015	科研費 若手研究 (B)	テキストマイニングと高 頻度データ解析による金 融バブル生成・崩壊過程 のマイクロ分析																		19.0
中西 泰人	2009～ 2012	さきがけ	空間的な情報システムの 設計開発支援システム																		40.0
梅谷 俊治	2010～ 2013	さきがけ	問題構造の解析に基づく 組合せ最適化アルゴリズ ムの自動構成																		40.0
	2019～ 2022	科研費 基盤研究 (B)	大規模データの特徴抽出 と再利用に基づくサービ ス最適割当アルゴリズム の開発																		10.4
大武 美保子	2010～ 2016	さきがけ	大規模会話データに基づ く個別適合型認知活動支 援																		40.0
	2020～ 2021	科研費 新学術領 域研究(研 究領域提 案型)	ベッドサイド見守り声掛 けロボットのための対話 システムの開発																		18.3

研究者	研究期間 (年度)	研究種目	研究課題	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	金額 (百万円)	
	2020～ 2022	戦略的創造研究推進事業 AIP：人工知能／ビッグデータ／IoT／サイバーセキュリティ統合プロジェクト(AIPプロジェクト) 日独仏 AI 研究	高齢社会のための AI：AI により促進される認知機能訓練のための基礎研究と実用的なツールの開発																		300.0	
鹿島久嗣	2010～ 2013	さきがけ	高精度でスケーラブルな多項関係予測の実現																		40.0	
	2010～ 2014	科研費 若手研究(A)	グラフ構造データの予測的分析のための機械学習手法の研究																			14.8
	2015～ 2019	科研費 基盤研究(A)	ビッグデータ時代の複雑構造データを扱う機械学習法の研究																			43.0
	2020～ 2023	科研費 基盤研究(B)	複雑な関係データに基づく意思決定のための機械学習研究																			17.6
河原吉伸	2010～ 2013	さきがけ	組合せ論的計算に基づく超高次元データからの知識発見																			40.0
	2014～ 2017	科研費 基盤研究(B)	離散凸解析に基づく機械学習アルゴリズム体系の構築とその応用																			16.0
	2018～ 2021	科研費 基盤研究(B)	データからの潜在ダイナミクス抽出のための統計的機械学習とその応用																			16.8
	2019～ 2024	CREST	作用素論的データ解析に基づく複雑ダイナミクス計算基盤の創出																			150～ 498
Alastair Butler	2010～ 2013	さきがけ	自然言語テキストの高精度で頑強な意味解析とその応用																		40.0	



研究者	研究期間 (年度)	研究種目	研究課題	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	金額 (百万円)
浜中雅俊	2010～2013	さきがけ	計算論的メディア操作の形式化																		40.0
	2013～2016	科研費 若手研究 (A)	音楽音響信号に対する音楽操作の蓄積と再利用の実現に関する研究																		10.7
	2017～2020	科研費 基盤研究 (B)	音楽操作の蓄積と再利用に関する研究																		15.7
山際伸一	2010～2013	さきがけ	高性能ストリーム・コンピューティング環境の構築																		40.0
	2012～2014	科研費 基盤研究 (B)	ストリーム指向プログラムのマクロ並列化の研究																		17.9
	2015～2018	科研費 基盤研究 (B)	データストリーム伝送路におけるロスレス圧縮技術の開発																		17.9
	2020～2023	科研費 基盤研究 (B)	大局的エントロピー予測によるデータ圧縮の最適化技法の開発																		17.4
山崎公俊	2010～2013	さきがけ	能動センシングによる日用柔軟物の情報知識化とその応用																		40.0
	2014～2017	科研費 若手研究 (A)	布製品の能動操作による布の知識表現と操作技能の自動獲得																		24.4
	2015～2019	次世代ロボット中核技術開発・次世代人工知能技術分野	不定形物操作システムの研究開発																		70.7
	2019～2020	次世代人工知能・ロボットの中核となるインテグレート技術開発, 人工知能技術の適用領域を広げる研究開発	モデル化難物体の操作知識抽出に基づく柔軟物製品の生産工程改善																		59.5
	2020～2022	科研費 基盤研究 (B)	手順のある不定形物作業における状態予測および操作方法獲得																		17.7

2020年7月16日調査

## 2.2.2 論文

論文発表件数は研究者の研究活動を示す重要な指標であるため、研究者について成果に該当する論文数と発展に該当する論文数とを調査した。検索はいずれも2021年1月28日に実施した。

成果論文数は全体で290報、発展論文数は全体で302報であった。また、それぞれの責任著者となっている論文は86報、29報であった。

論文数20報以上の研究者数は、成果論文では2名で、発展論文では5名である。研究者別では、各研究者間でばらつきはあるが、さきがけの研究成果の論文を多く発表したのは、成果論文数では、杉山が40報、福田が36報の論文を発表している。発展論文数では、佐久間が80報を超える論文を発表しており、次いで鹿島が42報であった。成果論文と比較をすると、佐久間、鹿島、山崎の論文発表活動が大きく発展していることがわかる。

また、Percentileについても調査を実施したところ、Top10%以内の論文数は、成果論文では45報、発展論文では21報であり、福田が5報、また佐久間および鹿島が4報であり、続いて杉山が3報と多い。

Top1%以内の論文数は、成果論文では8報、発展論文では0報であった。特に注目すべき論文としては、Top0.01%以内となっている論文に、松尾のSakaki et al., "Earthquake shakes Twitter users: Real-time event detection by social sensors," Proceedings of the 19th International Conference on World Wide Web, WWW '10, : 851-860 (2010)がある。

なお、本研究領域内での共著論文数は、さきがけ研究成果で10報、一方でさきがけ研究成果の発展に関しては4報あった。

表 2-3 さきがけの成果および発展の論文(原著論文)数

期 (採択年度)	研究者	①さきがけの成果							②さきがけの発展						
		論文数	責任著者 論文数	平均 FWCI 数	Top 論文数				論文数	責任著者 論文数	平均 FWCI 数	Top 論文数			
					10%	1%	0.1%	0.01%				10%	1%	0.1%	0.01%
第1期 (2008年度)	猪口 明博	7	4	1.65	1	0	0	0	5	2	0.13	0	0	0	0
	大野 和則	6	5	1.81	2	0	0	0	12(1)	1	0.86	0	0	0	0
	大羽 成征	10	3	0.63	1	0	0	0	4	1	3.01	2	0	0	0
	岸本 章宏	15	2	2.04	4	1	0	0	4	0	0.00	0	0	0	0
	島野 美保子	7	3	0.48	0	0	0	0	4	2	0.48	0	0	0	0
	寺沢 憲吾	5	5	0.88	1	0	0	0	4	0	0.00	0	0	0	0
	Nigel Collier	13(2)	7	2.03	5	0	0	0	6	1	1.40	2	0	0	0
	福田 健介	36	5	1.95	7	2	0	0	20	2	1.64	5	0	0	0
	星野 崇宏	12	3	0.59	1	0	0	0	2	0	0.72	0	0	0	0

	松尾 豊	8	1	44.23	5	4	1	1	0	0	0.00	0	0	0	0
第2期 (2009年度)	赤石 美奈	4	2	0.30	0	0	0	0	0	0	0.00	0	0	0	0
	伊藤 孝行	1	0	0.50	0	0	0	0	1	0	0.40	0	0	0	0
	上野 玄太	5	4	0.44	0	0	0	0	3	0	0.39	0	0	0	0
	宇野 毅明	6(1)	1	1.05	0	0	0	0	4	0	0.88	0	0	0	0
	岡部 誠	11	1	1.14	1	1	0	0	5	1	0.89	0	0	0	0
	北本 朝展	3	2	0.25	0	0	0	0	2	0	1.53	1	0	0	0
	坂本 比呂志	13(1)	4	0.91	1	0	0	0	9(2)	5	0.60	0	0	0	0
	佐久間 淳	11	2	2.24	4	0	0	0	89(1)	0	0.74	4	0	0	0
	杉山 将	40(7)	9	1.80	7	0	0	0	39	1	0.94	3	0	0	0
	鈴木 秀幸	6	0	0.76	0	0	0	0	1	0	1.10	0	0	0	0
	高田 輝子	1	1	0.17	0	0	0	0	1	0	1.97	0	0	0	0
	中西 泰人	7	3	0.55	0	0	0	0	1	1	0.00	0	0	0	0
第3期 (2010年度)	梅谷 俊治	2	1	0.33	0	0	0	0	3	2	0.26	0	0	0	0
	大武 美保子	4	3	0.39	0	0	0	0	12	1	0.91	0	0	0	0
	鹿島 久嗣	15(1)	1	1.99	4	0	0	0	42	0	0.86	4	0	0	0
	河原 吉伸	6	2	1.67	1	0	0	0	3	0	0.82	0	0	0	0
	Alastair Butler	6	3	0.06	0	0	0	0	0	0	0.00	0	0	0	0
	浜中 雅俊	8	1	0.95	0	0	0	0	4	3	0.20	0	0	0	0
	山際 伸一	15(3)	3	0.43	0	0	0	0	0	0	0.00	0	0	0	0
	山崎 公俊	12	5	1.36	1	0	0	0	22	6	0.57	0	0	0	0
領域全体	290 (5)	86	3.30	45 (1)	8	1	1	302	29	0.94	21	0	0	0	

1 各研究者の論文数は重複論文を含むため、領域全体の論文数の合計数は一致しない。( )中の数値は重複論文数。領域全体の論文数には重複論文数は含まない。重複論文の概要(関係する研究者名)を表2-4、表2-5に記す。

2 責任著者とはCorresponding Authorと同義。

2021年1月28日調査

表2-4 成果論文における共著関係

No	期生	研究者(報数)	共著の概要
1	第1期	ナイジェル・コリア(2)	杉山 将 <sup>A,A</sup> と2報
2	第2期	宇野 毅明(1)	杉山 将 <sup>A</sup> と1報
3	第2期	坂本 比呂志(1)	山際 伸一 <sup>A</sup> と1報
4	第2期	杉山 将(7)	ナイジェル・コリア <sup>A,A</sup> と2報、宇野 毅明 <sup>A</sup> と1報、鹿島久嗣 <sup>A,C,C</sup> と3報、佐久間 淳 <sup>B</sup> と1報

5	第3期	鹿島 久嗣(1)	杉山 将 <sup>A</sup> と1報
6	第3期	山際 伸一(3)	坂本 比呂志 <sup>A,B</sup> と3報

注) 共著論文に対して、各共著者が、Aは成果論文、Bは発展論文、Cはその他と位置付けていることを表す

表 2-5 発展論文における共著関係

No	期生	研究者(報数)	共著の概要
1	第1期	大野 和則(1)	山崎 公俊 <sup>C</sup> と1報
2	第2期	坂本 比呂志(2)	山際 伸一 <sup>A,A</sup> と2報
3	第2期	佐久間 淳(1)	杉山 将 <sup>A</sup> と1報

注) 共著論文に対して、各共著者が、Aは成果論文、Cはその他と位置付けていることを表す

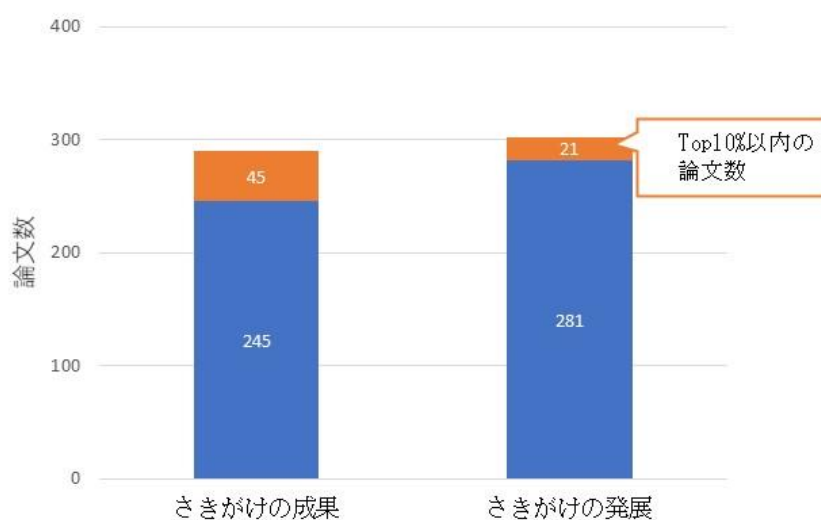


図 2-1 さきがけの成果および発展に関する論文数

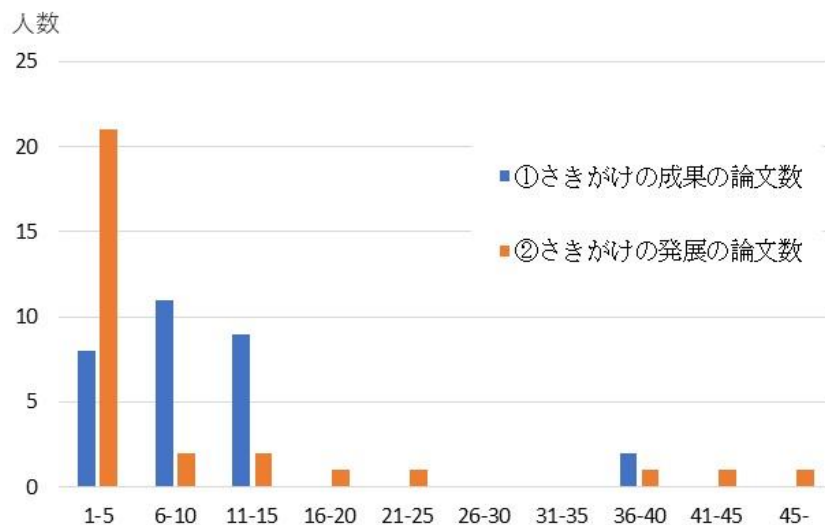


図 2-2 各研究者の論文数分布

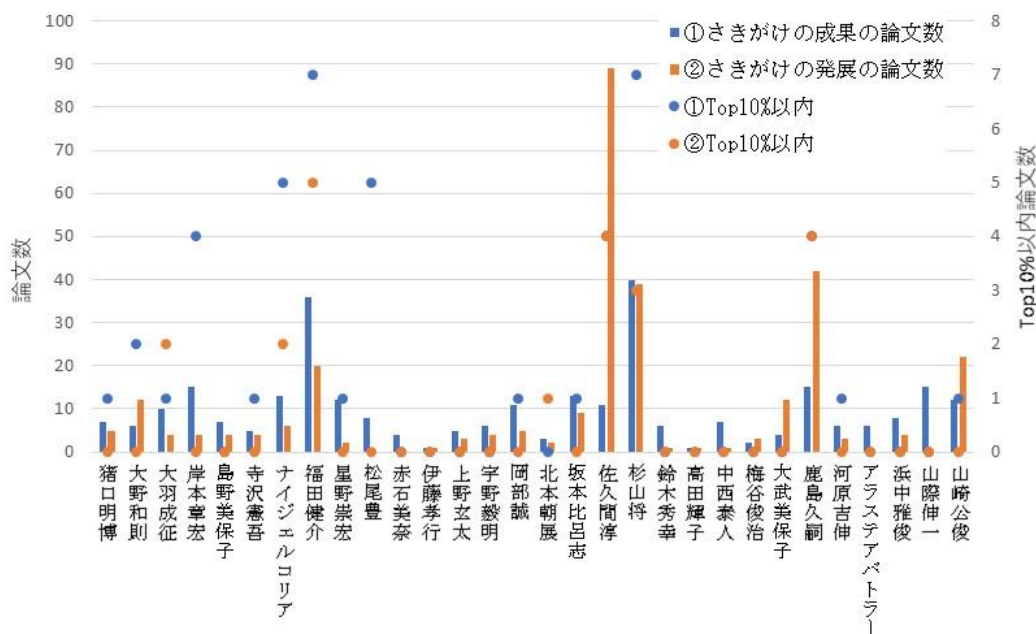


図 2-3 各研究者の研究領域期間中・終了後の論文数

### 2.2.3 特許

特許出願件数・登録件数は研究開発が応用に向けて進展していることを表す一つの指標であると考えられるため、研究期間中と研究終了後の状況について調査し、表 2-6 に示した。

研究期間中の研究者の特許出願は国内 49 件、海外 6 件であった。成立件数(期間中に出願した特許のうち、現時点で特許登録されている件数)は、国内 35 件、海外 6 件であった。期間中では、杉山の特許出願数が 12 件と多く、うち 8 件が登録されている。また、杉山は、海外にも 2 件出願しており、うち全件が登録されている。

研究期間後の特許出願は、国内 80 件、海外 38 件であり、うち国内 28 件、海外 12 件が特許として成立している。研究期間後では、大野と岸本の特許出願数が多い。大野は、国内で 15 件、海外で 1 件を出願し、国内で 7 件、海外で 1 件が登録されている。岸本は、国内では出願していない一方、海外へ 15 件を出願している。これは、岸本の所属が、研究終了後にアイルランド国の IBM に移ったためと思われる。

また、本さきがけ領域の研究者が発明者として含まれ、企業が出願人となっている特許は、出願 105 件であり、57 件(国内 43 件、海外 14 件)が成立している。

表 2-6 研究期間中・終了後の特許の出願と成立状況

採択年度	研究代表者	研究期間中				研究終了後			
		出願件数		登録件数		出願件数		登録件数	
		国内	海外	国内	海外	国内	海外	国内	海外
2008年度	猪口 明博	1	0	0	0	1	0	0	0
	大野 和則	1	0	1	0	15	1	7	1
	大羽 成征	0	0	0	0	0	0	0	0
	岸本 章宏	0	0	0	0	0	26	0	7
	島野 美保子	1	0	0	0	2	1	0	0
	寺沢 憲吾	1	0	0	0	0	0	0	0
	Nigel Collier	0	0	0	0	0	0	0	0
	福田 健介	1	0	0	0	6	0	4	0
	星野 崇宏	0	0	0	0	1	1	0	0
松尾 豊	3	0	2	0	5	1	1	0	
2009年度	赤石 美奈	0	0	0	0	0	0	0	0
	伊藤 孝行	3	0	1	0	7	1	0	0
	上野 玄太	3	0	3	0	0	0	0	0
	宇野 毅明	0	0	0	0	0	0	0	0
	岡部 誠	0	0	0	0	0	0	0	0
	北本 朝展	0	0	0	0	0	0	0	0
	坂本 比呂志	0	0	0	0	2	0	2	0
	佐久間 淳	1	1	1	1	7	1	1	1

	杉山 将	12	2	8	2	8	5	3	3
	鈴木 秀幸	0	0	0	0	0	0	0	0
	高田 輝子	0	0	0	0	0	0	0	0
	中西 泰人	1	0	1	0	0	0	0	0
2010 年度	梅谷 俊治	0	0	0	0	4	0	3	0
	大武 美保子	3	0	2	0	5	0	1	0
	鹿島 久嗣	0	0	0	0	0	0	0	0
	河原 吉伸	4	2	4	2	7	1	3	0
	Alastair Butler	0	0	0	0	0	0	0	0
	浜中 雅俊	2	0	1	0	1	0	1	0
	山際 伸一	4	0	4	0	2	0	2	0
	山崎 公俊	8	1	7	1	7	0	0	0
	領域全体	49	6	35	6	80	38	28	12

- 1) PCT 出願、海外国への個別特許申請のいずれかがあれば、海外としてカウント。
  - 2) 国内特許出願し PCT 出願あるいは直接 PCT 出願された場合は国内出願件数に含めてカウント。
- 2020年10月28日調査

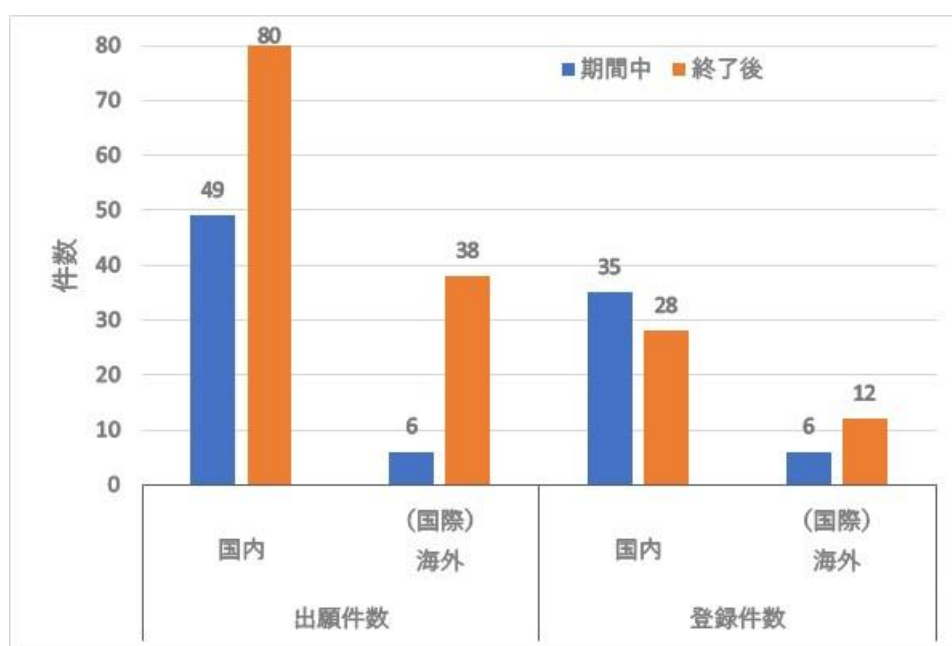


図 2-4 研究期間中・終了後の特許の出願と成立状況

#### 2.2.4 受賞

科学技術の進歩への貢献や研究成果に関する評価を示す指標の一つとして、受賞歴が挙げられる。研究領域終了後の受賞は93件あり、表 2-7 に、各研究者の受賞歴を示す。若手研究者に対する著名な賞である、文部科学大臣表彰若手科学者賞を、岸本が2012年

に、杉山が2014年に、山崎が2016年に、松尾が2018年に、河原が2020年に受賞した。また、伊藤は2014年に、星野と杉山は2017年に、鹿島は2019年に日本学術振興会賞を受賞している。本研究領域の研究者は、その他、財団などの様々な賞や、学会賞などを受賞している。

表 2-7 研究終了後の受賞リスト

No.	受賞者	賞の名称	授与機関	受賞年
1	大野 和則	競基弘賞特別賞	国際レスキューシステム研究機構	2012
2		教育賞	日本機械学会	2013
3		ROBOMECH 表彰	一般社団法人 日本機械学会 ロボティクス・メカトロニクス部門	2017
4		Best Paper Award Third Prize	IEEE Robotics&Automation Society	2018
5		優秀ポスター賞	建設ロボット研究連絡協議会	2019
6		Best Paper Award on Mechanisms and Design	IEEE Robotics&Automation Society	2019
7		第1回優秀研究・技術賞	日本ロボット学会	2020
8	岸本 章宏	科学技術分野の文部科学大臣表彰若手科学者賞	文部科学省	2012
9	寺沢 憲吾	第66回電子情報通信学会論文賞	電子情報通信学会	2010
10	星野 崇宏	2016年度日本学術振興会賞	日本学術振興会	2017
11		2017年日本統計学会研究業績賞	日本統計学会	2017
12	松尾 豊	科学技術への顕著な貢献2015(ナイスステップな研究者)	文部科学省 科学技術・学術政策研究所(NISTEP)	2015
13		科学技術分野の文部科学大臣表彰若手科学者賞	文部科学省	2018
14		2018年度情報処理学会論文賞	情報処理学会	2019
15	伊藤 孝行	最終予選出場	国際自動交渉エージェント競技会	2012
16		科学技術分野の文部科学大臣表彰科学技術賞(研究部門)	文部科学省	2013
17		基礎研究賞	日本ソフトウェア科学会	2014
18		日本学術振興会賞	日本学術振興会	2014
19		Best Paper Award	ICCSE	2016
20		業績賞	人工知能学会	2016
21		情報処理学会フェロー認定	情報処理学会	2018
22		The Best Video Award	IJCAI	2019
23		Best Paper Award	Pacific International Conference on Artificial Intelligence	2019
24		Best Paper Award	IEEE/ACIS	2019
25		Best Paper Award	IJCAI	2019
26	宇野 毅明	研究会優秀賞	人工知能学会	2013
27		山下記念研究賞	情報処理学会	2014



No.	受賞者	賞の名称	授与機関	受賞年
28		I-Scover チャレンジ 2014 最優秀賞	電子情報通信学会	2015
29	岡部 誠	対話発表賞	日本ソフトウェア科学会	2014
30		画像電子学会 VC 賞	Computing (VC) 研究会/グラフィクスと CAD (GCAD) 研究会	2014
31		優秀研究発表賞	Computing (VC) 研究会/グラフィクスと CAD (GCAD) 研究会	2014
32		優秀ポスター賞	画像関連学会連合会	2015
33		インタラクティブ発表賞	電子情報通信学会パターン認識・メディア理解 (PRMU) 研究専門委員会	2015
34		インタラクティブ発表賞	電子情報通信学会パターン認識・メディア理解 (PRMU) 研究専門委員会	2015
35		VC 賞	画像電子学会	2016
36		グラフィクスと CAD (GCAD) 研究会 優秀研究発表賞	情報処理学会	2016
37	北本 朝展	Geo アクティビティフェスタ奨励賞	国土交通省 (国土地理院、国土政策局)	2013
38		Linked Open Data チャレンジ Japan 2013 基盤技術部門 最優秀賞	LOD チャレンジ実行委員会	2014
39	坂本 比呂志	人工知能学会研究会優秀賞	人工知能学会	2014
40		Best paper award at VLDB2015 workshop: BPOE-6	BPOE	2015
41		大学発ベンチャー表彰	科学技術振興機構	2018
42	佐久間 淳	佐久間淳, 第 5 回データ工学と情報マネジメントに関するフォーラム (第 11 回年次大会) 日本データベース学会上林奨励賞	日本データベース学会	2013
43		研究会賞	電子情報通信学会	2014
44		コンセプト論文賞	情報処理学会	2015
45		最優秀論文賞	日本データベース学会	2015
46		最優秀論文賞	情報処理学会	2016
47		優秀論文賞	情報処理学会	2016
48		最優秀論文賞	情報処理学会	2017
49		ベストオーサー賞	電子情報通信学会	2017
50		優秀論文賞	情報処理学会	2017
51		学生優秀論文賞	情報処理学会	2017
52		学生論文賞	情報処理学会	2017
53	杉山 将	最優秀論文賞	Conference on Technologies and Applications of Artificial Intelligence (TAAI2013)	2013
54		優秀賞	第 11 回情報学ワークショップ (WiNF2013)	2013

No.	受賞者	賞の名称	授与機関	受賞年	
55		優秀論文賞	Seventeenth International Conference on Artificial Intelligence and Statistics (AISTATS2014)	2014	
56		科学技術分野の文部科学大臣表彰若手科学者賞	文部科学省	2014	
57		IBISML 研究会賞ファイナリスト	電子情報通信学会 IBISML 研究会	2015	
58		優秀論文賞	第 20 回ロボティクスシンポジア	2015	
59		最優秀論文賞次点	The 9th Asian Conference on Machine Learning (ACML2017)	2017	
60		日本学士院学術奨励賞	日本学士院	2017	
61		日本学術振興会賞	日本学術振興会	2017	
62		最優秀ポスター発表賞	The First Japan-Israel Machine Learning Meeting (JIML2018)	2018	
63		IBISML 研究会賞	電子情報通信学会 IBISML 研究会	2018	
64		IBISML 研究会賞	電子情報通信学会 IBISML 研究会	2018	
65		フェロー	電子情報通信学会	2018	
66		最優秀論文賞	ACML2019 Weakly-supervised Learning Workshop	2019	
67		KDDI Foundation Awards	KDDI 財団	2019	
68		大武 美保子	現場イノベーション賞	人工知能学会	2016
69		鹿島 久嗣	人工知能学会論文賞	人工知能学会	2015
70	日本学術振興会賞		日本学術振興会	2019	
71	河原 吉伸	Outstanding Reviewer Award	NIPS 2014	2014	
72		科学技術分野の文部科学大臣表彰若手科学者賞	文部科学省	2020	
73	浜中 雅俊	IJCAI-19 (the 28th International Joint Conference on Artificial Intelligence), Most Entertaining Video Award	IJCAI (International Joint Conference on Artificial Intelligence Organization)	2019	
74		SIGGRAPH2019 Emerging Technologies Laval Virtual Revolution Research Jury Prize	Laval Virtual	2019	
75		羽倉賞 奨励賞	一般財団法人 最先端表現技術利用推進協会	2019	
76		第 24 回情報処理学会シンポジウム インタラクシオン 2020 インタラクティブ発表賞	情報処理学会	2020	
77	山際 伸一	Embedded Technology Award 2015 「特別賞」	一般社団法人 組込みシステム技術協会	2015	
78		Best Paper Award	VLDB 2015/ BPOE-6	2015	
79		常陽ビジネスアワード「最優秀賞」(成長事業部門)	株式会社常陽銀行	2016	
80		TXアジア・アントレプレナーシップ・アワード 2017 国内代表選出	アジア・アントレプレナーシップ・アワード運営委員会	2017	

No.	受賞者	賞の名称	授与機関	受賞年
81		大学発ベンチャー表彰「科学技術振興機構長賞」	JST・NEDO	2018
82		第5回三菱UFJ銀行 Rise Up Festa「優秀賞」	三菱UFJ銀行	2018
83		IoT推進ラボ 第5回 IoT Lab Selection ファイナリスト選出	経済産業省、IoT推進ラボ	2018
84		MITANNI Business Contest2019「優秀賞」	三谷産業株式会社	2019
85	山崎 公俊	計測自動制御学会論文賞	計測自動制御学会	2014
86		日本機械学会奨励賞(研究)	日本機械学会	2015
87		科学技術分野の文部科学大臣表彰若手科学者賞	文部科学省	2016
88		GaiTech Best Paper in Robotics	IEEE ICIA	2017
89		Best paper award finalist	IEEE ICIA	2017
90		Best paper award finalist	IEEE ICIA	2017
91		ものづくりカテゴリー第2位(NEDO 理事長賞), World Robot Summit2018	NEDO	2018
92		Best paper award finalist	IEEE Robotics&Automation Society	2018
93		The Most Influential Paper over the Decade Award	IAPR	2019

### 2.2.5 招待講演

研究者の研究成果を、国際学会における招待講演として発表した件数が、研究終了後、合計 497 件に上った。特に、杉山が 222 件、伊藤が 77 件、大野が 49 件と多い。

杉山は、機械学習に関する講演が多く、国内外の様々な学会やシンポジウムで講演を実施している。また日本の企業での講演も数多く確認された。伊藤は、自動交渉技術、合意意見形成支援技術、マルチエージェントに関する講演が多く、大野はロボット技術に関連する講演が多い。

また特に伊藤、高田、浜中らは海外での講演実績が多く確認された。

### 2.2.6 報道

研究終了後に報道機関から報じられた件数は、総数が 628 件に上った。研究者別では、杉山が 214 件で最も多く、研究成果が社会的にも注目されていることを示している。次いで、松尾が 149 件、大野が 65 件となっている。

### 2.2.7 共同研究や企業との連携

本研究領域では、研究者自身のネットワークやさがけを通じてのつながり、国際共同プログラムへの応募などを通じて国内外の大学や企業と共同研究が数多く行われている。特に活発に共同研究を行っている事例として大武は千葉大学、京都大学、神戸大学、筑波大学、東北大学、国立国語研究所、産業技術総合研究所などと共同研究を実施した。海外

大学ではチューリッヒ大学と共同研究を実施し、また健康な加齢に関するワークショップをチューリッヒ大学、ETH、スイス大使館との共催で開催した。

またさきがけに参加した研究者同士の共同研究の事例として梅谷と河原らによる劣モジュラ関数最大化問題に対する分枝カット法の開発の共同研究が挙げられる。企業との共同研究を活発におこなった研究者としては山崎が挙げられ、トヨタ自動車株式会社、セイコーエプソン株式会社、富士紡ホールディングス株式会社、ワイエイシイマシナリ株式会社などと共同研究を実施した。また山崎は NEDO 次世代ロボット中核技術開発におけるプロジェクトを通じて数多くの大学と共同で研究を実施した。ここで挙げた研究者以外にも寺沢の国史館台湾文献館、台湾国立政治大学と連携した台湾総督府所蔵の日本近代公文書自動解読システム開発の研究や Bulter の Keyaki Treebank に関する NTT との共同研究など、多くの研究者が様々な大学や企業と活発に共同研究をしていることが確認できた。

## 2.2.8 実用化・製品化

本研究領域は、知の創生と情報社会であり、IT などと関連した事例が数多く確認された。例えば伊藤が開発した自動交渉技術を用いた大規模合意形成支援システムは名古屋市における次期総合計画策定において使用され、さきがけの研究の発展型のシステム D-agree も名古屋市の意見集約に利用された。また宇野が開発した類似性解析・クラスタリングアルゴリズムは愛媛県の婚活支援事業におけるマッチングレコメンデーションエンジンや株式会社 DAC のターゲティングアルゴリズムなどに用いられた。他にも岡部が IT 企業へ静止画に動きを付与する動画作成支援システムのノウハウとソフトウェアの提供した実績などもある。

さらに実用化への貢献として北本は、東日本大震災のデジタルアーカイブを 10 年にわたって拡充し続けており、基礎データとして活用され、福田もネットワークトラフィック中の異常をラベリングしたデータ (MAWILab) を公開し、当該分野のベンチマークデータとして活用されている。

また大野は現在自動運転の為の知能化ソフトを 2022 年までに公開することを目指して開発を実施しており、梅谷は開発した大規模な 0-1 整数計画問題に対する効率的な局所探索法が NTT データ数理システムの汎用数理計画法パッケージの新機能として実装される予定であるなど今後も本分野における研究成果が実用化される例は数多く出てくると考えられる。

## 2.2.9 ベンチャー

### 2.2.9.1 株式会社 Gunosy

松尾研究室の関喜史氏が創業者の 1 人となり起業されたベンチャー企業である。松尾は 2015 年からアドバイザーボードに着任している。

#### (1) (株) Gunosy 社設立の状況等について

##### ① 起業の理由・動機・狙い等

「情報を世界中の人に最適に届ける」を企業理念に掲げ、人工知能や機械学習を用いて適切な情報を届けるサービスの提供を目指して起業された。

#### ②起業の原資となった技術(登録特許、論文等)

公開情報からは情報が得られなかった。

### (2)会社の概要

#### ①基本データ(売上高、資本金(主たる株主と出資額)、従業員数等)

資本金：4077 百万円

主な株主：木村新司 5536 株(保有割合 23.39%)、KDDI 株式会社 3550 株 (保有割合 14.99%) 株式会社日本カストディ銀行(信託口) 3111 株 (保有割合：13.14%)

売上高：13987 百万円 (2020 年度)

#### ②事業の概要

##### (i)事業の内容

主に情報キュレーションサービスを通じて、インターネット上に存在する膨大な量の情報の中から、特定の基準に基づき情報を収集し配信するサービスを提供している。特に情報の収集・整理をアルゴリズムなどのテクノロジーを用いて、ユーザの必要とする情報を届けている。

##### (ii)事業の計画

サービスラインナップの拡充を順次行い、スマートフォンにおける情報のプラットフォームとしての地位確立に向け、アルゴリズム開発の技術やメディア運営のノウハウを生かし社会的な課題を解決できる独自のサービスの創出に取り組む。

##### (iii)製造の状況、量産の可能性等

公開情報からは情報が得られなかった。

#### ③市場開拓戦略・製品戦略について

公開情報からは情報が得られなかった。

#### ④その他

松尾研究室の関喜史氏が創業者の1人となり起業されたベンチャー企業である。また、アドバイザリーボードに慶応大学の遠藤典子氏、嘉悦大学の小野展克氏、経営共創基盤取締役 塩野誠氏とともに、松尾が着任している。

### 2.2.9.2 株式会社 PKSHA Technology

松尾研究室の上野山勝也氏が株式会社 PKSHA Technology をアルゴリズムライセンス事業の企業として2012年10月に起業した。

#### (1)(株)PKSHA Technology 社設立の状況等について

##### ①起業の理由・動機・狙い等

アルゴリズム領域の最先端技術の開発からその応用までをワンストップで行うことを目的として、アルゴリズムの研究を行う技術者・研究者によって機械学習技術を用いたデータ解析事業の企業として起業された。

## ②起業の原資となった技術(登録特許、論文等)

公開情報からは情報が得られなかった。

## (2)会社の概要

### ①基本データ(売上高、資本金(主たる株主と出資額)、従業員数等)

資本金：1896万円

株主：上野山勝也 957万株(保有割合 31.57%)，山田尚史 356万株(保有割合 11.74%)，株式会社 LUCE Capital 110万株(保有割合 3.65%)

売上高：739千万(2020年9月期)

### ②事業の概要

(i)各種製品の概要(名称、種別、機能、価格、生産量等)

#### a CELLOR

機械学習技術を用いた CRM ソリューションである。従来多くの時間やコストのかかっていたデータ分析を自動化、または半自動化することで時間やコストの削減につながるだけでなく、分析結果を基にユーザーに広告等を配信することにより優良顧客の離反防止や新規顧客のロイヤル化を行っている。

#### b HRUS

深層学習技術を用いた画像・動画の識別エンジンで主に企業向けのソフトウェアである。今後様々な業界・領域でカメラを中心としたイメージング機器が普及していくと想定されており、それらの様々なイメージング機器と連携して動作し、物体検知や物体認識を実現することでイメージング機器のサービス品質を高めることを目指している。

(ii)事業の計画

深層学習の技術をビジネスとして見たとき、その性能を高める上で目的に沿ったデータを集めることが重要であり、また使えば使うほど精度が向上していくという好循環構造をもつ。この技術特性を正しく理解し、事業成長につながる事業展開の戦略・戦術を採用していくことを目指し技術開発・研究開発・製品化を進めていく。

### ③市場開拓戦略・製品戦略について

公開情報からは情報が得られなかった。

### ④その他

松尾研究室との関係：松尾研究室の学生上野山勝也氏が起業した会社である。また松尾自身も技術顧問を務めている。

## 2.2.9.3 株式会社 DeepX

2016年4月に松尾研究室の那須野薫氏によって人工知能技術を応用して社会課題を解決することを目的に起業された。

### (1)(株)DeepX社設立の状況等について

#### ①起業の理由・動機・狙い等

人工知能の技術を応用して日本の産業に貢献したいという思いから創業された。具体的にはソフトウェア最適化技術で日本特有のモノづくり産業をレバレッジし産業界における課題を解決することを目的とした事業を行い、将来的には日本のみならず世界を社会的・経済的に変革することを目指す。

#### ②起業の原資となった技術(登録特許、論文等)

公開情報からは情報が得られなかった。

### (2)会社の概要

#### ①基本データ(売上高、資本金(主たる株主と出資額)、従業員数等)

資本金：9.1億円(資本準備金含む)

株主：那須野薫、富山翔司、初谷玲慈

#### ②事業事例

##### (i)土木建設×AI

建設現場でのさまざまな土木作業の自動化に向け、ゼネコンの株式会社フジタと連携し汎用重機である油圧ショベルの自動化を目指している。最先端のディープラーニングによる画像認識技術や強化学習の技術を駆使し、油圧ショベルによる切削作業の自動化を実現した。

##### (ii)食品加工×AI

従来食品は非定型性や軟性、粘性などの影響で機械では扱いが難しく、掴み上げ作業や盛り付け作業の自動化は困難だったが現在自動化機器メーカーの株式会社イシダと提携し盛り付け機器のアームをAIによって制御し自動化する技術を開発している。

#### ③市場開拓戦略・製品戦略について

公開情報からは情報が得られなかった。

#### ④その他

松尾研究室との関係：松尾研究室の学生那須野薫氏が起業した会社である。また松尾自身も技術顧問を務めている。

### 2.2.9.4 READYFOR 株式会社

松尾研究室の共同研究先、慶應義塾大学藤田ゼミの米良はるか氏が起業したベンチャー企業である。元々は松尾が代表取締役を務めていたオーマ株式会社の一事業であったが2014年に独立し、株式会社となった。

#### (1)READYFOR(株)設立の状況等について

##### ①起業の理由・動機・狙い等

日本初のクラウドファンディングサービスである。元々はオーマ株式会社の一事業であったが2014年に独立し株式会社となった。インターネットを通じた資金調達システムを日本に導入し、社会をより良くすることを目的に創業された。

##### ②起業の原資となった技術(登録特許、論文等)

公開情報からは情報が得られなかった。

## (2)会社の概要

### ①基本データ(売上高、資本金(主たる株主と出資額)、従業員数等)

資本金：6.3億円(資本剰余金含む)

代表者：米良はるか、樋浦直樹

### ②事業の概要

#### (i)クラウドファンディング事業

日本初・日本最大級のクラウドファンディングサービスである。2011年3月にサービスを開始して以来1万件以上のプロジェクトを掲載し100億円以上の資金が集まったという実績があり、またキュレーターという担当者が各プロジェクトにつきマンツーマンでサポートすることでプロジェクトの達成率は業界最高基準の75%を誇る。

#### (ii)法人向け事業

クラウドファンディングサポートから得られたノウハウを活用し、企業の寄付活動・SDGs推進をサポートしている。具体的には企業がパートナーシップを組む際の安心・安全な団体の紹介や、オンラインで気軽に参加できる従業員寄付、マッチングギフト、事業者ポイントを活用した寄付などの支援を行っている。

### ③市場開拓戦略・製品戦略について

公開情報からは情報が得られなかった。

### ④その他

松尾研究室の共同研究先の藤田ゼミの学生であった米良はるか氏が起業した会社である。元は松尾自身が代表取締役を務めるオーマ株式会社の一事業だったが2014年に独立した。

## 2.2.9.5 株式会社 ACES

「ヒトの知見を数式化する」会社として起業された。機械の眼とも呼ばれる画像映像認識AIアルゴリズムを用いて、人の眼を通して認識されていたリアルな情報を自動でデジタル化し、人の働き方を自動化・効率化することを目指す。

### (1)(株)ACES社設立の状況等について

#### ①起業の理由・動機・狙い等

松尾研究室に所属していた田中浩一郎氏を含む6人のメンバによってディープラーニングに特化した会社として起業された。6人のうち3人が松尾研究室に所属している。ビジネス課題を踏まえたAI導入の要件定義からデータの収集、モデルの開発・検証、システムの導入に至るまでの一連のフローをサポートしAIの社会実装を目指す。

#### ②起業の原資となった技術(登録特許、論文等)

公開情報からは情報が得られなかった。

## (2)会社の概要

### ①基本データ(売上高、資本金(主たる株主と出資額)、従業員数等)

資本金：1億7853万円

従業員数：21名



主なクライアント：電通、テレビ東京、損保ホールディングス、JFE エンジニアリング等

## ②事業事例

### (i) SHARON

リアル産業をデジタル化する為に必要な最先端の動画認識・解析のアルゴリズムを API/SDK として提供するサービスである。特に3つの点を強みとしており1つ目は AI に関する専門知識がなくとも導入可能である点、2つ目は導入前の検証が可能である点、そして3つ目は学習データなしで導入可能である点である。具体的には身体動作の解析や心の動きの可視化、健康状態の管理などの自動化を可能とし、運転中の眠気検知や顔モザイク処理などに利用されている。

### (ii) スポーツにおける身体情報のデジタル化

最先端の 2D/3D Pose Estimation アルゴリズムをはじめとする Deep Learning 技術によっていままで「センス」とされていたスポーツにおける身体行動をデジタル化する。バイオメカニクスやスポーツ科学の知見も取り入れることで、ケガ防止や特徴把握パフォーマンスの向上を目指している。現在国内プロ野球球団に対してアルゴリズムの導入を行っている。

## ③市場開拓戦略・製品戦略について

公開情報からは情報が得られなかった。

## ④その他

松尾研究室に所属していた田中浩一郎氏が CEO を務める。また創業メンバ6人のうち田中浩一郎氏を含む3人が松尾研究室所属である。松尾は技術顧問を務める。

### 2.2.9.6 Bestat 株式会社

人が行う単純作業を AI が代わりにこなすことで、人が人らしく生きられる世界を目指して松尾研究室に所属の松田尚子氏によって起業された。特に画像 AI に重きをおいて事業を展開している。

#### (1) bestat (株) 設立の状況等について

##### ① 起業の理由・動機・狙い等

人が人らしく生きられる世界を画像 AI の技術を用いて創りあげることがを目的に松田尚子氏によって起業された。画像 AI は先端技術であり、いまだに何が可能なのかを正しく理解している人は限られている。そこで同社は画像 AI を組み込んだシステムとアプリケーションの開発を行い様々な空間でリモートサービスを提供するシステムを扱う事業を取り扱っている。

##### ② 起業の原資となった技術(登録特許、論文等)

公開情報からは情報が得られなかった。

#### (2) 会社の概要

##### ① 基本データ(売上高、資本金(主たる株主と出資額)、従業員数等)

社員数：10人以下

## ②事業の概要

画像 AI エンジン開発・提供、画像 AI の SaaS とデバイスの提供、監視カメラ用画像 AI の開発・提供

## ③市場開拓戦略・製品戦略について

公開情報からは情報が得られなかった。

## ④その他

松尾研究室所属の松田尚子氏が創業した。また松尾自身も同社の技術顧問を務める。

### 2.2.9.7 株式会社 ELYZA

松尾は、2018年9月に松尾研究室出身のメンバによって設立された(株)ELYZA社(CEO、曾根岡侑也氏)において2020年10月から技術顧問を務め、AI領域における研究開発や事業開発に関する助言を行っている。

#### (1) (株)ELYZA 社設立の状況等について

##### ①起業の理由・動機・狙い等

松尾研究室発の Deep Learning を軸に、未踏の問題の解決を目指し、技術領域では自然言語処理、社会実装では小売の需要予測に焦点を当て、現在の Deep Learning の先にあるイノベーションに向けて取り組むことを目指し、設立された。

##### ②起業の原資となった技術(登録特許、論文等)

公開情報からは情報が得られなかった。

#### (2) 会社の概要

##### ①基本データ(売上高、資本金(主たる株主と出資額)、従業員数等)

公開情報からは情報が得られなかった。

##### ②事業の概要

###### (i) 各種製品の概要(名称、種別、機能、価格、生産量等)

ELYZA Brain は、同社が開発した、NLP(自然言語処理技術)領域における国内最大の日本語 AI エンジンである。国内最大の日本語テキストデータの学習量及び国内最大のモデルの大きさを有しており、特定の技術課題ではネイティブな日本語話者の精度を超える。

###### (ii) 競合する会社と競合他社に対する強み

公開情報からは情報が得られなかった。

##### ③市場開拓戦略・製品戦略について

###### (i) 市場開拓、顧客開拓について

2020年9月から、ELYZA Brain を活用した企業との事業共創プログラム、「パートナープログラム NLP30」を開始(検証フェーズの開発リソースを無償提供)。ビジネスパートナーを募集し、テキストデータを多く持ち、分析による業務自動化や新規ビジネス創出の意欲があるパートナーとの協業を通じて、ELYZA Brain の販売につなげる。

###### (ii) 新製品開発の概要

公開情報からは情報が得られなかった。

##### ④その他

松尾研究室に伴走し、大学を中心にイノベーションを生み出す「エコシステム」を作り、大きく発展させることを目的に設立された株式会社松尾研究所と密に連携。2020年10月からは松尾が技術顧問を務め、AI 領域における研究開発や事業開発に関して助言を行い、事業拡大の加速化を図る。法律事務所と契約書の領域で共同研究を行っている。

#### 2.2.9.8 株式会社 O11o

松尾は、東京大学松尾研究室のメンバを中心に 2019 年 2 月に設立された(株)O11o 社(代表取締役 CEO、川合健斗氏)において、自身の研究成果を商品化している。

##### (1) (株)O11o 社設立の状況等について

###### ①起業の理由・動機・狙い等

Deep Learning 技術を活用し、世界最高レベルの顔認証など様々な画像認識をエッジデバイスで実現することを目指し、設立された。

###### ②起業の原資となった技術(登録特許、論文等)

公開情報からは情報が得られなかった。

##### (2) 会社の概要

###### ①基本データ(売上高、資本金(主たる株主と出資額)、従業員数等)

資本金：800 万円

株主：Deep30 有限責任事業組合、株式会社経営共創基盤※

※2020 年 2 月、第三者割当増資を引受

従業員数：15 名

###### ②事業の概要

###### (i) 各種製品の概要(名称、種別、機能、価格、生産量等)

OLLO BOX 製造作業の自動分析や従業員の出勤管理ができる、追加学習不要の画像認識エッジデバイス(価格、生産量不明)。

###### (ii) 競合する会社と競合他社に対する強み

公開情報からは情報が得られなかった。

###### ③市場開拓戦略・製品戦略について

公開情報からは情報が得られなかった。

###### ④その他

東京大学松尾研究室発のベンチャー企業。2019 年 11 月、株式会社 IGPI ビジネスアナリティクス&インテリジェンスと共同研究を実施。最先端のディープラーニング技術を用い、従来の画像認識技術では困難であったフィルムで包装されたパッケージの外観検査技術を開発し、アルビオンの工場に導入。

2021 年 1 月、住友電気工業株式会社と株式会社 O11o は、O11o 独自のエッジ AI 技術を活用した工場の生産性向上に向け、研究開発を共同で実施することに合意。2020 年度末までに、住友電工グループの複数の生産拠点において実証実験を開始し、2021 年度から製造現場の DX 基盤技術として運用展開することを目指す。

## 2.2.9.9 Agreebit 株式会社

伊藤は2019年3月に設立された Agreebit(株) (代表取締役社長/CEO、桑原英人、取締役/CTO、伊藤孝行)において、自身の研究成果を商品化している。

### (1) Agreebit(株) 設立の状況等について

#### ① 起業の理由・動機・狙い等

さきがけ「マルチエージェントの交渉と協調に基づく集合的コラボレーション支援システムの開発」の成果である AI エージェント技術を、持続可能な世界実現のために社会に提供し、世界の課題、国の課題、地域の課題、企業の課題などの解決に貢献することを目標として設立。

#### ②起業の原資となった技術(登録特許、論文等)

特願 2018-148665(伊藤孝行、白松俊、合意形成支援装置および合意形成支援装置用のプログラム、2018. Aug. 07、は国際 PCT にも出願)

### (2) 会社の概要

#### ①基本データ(売上高、資本金(主たる株主と出資額)、従業員数等)

公開情報からは情報が得られなかった。

#### ②事業の概要

##### (i) 各種製品の概要(名称、種別、機能、価格、生産量等)

D-Agree AI によるファシリテーション支援を実現した、インターネット上の議論・合意形成支援システム。「いつでも」「どこでも」「何人でも」気軽に参加でき、行政課題、企業課題、組織課題の合意形成を達成可能(価格、生産量不明)。

##### (ii) 競合する会社と競合他社に対する強み

公開情報からは情報が得られなかった。

#### ③市場開拓戦略・製品戦略について

公開情報からは情報が得られなかった。

#### ④その他

名古屋市の次期総合計画策定における市民からの意見集約に2度使われている。1つは、COLLAGREE システムで、これは、さきがけ期間終了後に作ったものである。内閣府の最先端次世代研究支援プログラムに採択され、その専任義務により、さきがけは中断したが、COLLAGREE を実装し、名古屋市と共同プロジェクトとして継続した。この成果に基づき JST CREST に採択され、JST CREST で開発した、発展型のシステム D-agree も名古屋市の意見集約に利用された。

2020年6月、株式会社 HYPER CUBE、株式会社 HackCamp と共に3社の経営資源を統合し、AI ファシリテータを活用した完全オンラインによるファシリテーション&コンサルティングサービス「サイバーコンサルティング」を開始。

2021年1月、経済産業省が実施する「J-Startup プログラム」の地域版である「J-Startup CENTRAL」参加スタートアップに選出された。

## 2.2.9.10 ストリームテクノロジー株式会社

山際は 2015 年 8 月に設立したストリームテクノロジー(株)において、自身の研究成果を商品化している。

### (1)ストリームテクノロジー(株)設立の状況等について

#### ①起業の理由・動機・狙い等

JST さきがけ領域会議にて、データ圧縮理論を専門とする坂本との出会いが当社技術のブレークスルーの端緒となった。JST や産業界のつながり、イノベーション・ジャパンへの出展といった機会を利用し、知財を活用する企業を活発に探索したが、当時は景気の上昇初期段階にあり、思い通りの展開を確約する協業先が見つからなかった。タイミングを失いかねないという思いから、さきがけ「高性能ストリーム・コンピューティング環境の構築」の成果である LCA-DLT 技術を社会に還元し、ハードウェア、ソフトウェアの高い実装能力の市場提供を目標に創業した。

#### ②起業の原資となった技術(登録特許、論文等)

(i)特許 6256883、特許 6168595

(ii)論文 Information (2016), 7(4), 63 LNCS (2016) 9495, 133-146

### (2)会社の概要

#### ①基本データ(売上高、資本金(主たる株主と出資額)、従業員数等)

資本金：300 万円

従業員：3 名

#### ②事業の概要

(i)各種製品の概要(名称、種別、機能、価格、生産量等)

ストリームデータ圧縮評価キット 従来比 1/30 の超低消費電力・IoT 向け小型コンピュータ用圧縮アクセラレータ(価格不明、2019 年 8 月 20 日販売開始)

(ii)競合する会社と競合他社に対する強み

公開情報からは情報が得られなかった。

#### ③市場開拓戦略・製品戦略について

市場開拓、顧客開拓について、大手半導体メーカーや大手電機メーカーとストリームデータ圧縮技術の社会実装に向けて共同研究を実施。

#### ④その他

2018 年 6 月 12 日、株式会社アバールデータ(代表、広光勲氏)は、ストリームテクノロジー(株)の可逆データ圧縮技術を FPGA に実装し、ストリームデータを 1.5GByte/sec のスピードで、データの取り込みと圧縮データを同時にシステムへ転送可能な高速ハードウェア圧縮ボード「APX-LLC01」(販売予定価格 298,000 円(税別))を販売開始。

2016 年常陽ビジネスアワード「グランプリ」、2018 年 IoT Lab Selection ファイナリスト、2018 年大学発ベンチャー表彰「科学技術振興機構理事長賞」受賞等している。

## 2.3 研究成果から生み出された科学技術や社会・経済への波及効果

本研究領域では、2008年度から2010年度にかけて合計30件の研究課題を採択し、「多様で大規模な情報から『知識』を生産・活用するための基盤技術の創出」という戦略目標の下で研究を遂行した。これらの研究は、まだディープラーニングがブームの端緒とともに進められた研究であり、大規模データを処理する技術や統計数理科学を応用した分析・モデル化技術など当時の最先端の研究成果が得られた。本研究期間終了後、他の研究プロジェクトや自治体との連携プロジェクト、民産官学連携研究拠点のプロジェクトへの展開などを通じて、幅広い実社会への応用を見据えた具体的な取り組みが精力的に行われている。また、機械学習が普及し、本研究領域の研究成果が活用される社会を想定した先行的な課題(秘匿情報に対する知識抽出等)への取り組みも実施されている。さらに、関連技術の創出と併せて、今まさに名実ともにディープラーニングの第一人者となった著名な研究者や主な国研の要職に就く研究者など多くの人材を輩出した。本研究領域は、キャリアアップして教授・教授相当になった研究者が14名おり、本研究領域の日本の中心的な研究者を輩出している。例えば、松尾は新しい組織(一般社団法人日本ディープラーニング協会)を設立した他、複数の要職(同協会 理事長、ソフトバンクグループ株式会社 取締役、人工知能学会 理事、情報処理学会 理事等)に就いている。杉山は機械学習研究部会を立ち上げたほか、理化学研究所革新知能統合研究センター長に就いている。

### 2.3.1 研究領域の展開状況(展開図)

展開と発展の展開図として図 2-5に示した。

研究課題は戦略目標を基に、主に「(1)現在の計算機環境を活用した、人工知能の創出」「(2)現在のデータ爆発環境下での情報圧縮やリアルタイムセンシングという情報学的技術の適用」「(3)データ爆発環境下における統計的手法の社会適用」の3つの分類で取り組まれた。

「(1)現在の計算機環境を活用した、人工知能の創出」では、さきがけの研究成果やその発展成果を基に、WEB上(ニュースサイトやSNS等)の情報を抽出・分析する技術が開発され、ネットワーク理論と機械学習を用いたウェブ情報の知識化が行われた。結果、Twitterを分析し、意見集約する等、マーケティングビジネスへの応用等に結びついている。また、確率密度比の直接推定を活用した機械学習分野の解法が開発され、製造プロセス異常診断、部品品質検査等、企業と共同で多数の実用化に結びついている。今後、当分類では、テキストや映像、ウェブ情報の構造化、機械学習の高度化が期待でき、社会的な様々なテーマを対象に分析、予測できるようになり、金融分野、マーケティング分野、製造業、ロボット等の発展に貢献できる。大規模なウェブ情報を中心としたデータを収集し、機械学習を応用することで、今までノウハウとして蓄積されてきた分類や予測などが自動化される。

「(2)現在のデータ爆発環境下での情報圧縮やリアルタイムセンシングという情報学的技術の適用」では、データ圧縮下における情報処理や大量データフローを扱う高性能・コンパクトなハードウェア構成の基盤技術が開発された。その技術は製品化され、ベンチャ

一企業を通じて社会実装されている。また、視覚や触覚センサを活用したロボットによる自律的な環境情報獲得手法の開発が実施された。今後、今まで以上に大量のデータを扱う社会が来ると想定されるが、その大量データの生成・通信・保存に対応可能な圧縮技術への応用や圧縮データのリアルタイム処理技術への応用が期待される。また、実環境での自律ロボット向けセンシング、高次元の測定データを学習した知能による技術発展も期待される。これらの技術は、今後も大量に生成・通信・保存される文書・映像などの大規模データをリアルタイムに・効率的に扱う技術の確立に貢献する。

「(3)データ爆発環境下における統計的手法の社会適用」では、高齢者同士の会話に介入し、認知症予防を支援する技術が開発されている。その他、本研究領域の成果であるグラフ構造を用いたプライバシー保護アルゴリズムは医療分野におけるデータ(医療データ、ゲノムデータ等)のプライバシー保護に関する応用研究に活用されている。今後、情報社会における各種課題(震災情報の分析・活用、都市空間の高度化、感染症伝搬の分析、プライバシー保護、高齢者の認知症予防等)に対して統計的手法が適用されていくと期待され、実社会データから知識を抽出し、社会問題の解決に貢献する。



戦略目標、分類	インプット	アクティビティ/アウトプット	アウトカム (short/mid-term)		アウトカム (long-term) /インパクト																									
			～追跡調査時点	今後予想される展開	今後想定される波及効果																									
<p>戦略目標： 多様で大規模な情報から『知識』を生産・活用するための基盤技術の創出</p> <p>分類： (1)現在の計算機環境を活用した、人工知能の創出 (2)現在のデータ爆発環境下での情報圧縮やリアルタイムセンシングという情報学的技術の適用 (3)データ爆発環境下における統計的手法の社会適用</p>	<p>研究総括： 中島</p> <p>研究代表者 猪口明博 大野和則 大羽成征 岸本章宏 島野美保子 寺沢憲吾 ナイジェル・コリアー 福田健介 星野崇宏 松尾豊 赤石美奈 伊藤孝行 上野玄太 宇野毅明 岡部誠 北本朝展 坂本比呂志 佐久間淳 杉山将 鈴木秀幸 高田輝子 中西泰人 梅谷俊治 大武美保子 鹿島久嗣 河原吉伸 アラステラバトラ 浜中雅俊 山際伸一 山崎公俊 30名</p>	<p>研究成果</p> <p>論文</p> <table border="1"> <tr> <td>①さきがけ研究成果の論文数</td> <td>②さきがけ研究成果の継続発展の論文数</td> </tr> <tr> <td>290 (45)</td> <td>302 (21)</td> </tr> </table> <p>( )の値はTop10%以内論文数</p> <p>特許申請・登録</p> <table border="1"> <thead> <tr> <th></th> <th>期間中</th> <th>終了後</th> </tr> </thead> <tbody> <tr> <td>出願</td> <td>49</td> <td>80</td> </tr> <tr> <td>国内</td> <td>49</td> <td>80</td> </tr> <tr> <td>国際</td> <td>6</td> <td>38</td> </tr> <tr> <td>登録</td> <td>35</td> <td>28</td> </tr> <tr> <td>国内</td> <td>35</td> <td>28</td> </tr> <tr> <td>国際</td> <td>6</td> <td>12</td> </tr> </tbody> </table> <p>ソフトウェア</p> <ul style="list-style-type: none"> <li>自動化のための知能化ソフト (大野)</li> <li>画像特徴量によるくずし字認識システム (寺沢)</li> <li>COLLAGREE、D-agree (伊藤)</li> <li>類似性解析・クラスタリングアルゴリズム (宇野)</li> <li>GeoNLP (北本)</li> </ul>	①さきがけ研究成果の論文数	②さきがけ研究成果の継続発展の論文数	290 (45)	302 (21)		期間中	終了後	出願	49	80	国内	49	80	国際	6	38	登録	35	28	国内	35	28	国際	6	12	<p>科学技術的および社会・経済的な波及効果</p> <p>大規模データを獲得したり処理したりする新しい仕組みを考え、新しい仕組みを実現する。社会を効率化したり現状の問題点を解決したり、人間の知的作業の質や量を向上させたりする。</p> <ul style="list-style-type: none"> <li>ネットワーク理論と機械学習を用いたウェブ情報の構造化・知識化 → Twitterを分析し、意見集約、マーケティングビジネスへ応用</li> <li>密度比推定による大規模・高次元データの知的処理技術の創生 → 製造プロセス異常診断、部品品質検査等企業と共同で多数の応用事例を解決</li> <li>高性能ストリームコンピューティング環境の構築 → データを圧縮したままでの情報処理、リアルタイム圧縮が可能なハードウェアを開発、ベンチャ企業で販売</li> <li>ロボットの視覚・触覚を用いた環境情報獲得手法の開発 → 実世界の電子化されていない情報をロボット技術で自律的に収集し知識を獲得する技術を開発</li> <li>大規模会話データに基づく個別適応型認知活動支援 → 高齢者の会話データを基にした認知症の検出と改善の支援</li> <li>実世界情報ネットワークからのプライバシー保護データマイニング → 発展的に予防医療のリスク予測部分やゲノム疫学に応用</li> </ul> <p>展開</p> <p>科研費、JSTや他機関での研究助成金獲得</p> <p>CREST研究領域「ビッグデータ統合利活用のための次世代基盤技術の創出・体系化」研究課題「自己情報コントロール機構を持つプライバシー保護データ収集・解析基盤の構築と個別化医療・ゲノム疫学への展開」(佐久間) JSPS新学術領域研究「会話支援技術と認知行動療法に基づく主体価値発展支援システムの開発」(大武) など複数の事業に展開</p> <p>若手研究者間で連携した研究者ネットワークの形成</p> <ul style="list-style-type: none"> <li>山際、坂本 (ストリーム型圧縮技術のハードウェア化を共同で研究を進め、ベンチャ企業を創立)</li> </ul> <p>受賞/人材育成等</p> <p>電子情報通信学会論文賞 (寺沢) を含め、計18名</p> <p>AI人材の輩出 杉山 (理化学研究所革新知能統合研究センター長) 松尾 (AIベンチャを多数創出)</p>	<p>実社会への応用を見据えた新しい基盤技術。大規模情報を対象とし、さらに情報を現実世界から取り込むための手法</p> <ul style="list-style-type: none"> <li>テキストや映像、ウェブ情報の構造化、機械学習の高度化により、社会的な様々なテーマを対象に分析、予測できるようになる。 <ul style="list-style-type: none"> <li>金融分野</li> <li>マーケティング分野</li> <li>製造業分野</li> <li>ロボット分野</li> </ul> </li> <li>大量データの生成、通信、保存環境に対応できる圧縮技術、圧縮したままでのリアルタイム処理</li> <li>実環境での自律ロボット向けセンシング、高次元な測定データに基づく知能による高度化技術</li> <li>情報社会における各種問題に対して統計的手法の適用 <ul style="list-style-type: none"> <li>高齢者の認知症予防</li> <li>震災情報の分析、活用</li> <li>都市空間の高度化</li> <li>感染症伝搬の分析</li> <li>プライバシー保護</li> </ul> </li> </ul>	<p>大規模データの収集、処理するための新しい仕組みを実現</p> <ul style="list-style-type: none"> <li>大規模なウェブ情報を中心としたデータの収集し、機械学習を応用することで、ノウハウとして蓄積されてきた分類、予測などが自動化され、例えば、金融市場等の将来予測の高度化によって、消費者リスクの少ない金融商品が実現する可能性がある。その他、製造業においては生産プロセスの効率化が図られ、労働者の就労環境が改善する。</li> <li>大規模データを処理するための革新的な技術</li> <li>今後も大量に発生するデータや保存されている大量の文書や映像などの大規模データをリアルタイムに、効率的に扱うための技術が確立することによって、例えば、情報やサービスへのアクセシビリティが向上し、消費者や事業者が知りたい情報にいつでも簡単にアクセスできるようになる。</li> <li>実社会における多様なデータから知識を抽出するための技術</li> <li>実社会での多様なデータから知識を抽出し、社会の効率化や問題点を解決する。例えば、高齢者の認知症予防が実現し、健康寿命が伸びることで、介護者不足の問題解決に貢献できる。他にも、震災や感染症対策が高度化することで、安心・安全な暮らしの実現に貢献できる。</li> </ul>
①さきがけ研究成果の論文数	②さきがけ研究成果の継続発展の論文数																													
290 (45)	302 (21)																													
	期間中	終了後																												
出願	49	80																												
国内	49	80																												
国際	6	38																												
登録	35	28																												
国内	35	28																												
国際	6	12																												

図 2-5 さきがけ研究領域の展開図



### 2.3.2 研究成果の科学技術の進歩への貢献

本研究領域では、その研究成果を用いて、アルゴリズムの開発など基盤技術の高度化が進んだだけでなく、社会実装を目標とした実証実験が複数の分野(自動運転、防災、情報検索、プライバシー保護、医療・ヘルスケア等)で展開された。以下に、研究成果の科学技術の進歩への貢献について幾つかの事例を示す。

宇野は、本研究領域での研究終了後、類似性を計算する高速なアルゴリズムをもとに、細かい構造を網羅的に抽出できるクラスタリングアルゴリズム、データ研磨アルゴリズムを開発。データ群の中から似たもの同士を探すというアプローチにより、データ群の中で他と大きく異なるデータがどこのグループにも属さない状況も生じるが、一方でデータを同じようなスケール感でクラスタリングすることに成功した。

大野は本研究領域の研究成果を複数の研究分野に応用した。例えば、自動運転車の駐車に関する研究、工場現場等における大型トラックの自動運転に関する研究、サイバー救助犬に関する研究、ドローンを用いたインフラ検査システムに関する研究を実施した。

坂本は巨大なデータベースをすべてインデックス化し、そのデータベースから任意の組み合わせで、似ている論文のペア(あるいはグループ)を見つけ出す技術を開発し、arXiv上の全論文に対して、高速な剽窃検出を実現した。また、圧縮索引技術のプライバシー保護マイニングへの応用研究も実施しており、データがプライバシー情報を含むものであった場合、暗号化したままデータの比較を可能にする技術を開発した。

赤石は統合された知識を分解・再編成することで新しい知識を生み出し、「歴史知識学」を発展させるためのシステムの設計・開発を行った。

伊藤はマルチエージェントシステムの自動交渉技術とその応用について、理論的アルゴリズムの提案と、応用システムへの展開を行なった。その後、それら成果はCREST研究領域「人間と調和した創造的協働を実現する知的情報処理システムの構築」、研究課題「エージェント技術に基づく大規模合意形成支援システムの創成」(2015年度)に受け継がれ、そのシステムの基礎技術に発展している。

大武は話の内容が過去、現在、未来のうちどの部分に相当する話をしているのかを判定する技術を開発した。その後、最近の話をする人とその人の認知機能の高さに相関関係が成り立つことを明らかにし、認知症予防の研究分野に貢献した。その他、ベッドサイド声掛け見守りロボットの開発やデイケア施設での実証実験等を実施した。

CollierはThe SIPHS(Semantic Interpretation of Personal Health MessageS) projectに取り組んでいる。ブログや掲示板等のオープンなオンラインデータを用いて、医療従事者が独自の臨床研究等を実施する際の支援を行う。

佐久間はさきがけで考案したアルゴリズムを医療分野に応用した。他の研究者と共同で、完全順同型暗号を用いてビッグデータを解析するためのアルゴリズムに関する研究を実施するとともに、医療従事者等とともに実証実験を実施。医療データ・ゲノムデータなどのプライバシー保護に関する研究に貢献した。

寺沢は、未翻刻文書画像から文字認識を経由せずに画像解析によって頻出語を抽出する既往の研究に対し、抽出した頻出語に対して重要度を評価することで、重要語の抽出が行

えることを示した。また、文字画像クラスタリング手法を改良することで、その精度が向上することを示した。その研究成果を活かし、文書画像検索システムを開発した。このシステムは後述の通り民間サービスとして展開されているほか、大学共同利用機関法人人間文化研究機構との古典籍に関する共同研究や中京大学と国史館台湾文献館(台湾)、国立政治大学(台湾)との共同研究等で活用され、人文学研究の発展に貢献している。

### 2.3.3 研究成果の社会・経済への貢献

本研究領域が開始された頃、センシング技術やインターネット技術等の高度化により、大規模なデータの取得が容易になりつつあった。そこで当時、本研究領域では「いかに大量のデータの中から知識を抽出するか」「いかに情報社会を作るか」を目指し、計算機、人工知能、情報学的技術、統計的手法の開発等に取り組む研究者が多かった。本研究期間終了後、それら研究成果は複数の分野の発展に貢献しており、マーケティングやプロセス改善、プライバシー保護など分野横断で幅広く活用されているものもあれば、医療、土木、災害等の特定分野で貢献しているものもある。

具体的には、宇野の研究成果やその発展成果によって、ネットニュースやツイッターのコメントを解析し、どんなコメントがどの程度寄せられているかを定量的に把握できる(マーケティング分野への貢献)。また、類似のコミュニティを見つけることができるという特徴を活かし、その研究はインターネット広告や婚活マッチング等のビジネスにも応用されている。大武の研究成果は、高齢者の認知機能改善等、高齢者支援に結びついており、今後の超高齢社会、健康寿命延伸という社会課題の解決に貢献している。現在、大武は「新しいコミュニケーションツールの台頭によって、人と人とのコミュニケーションが変わってきていること」「そして、好ましくない変化も生じてしまっていること」を危惧し、人間の言語能力向上の課題に取り組もうとしている。寺沢の文書画像検索システムはインターネットで公開されているほか、民間企業とライセンス契約し、市販のソフトウェアとして商品化されている。また、凸版印刷株式会社の既存技術と寺沢のシステムを組み合わせ、同社は「くずし字」を高精度でテキストデータ化する光学文字認識技術を開発している。伊藤は名古屋市と社会実験を行い、自動ファシリテーションエージェントを導入した合意形成支援システムを用いて名古屋市次期総合計画の中間案に対する市民の意見集約を行った。大羽は転職支援企業(ミイダス株式会社)の企業研究者に転身し、民間企業に転職することで、自身の研究成果の社会実装を進めている。その他、大野や佐久間のように、本研究領域で得た研究成果を多方面の分野(医療、土木、災害分野、プライバシー保護分野等)に応用している研究者を本研究領域は輩出している。

また、ベンチャー企業の設立やベンチャー企業へのノウハウ・技術移転が本研究領域では多く見られる。例えば、伊藤は、複数の企業に技術移転やノウハウ提供を行うとともに、AgreeBit 株式会社を設立し、議論・合意形成のための D-Agree サービス、ソフトウェアの販売を行っている。坂本は本研究領域の山際が設立したストリームテクノロジー株式会社(大学発ベンチャー表彰 2018 にて科学技術振興機構理事長賞を受賞)と共同研究を実施し、坂本の技術・ノウハウを移転。ストリームテクノロジー株式会社を通じて技術の社会実

装を進めている。松尾は自身の研究室で複数の企業と共同研究を実施するとともに、研究室から多くのベンチャー企業(株式会社 Gunosy、株式会社 PKSHA Technology、株式会社 ELYZA 等)を創出し、自らも技術顧問等の立場でアドバイスをを行っている。株式会社 PKSHA Technology は文部科学省 大学発ベンチャー表彰 2018 文部科学大臣賞を受賞している。杉山は複数のベンチャー企業(AI CROSS 株式会社、スキルアップ AI 株式会社等)の技術顧問に就いている。また、機械学習プロフェッショナルシリーズ(株式会社講談社販売)のシリーズ編者として、AI 技術の普及に貢献している。このように、自身でベンチャー企業を設立、または、技術顧問等のアドバイザーに就任する等、研究成果の社会実装を行う研究者を本研究領域は輩出している。

本研究終了後からビッグデータや AI の大ブレイクが起こり、関連技術が社会に普及してきたことを踏まえると、まさに本研究領域で狙ってきた時代になってきたといえ、さきがけの研究成果はそのような時代の形成に貢献した。

#### 2.3.4 その他の特記すべき事項(新たな展開や分野間融合)

前述の通り、松尾や杉山は新しい組織の立ち上げや要職への就任を通じて、社会の発展に貢献している。

また、本研究領域では研究者間のネットワーク強化が図られた。宇野は「さきがけは単に人と人を繋げるだけでなく、それぞれの個人が他の人から様々な考え方や文化、視点を交換する機会になったため、そういう点で成長する大きな機会になった」と考えている。その効果として、例えば、大野はさきがけ研究領域「情報環境と人」の原田達也(東京大学先端科学技術研究センター)等と CREST 研究領域「ビッグデータ統合利活用のための次世代基盤技術の創出・体系化」、研究課題「膨大なマルチメディアデータの理解・要約・検索基盤の構築」(2014 年度)にて、共同研究を実施している。その他にも、坂本は本研究領域の山際が設立したストリームテクノロジー株式会社と共同研究を実施している。このように、本研究領域がきっかけとなり、研究者のコミュニティが形成され、研究終了後の本研究領域の発展に貢献している。

**大規模グラフ系列からの知識体系化と理解支援手法の開発**

猪口 明博(関西学院大学理工学部 教授)

研究期間 2008 年 10 月～2012 年 3 月

展開している事業:

科研費若手(A)

**さきがけの成果:**

グラフ形式を対象とした、以下の分析手法の開発、およびその応用を行った。

- ・グラフ系列データから頻出する変化を抽出するアルゴリズム GTRACE、FRISMINER の開発
- ・自然言語処理分野における係り受け解析器の間違い動作パターンの抽出



**発展:**

**1. グラフ系列を対象としたクラスタリング手法の開発<sup>1</sup>**

Web ページのリンク関係や人間関係など時間とともに変化するグラフを対象として、時間とともに変化する特徴的なパターンをクラスタとして検出する手法を開発した。この手法の特徴は、ある時刻の点の分布だけでなく、前後の時刻の分布を加味しながらクラスタリングすることで、クラスタの時刻変化(例えば、クラスタの発生、消滅、併合、分裂)を検出することができる。例えば、図 1 に示す点の分布を開発した手法の入力として与える(入力の時点で、点の色情報はない)と、時刻  $t=3, 4$  において、灰色の点(クラスタ 1)内に新たに生じた点(オレンジ色はついていない)をオレンジ色(クラスタ 3)の新しいクラスタの発生を検出できている。従来手法では、このようなクラスタの検出は容易ではない。

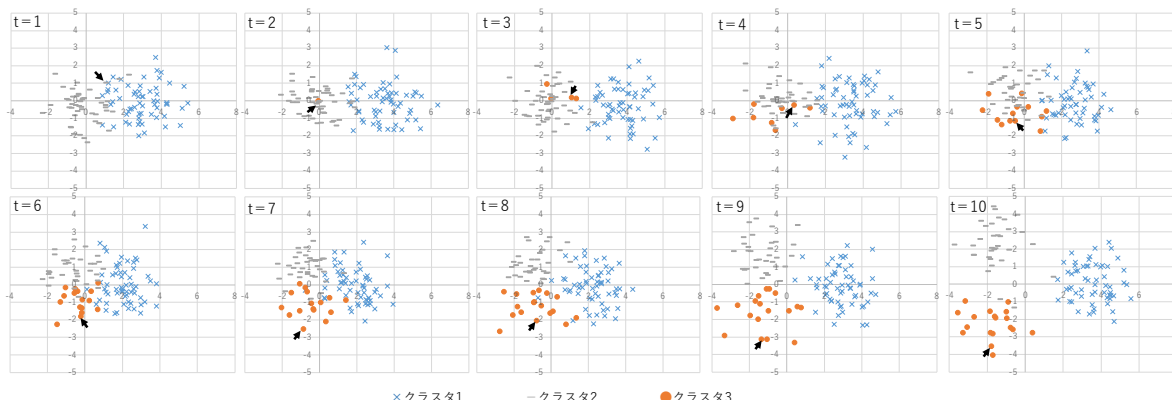


図 1. 検出されたクラスタ系列の頂点の分布

**2. グラフデータを対象としたデータベースの開発<sup>2</sup>**

グラフ集合からなるデータベース  $D=\{g_1, g_2, \dots, g_n\}$  とクエリグラフ  $q$  が入力として与えられたときに、 $q$  に部分グラフとして含まれる  $D$  の要素を高速に検索するデータベース技術を開発した。部分グラフ同型判定問題は NP 完全問題であるため、逐次的な処理では、高速な検索はできない。そこで、 $D$  の要素の部分グラフを事前に抽出し、それと  $q$  との包含関係を調べることで、検索速度の高速化を可能にした。これにより複数の商品を一緒に買った顧客や、人間関係が同じ集団を抽出といったデータマイニングを高速に行える。例えば、図 2 では、(a)の  $g_1$  が(c)の  $q$  の部分グラフであるが、 $g_1$  の部分グラフとして、(b)の  $p_1$  を事前に抽出しておく。これにより、 $q$  が与えられた際には、 $q$  と  $p_1$  の包含関係を調べることで、 $g_2$  や  $g_3$  と  $q$  の包含関係を調べる過程を省くことができ、検索速度が高速になる(ベンチマークテストの内容によっては処理時間が最大 97.5%削減)。

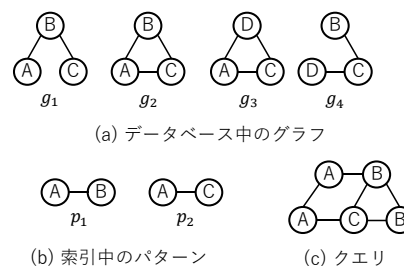


図 2. 検出されたクラスタ系列の頂点の分

<sup>1</sup>. Sohei Okui, Kaho Osamura, Akihiro Inokuchi: Detecting Smooth Cluster Changes in Evolving Graphs. ICMLA 2016: 369-374

長村 佳歩, 奥井 颯平, 猪口 明博: 滑らかな変化を検出するためのグラフ系列クラスタリング情報処理学会論文誌 58(1) 278 - 287

<sup>2</sup> Shun Imai, Akihiro Inokuchi: Efficient Supergraph Search Using Graph Coding. IEICE Trans. Inf. Syst. 103-D(1): 130-141 (2020)

### 3.1.2 ロボットの視覚・触覚を用いた環境情報獲得手法の開発 (大野和則)

#### ロボットの視覚・触覚を用いた環境情報獲得手法の開発

大野 和則(東北大学未来科学技術共同研究センター 准教授)

研究期間 2008年10月～2012年3月

展開している事業:

内閣府 SIP インフラ、科研費若手(A)、科研費基盤(B)

#### さきがけの成果:

ロボットが、現場の撮影データや、触ったり動いたりすることで得られた情報を基に周りの環境を把握することで、周囲に存在する未知の物体の詳細な形状をモデリングする方法の開発を行った。具体的には、①3次元点群計測と認識処理、②押し動作を用いた未知物体のモデリング、③透明物体の認識、④移動台車の整備と高精度位置推定の4つの研究実績より、ロボットの獲得情報を元に現場の状況のより詳細な把握を実現した。



#### 発展:

##### 1. 実世界の電子化されていない情報をロボット技術で自律的に収集し知識を獲得する技術の研究開発<sup>1</sup>

ロボットが自律的に情報収集を行うことが難しいとされる工場内の点検通路などの狭隘路を、移動ロボットが自律的・半自律的に移動する技術を開発し、LiDAR と画像とロボットの動きなどのセンサ情報や、各情報の変化に基づいて周囲の壁や障害物の形状や損傷といった現場の知識を獲得する方法を開発した。また、これにGPS情報を付加することで、観測した壁や障害物を地理的場所の情報に結び付けることができ、地図を作成することが可能となった。

##### 2. サイバー救助犬の研究開発<sup>2</sup>

災害時にレスキュー犬による捜索は非常に有効であるが、一方でレスキュー犬が何を見つけたのかといった情報について人間は犬のリアクションを通じての把握しかできない。そこで本研究では犬の捜索行動や搭載カメラの映像から捜索情報をリアルタイムに可視化する技術、犬の心拍変動から情動を推定する技術、光刺激を利用して犬を長距離ナビゲーションする技術の開発を行った。これによって犬が獲得した一次的な情報をレスキュー隊員も把握できるため、迅速かつ正確な救助へとつなげることが可能となる。

##### 3. 老朽化した橋梁の目視・打音点検を支援するドローンの研究開発<sup>3</sup>

橋梁などのインフラの点検にロボットの活用が期待されている。橋梁点検の重要な仕事の一つに、橋梁の各部の鮮明な画像を取得して、さらに損傷評価を行うことがある。橋梁構造物は複雑で常に衝突などのリスクがあることから、ロボット自身を保護しながら容易に操縦できるシステムが必要であった。そこで保護用の受動回転球殻を有するドローンを開発し、人が容易に近づくことが出来ない高所の損傷の情報を画像から収集する方法を開発した(図1)。本システムは、第三者機関による評価で有効性が確認され、さらに次世代社会基盤ロボットプログラム(NGRSI)の必須要件を満たすため、今後の実用化が見込める。



図1. 目視・打音点検をするドローンロボット

##### 4. 簡易取り付け装置を利用した建機の自動運転<sup>4</sup>

日本では、建設業の人手不足解消のために建設機械の自動化への期待が高まっている。本研究では、従来の建設機械にロボットを後付けすることで自動運転する技術を開発し、建設機械の動きのデータを解析することで、人間と機械が混在する環境において機能する人工知能の開発に取り組んでいる。本研究が発展すれば導入コストの高い建築現場の自動化の障壁を取り除くことができ、地域建設企業の導入障壁を下げるなどの効果や、実世界情報をロボット技術で収集する新たな応用先の開拓が期待される。

<sup>1</sup> 1)IEEE Robotics and Automation Letters (RA-L), 3, 4, 3371-3378, 2018、2)ROBOMECH Journal, 5, 1, 2018、3)JST CREST

<sup>2</sup> 1)Springer Tracts in Advanced Robotics 128, 143-193, 2019、2)IEEE Transaction Transactions on Medical Robotics and Bionics, 1, 3, 189-198, 2019、3)科研費(新学術領域)

<sup>3</sup> 1)Journal of Field Robotics, 35, 6, 850-867, 2018、2)内閣府 SIP

<sup>4</sup> 1)Autonomous Driving of Six-Wheeled Dump Truck with Retrofitted Robot", In Preprints of the 12th International Conference on Field and Service Robotics., pp.11, 2019.、2)NEDO 次世代人工知能・ロボットの中核となるインテグレート技術開発



## 仮説世界と物理世界の相互浸透モデリングによる知の創生

大羽 成征(ミイダス株式会社 HR サイエンス研究所 シニアリサーチャー)

展開している事業:

AMED SICP

研究期間 2008 年 10 月～2014 年 3 月

### さきがけの成果:

仮説世界と物理世界の相互浸透モデリングという観点から、個別仮説ベースの仮説検定とベイズ的な事前知識を融合させることを試みた。理論的成果として相互浸透型の構造を与えることで、検定の検出力を改善できることを示した。また、遺伝子発現量に基づく癌悪性度関連遺伝子の検出力を大幅に改善させることなどに成功した。

### 発展:

#### 1. 非定常環境におけるロバスト適応 BCI<sup>1</sup>

ユーザの脳内状態を読み取ることでコンピュータを操作するブレイン・コンピュータ・インターフェイス(BCI)は従来、ユーザの違いや実験環境自体の変化、同一ユーザでも脳内状態の時間的な変動の影響を受けて、使用中に少しずつ精度が落ちていくという問題があった。そこで本研究では、BCI を非定常な環境へ適用できる技術開発を目的として、大羽らが脳内状態変動対策技術、ドイツベルリン工科大学が実験環境の変動対策技術について共同研究を実施した。ベータダイバージェンスによるフィルタにより脳波の特徴を表すロバストな特徴量の抽出に成功した。また、抽出された特徴量から脳内情報の判別手法として、マルチタスク学習とアンサンブル学習を組合せた手法によって従来の手法を上回る性能を示した。更に、転移学習の仕組みを採用し、非定常環境においてもロバストな性能を示すことが確認できた。

#### 2. 脳の構造と活動の大規模データ解析<sup>2</sup>

本研究は「思考を実現する神経回路機構の解明と人工知能への応用」というプロジェクトの中で実施された。特に大羽は哺乳類の脳を構成する神経回路の構造の同定を全脳レベルで行う高度アルゴリズムを開発することを目的として研究を実施した。本研究では、以下の3つの成果が得られた。初めに、全脳の拡散 MRI 画像に基づく構造同定を実施するアルゴリズムを開発し、実装した。次に、連続切片蛍光顕微鏡画像を前処理することで拡散 MRI 画像と等価な粗視化画像を構築し構造同定を実施(図 1)し、拡散 MRI の同定結果と比較することによりマクロ及びメソスケール構造の信頼性の検証を行った。その結果、拡散 MRI 画像と前処理を実施した連続切片蛍光顕微鏡画像での同定結果がメゾレベルにおいて矛盾しないことを確認できた。最後に、構造同定した結果を本研究が属するプロジェクトの他研究者が実施する全脳ネットワークシミュレーションに提供した。回路構造に着目することで神経回路シミュレータ MONET を開発し、大脳皮質や小脳のシミュレーションを実施。

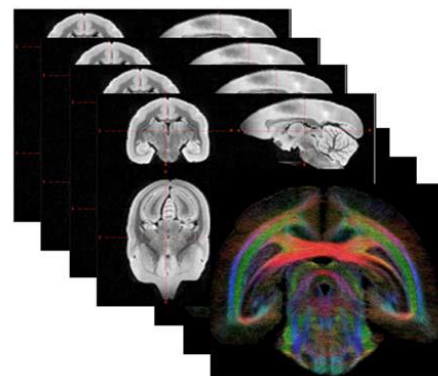


図 1. 拡散 MRI データから構造同定

結果を本研究が属するプロジェクトの他研究者が実施する全脳ネットワークシミュレーションに提供した。回路構造に着目することで神経回路シミュレータ MONET を開発し、大脳皮質や小脳のシミュレーションを実施。

#### 3. 経験的ベイズに基づく神経のスパイク反応の応答関数の推定法

複数のニューロンの機能的な接続の解析はこれまで、ニューロン活動の同時記録を用いて実施されてきた。その際のニューロンのスパイクの応答は一般化線形モデル(GLM)によって推定されてきたが、計測装置の限界やニューロンのペアがニューロンの数に対して2次関数的に増加することなどが原因となり、これまでは有効な推定ができていなかった。本研究では、GLM と経験的ベイズ検定を組み合わせた新しいスパイク応答関数推定法を提案した。この手法によってニューロンのペアの機能的なつながりを高い検出性能で検出でき、更に計算コストを削減できることを示した。本研究の手法をラット海馬の機能的多ニューロンカルシウムイメージングデータセットに適用したところ、AMPA 受容体と NMDA 受容体が介在している可能性があると考えられる有意な機能的接続を発見した。

<sup>1</sup> JST 戦略的国際科学技術協力推進事業評価資料

<sup>2</sup> 文部科学省「思考を実現する神経回路機構の解明と人工知能への応用(脳のビッグデータ解析、全脳シミュレーションと脳型人工知能アーキテクチャ)」成果報告書

<sup>3</sup> BMC Neuroscience volume 17, 27 (2016)

## 大規模並列化によるハイパフォーマンス人工知能技術

岸本 章宏(IBM 東京基礎研究所)

研究期間 2008 年 10 月～2012 年 3 月

### さきがけの成果:

大規模なデータの探索において、効率のよい並列化アルゴリズムを開発した。更に当初研究計画になかった逐次探索アルゴリズムの改良、Binary Decision Diagram (BDD) 構築の並列化も実施した。



### 発展:

#### 1. クラウド環境における、データ駆動型スケジューリング(DDS)に基づく大規模並列探索アルゴリズムの開発<sup>1</sup>

組み合わせ問題を解くメモ化手法<sup>注1</sup>の分散並列化は、メモの効率の良い共有が必要なため、多数の計算機利用による並列化の効果を出しにくい。本研究ではプランニング<sup>注2</sup>と多重シーケンスアライメント<sup>注3</sup>をテストベッドに用いて、クラウド環境での計算機利用コストを最適化しつつ、DDSを大規模並列化できるように拡張し、理論的および実験的な成果を得た。また、本研究で開発したアルゴリズムでは、計算時の必要利用コストの最悪ケースをある値以下に抑えることができることを理論的に示した。さらに、プランニングやバイオインフォマティクスで重要な題材である、与えられた複数の配列の類似領域を特定する多重シーケンスアライメント問題で、上記の理論的な想定上で最適に近い計算機利用コストで並列化できることを実験的にも示した。クラウド環境は社会全体に広まっており、クラウドサービスユーザーは、利用サービス・コストを常に考慮するため本研究の手法は今後着目される方法の1つになると考えられる。

注1: 組合せ最適化問題を解く手法の1つ。表に記録した値を順次更新することで高速に解を得る手法で動的計画法とも密接に関連する。

注2: プランニング言語で記述された問題と初期状態、目標に対し、初期状態から目標を達成するための一連の行為(action)を自動生成すること。

注3: DNAの塩基配列やタンパク質のアミノ酸配列について、複数の配列間で類似する部分が並ぶように整列したもの

#### 2. 並列 SIMPATH アルゴリズムの開発<sup>2</sup>

Knuthによって開発された SIMPATH アルゴリズムは、グラフ上の2点を結ぶ全てのパスを列挙する効率的なアルゴリズムで、例えば、電力ネットワークの送配電コストの最適化にも利用できる重要なアルゴリズムである。本研究では、共有メモリ計算機上で効率の良い並列 SIMPATH アルゴリズム、NBA、RBA、EBAを開発した。この3つの手法について、グリッドグラフの解法によって性能の検証したところ、EBAが最も良い性能を示した。ベストケースでは、32コアを利用した際、逐次アルゴリズムと比べて、約7倍の高速化を実現した。

#### 3. 化学合成計画への証明数探索アルゴリズムの応用<sup>3</sup>

モンテカルロ木探索(MCTS)や証明数探索(PNS)などの探索技術は、ゲームでは有効である一方で産業応用における性能評価は十分に行われていない。本研究では MCTS と PNS の一種である Depth-First Proof-Number(DFPN)探索を、有機化学分野の逆合成解析(Retrosynthetic Analysis)の分野に適用し、その性能を検討した。その結果、DFPN よりも先行研究のアルゴリズムの方が、性能が良いことが判明した。そこで、DFPN の欠点を修正するために、ヒューリスティックエッジ初期化を組み合わせた新しいアルゴリズム DFPN-E を開発した。実験の結果、DFPN-E は、先行研究の手法よりも平均して3倍の速さで計算を行うことができた。

<sup>1</sup> "Iterative Resource Allocation for Memory Intensive Parallel Search Algorithms", In Proceedings of the 26th AAAI Conference on Artificial Intelligence, pages 478-485..

<sup>2</sup> "Shared-Memory Parallel Frontier-Based Search. WALCOM: Algorithms and Computation," pp. 170-181, Springer, 2013

<sup>3</sup> "Depth-First Proof-Number Search with Heuristic Edge Cost and Application to Chemical Synthesis Planning," NeurIPS 2019

## 大規模画像データの潜在情報抽出に基づく画像生成

島野 美保子(国立情報学研究所コンテンツ科学研究系 特任助教)

研究期間 2008 年 10 月～2014 年 3 月

### さきがけの成果:

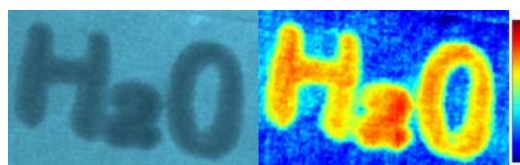
映像の高時間分解能化を実現する物理モデルと非制御下の画像・映像データの統計的学習を融合した手法を提案した。また、空間高解像度化や高階調化と拡張し、画質改善映像の符号化への発展も実現した。



### 発展:

#### 1. 分光画像を用いた濡れと物体色推定<sup>1</sup>

物体の濡れた表面の認識や濡れの程度の推定は多くのコンピュータビジョンアプリケーションにおいて重要な課題の 1 つで、例えば自動運転では路面上の濡れ具合から滑りやすい場所を特定できる。本研究では濡れた表面を認識するために物体の色の変化、特に色のスペクトル変化が表面の濡れ具合に関する豊富な情報をもつことを示した。また物体が濡れると暗く見える現象や、多重散乱の繰り返しで色味がシャープになる現象も確認できた。これらからマルチスペクトル画像を用いて表面の色と濡れ具合を推定するモデルを開発した(図 1)。また実際に濡れた多数の物体を用いて開発したモデルを包括的に検証する実験にて、様々な物体で濡れ具合や色を十分な精度で推定できることを確認した。物体内部の散乱に着目し、濡れることによる物体表面の分光特性の変化をモデル化した点が評価され、CV 分野のトップ国際会議のオーラルに採択され、新聞でも報道された。



(a) 入力画像 (b) 濡れ度合いの分布

図 1. 濡れ度合いの分布推定<sup>1</sup>

#### 2. 可変リングライト撮像による物体表層構造の可視化<sup>2</sup>

光は物体などに作用し、反射、屈折、散乱などの光学的過程を繰り返しながら伝播する。特に光が物体表面に入った際には、物体表面の粒子等にぶつかる度に吸収と反乱を繰り返すことで複雑な伝播をする。そのため、物体表面の光の伝播を撮像・可視化する方法はこれまでなかった。本研究では新たに導出した物体の外観ごとの光源の照射方法を用いて、光の伝播距離の制御にて可視化する方法を開発した(図 2)。本方法は、通常の可視光光源とカメラで物体表層内の光の拡散や散乱の伝搬過程を可視化した点が特徴で、果物や人間の皮膚など様々な自然物体にて可視化を検証した。本研究の結果は X 線や高速カメラと異なる簡便な実装で物体表面の状態を非接触で捉えられるため、とりわけ医療や美容などの領域での非接触検査への応用が期待される。

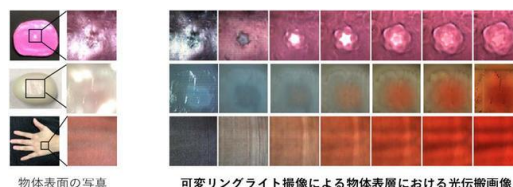


図 2. 物体表面における光伝播画像<sup>2</sup>

#### 3. 染色を行わない生体組織等における顕微鏡観察の実現<sup>3</sup>

透過光顕微鏡では散乱光の総和を撮像素子が捉えるために、試料の観察が不鮮明な状態で行われる傾向がある。本研究では透過光顕微鏡において、高周波照明を用いて直接透過光と散乱光を分離する新しい計算機撮影法を提案した。本研究で開発した手法を実際の生体組織を用いた実験にて検証し、本手法の有効性を実証した。また、本手法は染色を伴わずに透過光と散乱光を分離し、生体組織の吸収特性を解析できる画期的な計測技術として高い評価を得た。さらに近年がんの検出等に有効と注目されている散乱角度情報に着目した新たな散乱光分析法の可能性を見出した。

<sup>1</sup> "Wetness and Color from a Single Multispectral Image," CVPR2017, pp.321-329

<sup>2</sup> "Variable Ring Light Imaging: Capturing Transient Subsurface Scattering with An Ordinary Camera," ECCV (11) 2018, pp. 624-639、国立情報学研究所プレスリリース(2018 年 9 月 7 日)、特願 2019-079588

<sup>3</sup> "Separation of Transmitted Light and Scattering Components in Transmitted Microscopy," MICCAI (2) 2017, pp.12-20、特願 2019-537702



### 3.1.6 擬似コード変換と統計解析による文書画像からの知識抽出 (寺沢憲吾)

#### 擬似コード変換と統計解析による文書画像からの知識抽出

寺沢 憲吾(公立はこだて未来大学システム情報科学部 准教授)

研究期間 2008 年 10 月～2012 年 3 月

展開している事業:

科研費若手(A)

#### さきがけの成果:

機械による文字認識が困難な文書画像に対し、これをテキスト化してから解析する手法のみに頼るのではなく、文字を含む画像を画像のまま解析する手法を開発し、高速な全文検索を実現するとともに、頻出語句の抽出や重要度の評価法などについての知見を得た。



#### 発展:

##### 1. 未翻刻文書画像からのキーワード(重要語)の抽出法を開発<sup>1</sup>

本研究では、歴史的な新聞画像から特徴的なキーワードを抽出する新しいアプローチを提案した。未翻刻文書画像から、既存研究で行われている光学的文字認識(OCR)による文字認識を経由せずに、画像解析の手法によって頻出語を抽出する。具体的には、単文字画像の画像特徴量のクラスタリングと離散化による擬似コード化によって抽出された単語のうち、出現頻度の高い文字列を特徴的なキーワードとして抽出した。また、これらの出現頻度の高いキーワードは集中的に出現することに着目し、その出現間隔の分布の違いに着目して重要度を評価する方法を開発した。この手法を 19 世紀に発行された日本の新聞画像のアーカイブに適用したところ、有望な結果が得られた。

##### 2. 未翻刻文書画からの頻出語抽出の精度向上<sup>2</sup>

大量の文献資料画像が貯蓄・公開されているデジタルアーカイブの活用を考える上で、各文献資料画像の索引やタグ付けされたキーワードがあると便利である。本研究では文字認識が困難な文献資料画像に対して、文字認識を行わずにキーワードを抽出するシステムを開発した。本システムは、特徴ベクトルで表された文字画像群を k-means 法によってクラスタリングすることで、画像データのみからの頻出語抽出を可能とした(図 1)。しかし一方で、文献資料画像内の文字数が非常に多い場合に、クラスタリングの膨大な計算コストや 1 つの字種が複数のクラスタに分離する問題が生じるため、本研究ではそのような場合でも適用可能なクラスタリング手法を提案した。

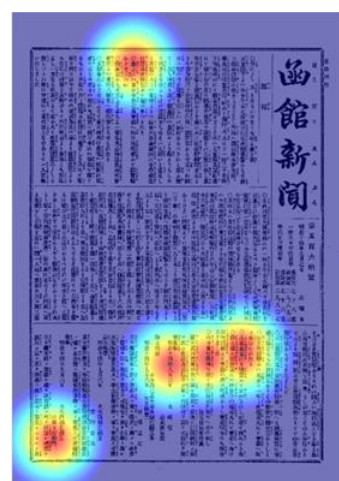


図 1. 古文書から抽出された頻出語

##### 3. 高次元特徴量ベクトルの最近傍探索を行う改良型 locality sensitive hashing (以下、LSH)アルゴリズム<sup>3</sup>

近年、画像、動画、音声などの多様なデータが氾濫しており、大量のデータから必要なデータを適切に選択する技術がより重要となってきた。本研究では問い合わせデータを高速に探索する最近傍探索アルゴリズムの開発を目指し、1000 万の画像からなるデータベースの画像検索において、Spherical LSH (以下、SLSH)という手法を用いることで、従来手法の Exact Euclidean LSH (以下、E2LSH)と比較して、より高速に探索が行えることを確認した。また、データセットサイズが大きくなると、メモリ制約から SLSH の高速特性の低下を確認し、インタセクション類似度を用いる場合は改良版 LSH という手法が最近傍探索のアルゴリズムとして有効であることを確認した。

<sup>1</sup> "Extraction of Distinctive Keywords and Articles from Untranscribed Historical Newspaper Images,"

International Workshop on Advanced Image Technology, IWAIT2020, Jan. 2020, KAKEN「文書画像アーカイブに対するテキスト情報に依存しない内容解析」

<sup>2</sup> "文字認識が困難な文献史料画像の解析のための文字画像クラスタリング手法," 電子情報通信学会 PRMU 研究会, (2018.9), KAKEN「文書画像アーカイブに対するテキスト情報に依存しない内容解析」

<sup>3</sup> 科研費「高次元特徴量ベクトルの最近傍探索を行う改良型 LSH アルゴリズムの研究」研究成果報告書

### 3.1.7 健康被害を監視するための多言語ウェブサーベイランスシステム (Nigel H. COLLIER)

#### A Multilingual Surveillance System for Detecting and Tracking Public Health Threats

Nigel H. COLLIER (Professor of Natural Language Processing)

Research Period: 10/1/2008 – 3/31/2012

#### Research achievement of the PRESTO :

In order to enable public health experts and the government to collect timely and reliable information from many news sources, a system with algorithms and resources effective for detecting public health threats was developed. We also constructed a schema for automatic annotation of geo-temporal information and to develop a knowledge model for classifying reported outbreak cases for topical relevance.



#### Further research progress based on the PRESTO achievement :

##### 1. Global Health Monitor – a Web-based system for detecting and mapping infectious disease outbreaks<sup>1</sup>

The Global Health Monitor was developed, which was an online Web-based system for detecting and mapping infectious disease outbreaks that appear in news stories during this research(fig.1). The system analyzes English news stories from news feed providers, classifies them for topical relevance and plots them onto a Google map using geo-coding information, helping public health workers to monitor the spread of diseases in a geo-temporal context. Running on a cluster-computer, it monitors more than 1500 news feeds 24/7, updating the map every hour.

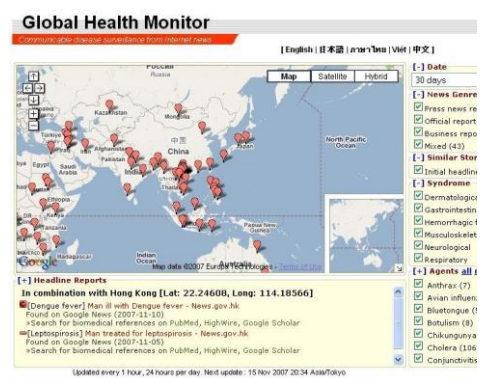


Figure 1. Global Health Monitor system<sup>1</sup>

##### 2. Completing the PheneBank project<sup>2</sup>

Existing scientific literature has the potential to be an incredibly valuable source of data for uncovering the relationships between diseases, phenotypes and genes. This project seeks to extract statistically significant associations between phenotypes, diseases and genes. Earlier approaches deal with texts without classifying words into distinct notions of genetic, pharmacological, and epidemiological relations etc. We propose a technique that de-conflates the representations of words based on the knowledge which derives from a semantic network. Our approach provides multiple advantages in comparison to the conventional works, including the ability to perform more accurate natural language processing in terms of semantic representations.

##### 3. Semantic Interpretation of Personal Health Messages<sup>3</sup>

Open online data such as microblogs and discussion board messages have the potential to be an incredibly valuable source of information about health in populations. On the other hand, social media poses a challenging area for the use in the medical field for a variety of reasons: short de-contextualised messages, high levels of ambiguity/out of vocabulary words, as well as inherent bias towards sensational topics. In this work, we adopted two techniques using neural networks in order to handle the transition between text language used in social media messages and formal medical language at the semantic level. Our approaches evaluated using three different datasets, where social media texts are extracted from Twitter messages. Our experimental results show that our proposed approaches outperform existing effective baseline.

<sup>1</sup> "Global Health Monitor – A Web-based System for Detecting and Mapping Infectious Diseases," The International Joint Conference on Natural Language Processing, 2008, January 2008

<sup>2</sup> Personal lab website, <sup>3</sup> Personal lab website

\*As of 9 March, 2021, the contents is currently under confirmation. This draft has been shared from the same day.

**時空間解析に基づくインターネット異常トラフィックの検出とそのデータベース化**

展開している事業:  
 科研費基盤(B)2件、科研費若手(A)

福田 健介(情報・システム研究機構国立情報学研究所アーキテクチャ科学研究系 准教授)

研究期間 2008年10月～2012年3月

**さきがけの成果:**

理論的な背景の異なるインターネットトラフィック異常検出器を複数開発し、これらの異常検出器を組み合わせることで検出精度を向上させる技術を開発した。さらに、オープンデータである MAWI トラフィックデータへ異常検出器を適用し、検出した異常をラベルとして提供するデータベースを開発して、公開した。



**発展:**

**1. グラフレットに基づくインターネットトラフィック中のアプリケーション同定手法の開発<sup>1</sup>**

インターネットにおけるエンドホストの特定はトラフィックエンジニアリングにおいて重要な課題である。そこでパケットの中身を見ることなく、ネットワークのグラフ構造の統計的なマッチングにより、トラフィック中のエンドホストを同定する手法の研究を実施した。特に本研究ではグラフレットという手法に着目した。このグラフレットにはホスト情報の定義を利用する教師あり手法とホストをクラスタリングする教師無し手法がある。前者は定義済みのホストを確実に特定できるが、未定義のホストには適応できないことに対して、後者は未定義のホストにも有効だが、数値的な特徴量に基づく解釈性の低いグループを生成してしまい、教師あり手法より特定力で劣る。本研究では、この2つの手法を組み合わせ、上記の性能トレードオフを解消したシノプティックグラフレットという手法を開発した。この手法を既存の手法と比較したところ、クラスタリングの有効性とその後解釈の妥当性を実証することができ、未定義のホストの特定に対してもある程度の解釈性を伴った特定が可能となった。

**2. DNS(domain name system)グラフ上での悪性ドメイン特定手法の開発<sup>2</sup>**

インターネット上で悪意のあるユーザを特定することは、一般の人々が安全にインターネットを利用するうえで重要なタスクである。本研究では DNS トラフィック中のドメイン名・IP アドレスからなる二部グラフに確率伝搬法を用いて、悪意のあるドメイン名を検出する手法を提案した。本手法を既存手法である N-gram と比較したところ、約 40%の性能向上を実現した。また提案手法を用いて推定した DNS グラフを可視化することで、悪性ノードが関わるグラフ構造を明らかにした。今後は、本研究外のデータでの検証にて異なるタイプの悪性ドメインの検出などを通じて、精度の向上が期待される。

**3. DNS backscatter によるネットワークイベント検出手法の開発<sup>3</sup>**

ネットワーク全体の活動は1台のコンピュータが他の多くのコンピュータに接触することで構築される。その活動には良性的なものから悪意のあるものまで様々あるが、これらの活動の特定はシステムの向上やセキュリティの観点で重要である。そこで本研究では、ネットワーク中で生じる大規模イベント(ネットワークスキャン、スパム送信等)の良性的悪性を権威 DNS サーバーのトラフィックの情報を用いた機械学習によって同定する手法を開発した(図1)。また、この手法を用いて構築されたネットワーク上の活動の識別器は、長期間、自律的に、有効に機能することを確認した。

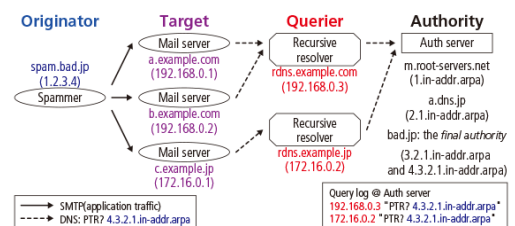


図 1. ネットワークの活動を DNS のデータを用いた機械学習にて推定

<sup>1</sup> "Synoptic Graphlet: Bridging the Gap Between Supervised and Unsupervised Profiling of Host-Level Network Traffic," IEEE/ACM Transactions on Networking, Vol.21, no.4, 2013.

<sup>2</sup> "DNS グラフ上でのグラフ分析と脅威スコア伝搬による悪性ドメイン特定," 日本ソフトウェア科学会 コンピュータソフトウェア, vol.33, no.3, 2016

<sup>3</sup> "Detecting Malicious Activity with DNS Backscatter Over Time," IEEE/ACM Transactions on Networking, vol.25, no.5, 2017.



## マルチソースデータ高度利用のための統計的データ融合

星野 崇宏(慶應義塾大学経済学部・大学院経済学研究科 教授)

研究期間 2008年10月～2012年3月

展開している事業:

科研費基盤(B)2件、科研費国際共同研究強化(A)、科研費若手(A)

### さきがけの成果:

別々の目的で収集された情報源を統合して取得されるマルチソースデータは、解析に必要な変数が得られていない欠損部分が多く存在する。本研究では各データセットで測定対象となっている個々人の背景情報を利用することでマルチソースデータからシングルソースデータとして統計的データ融合により補完する手法を開発した。



### 発展:

#### 1. 異質性を考慮した因果効果の推定法の開発とその応用<sup>1</sup>

欠損データの補完に関する問題は、観察研究や実験研究では非常に重要である。特に連続変数と離散変数が混在するデータセットに対する手法として、連鎖方程式に対する多重入力(MICE)という方法が広く用いられている。しかしこの MICE という手法はある特定の条件下においては著しく偏った推定値を出力する可能性がある。そこで本研究では SB-MI (Semiparametric Bayes multiple imputation) という新しい半パラメトリックベイズ多重入力法を提案した。この手法によって MICE の欠点を克服することができ、実際に提案手法は MICE と比較して MSE(mean squared error : 最小二乗誤差)の観点から最良の推定を行えることが確認できた。一方、本研究の提案手法は MCMC (Markov-chain. Monte Carlo)のアルゴリズムに基づいていることから計算時間が長くなる傾向にあるため、今後はアルゴリズムの効率を向上させていく必要があると考えられる。

#### 2. 行動ログと調査回答の乖離の理解及び介入法による改善<sup>2</sup>

一般に行われる調査において、調査の回答と問われている実際の行動の乖離が大きくなる点が、調査結果に大きく影響する。例えば「どのくらいの頻度でテレビを見ますか?」という質問への回答がその回答者の実際の行動とどの程度乖離しているのかという問題は、これまで先行研究でも解明されていない。そこで本研究では過去の行動についての調査の回答が実際の行動とどの程度乖離するのかを調べると共に、その乖離を調査時の心理的介入によってどの程度減少させることができるのかを調べた。購買行動についての実験結果から購買回数や購買金額、利用時間などの回答は平均として大きな乖離があること、またその乖離には個人差が大きくありそれらが個人の属性情報や行動経済学的な指標によって説明できること、心理的介入により乖離を減少させることができることが確認できた。

#### 3. ソーシャルメディア発信情報のバイアス補正法の研究<sup>3</sup>

近年マーケティングなどの分野において SNS などのソーシャルメディア上の情報は非常に重要なものとなりつつある。一方でネット上の情報発信は「書きたい人が書き、答えたい人が答える」といった形のものであるため、情報を扱う上で対象としたい母集団からのサンプリングとしては不適切であり、回答にはバイアスが生じていると考えられる。そこで本研究では、ネット情報の偏りを補正するために、セミパラメトリックベイズモデリングを用いた選択バイアスモデルを開発した。この手法を実際のインターネット閲覧情報に適用してロコミの効果について調べた結果、先行研究でも言われていたロコミは似たような人々を集める効果が高いこと、またその効果を除去した場合ロコミの効果は小さいことが示された。

<sup>1</sup> KAKEN「異質性を考慮した因果効果の推定法の開発とその応用」

<sup>2</sup> 科研費「行動ログと調査回答の乖離の理解及び介入法による改善」研究成果報告書

<sup>3</sup> 科研費「ソーシャルメディア発信情報のバイアス補正法の研究」研究成果報告書

**ネットワーク理論と機械学習を用いたウェブ情報の構造化・知識化**

展開している事業:

松尾 豊(東京大学大学院工学系研究科 教授)

科研費新学術、科研費若手(A)

研究期間 2008 年 10 月～2012 年 3 月

**さきがけの成果:**

ネットワーク理論と機械学習を用いて、ウェブ情報を構造化・知識化する技術を開発した。その結果、ウェブ上の情報を自動的に統合できるようになり、逐次得られるウェブ上のデータを基に、現在発生している社会現象を予測(地震情報の推定や選挙予測など)または把握する先進的な取り組みとしてインパクトを残した。



**発展:**

**1. 深層生成モデルによる異なるモダリティ間の双方向生成に関する研究<sup>1</sup>**

現在、データを基に学習したニューラルネットワークが新たに別のデータを生成する手法が注目を集めている。その生成モデルの中でも画像から文書を、逆に文書から画像を生成するというように異なるモダリティ間を双方向に生成するモデルの構築には、モダリティ間共有の表現の獲得が必要である。従来は Joint Multimodal Variational Auto-Encoder(JMVAE)という手法で共有表現を獲得していたがデータに欠損がある場合に対応できなかった。そこで JMVAE-kl と階層的 JMVAE という手法を開発し、これらが既存の手法と同等以上の精度で双方向生成できることを確認した。本研究では、人の顔画像と笑っている男性というような簡潔な属性情報を用いて双方向の生成が実現されている(図 1)が、今後は画像と文書情報というような構造の異なるデータ間の双方向生成の検証で、より実用性の高い技術への発展が期待される。



図 1. 生成モデルによる属性の変更

**特記事項<sup>2</sup>**

- ・東大松尾研究室では高度なディープラーニング技術者育成を目的としたエンジニア向けオンライン教育プログラムを無償公開。2018 年から 2 期の開講でのべ 3800 人以上の申込みがあり、累計 500 人が受講するなど技術者養成に貢献<sup>2-1)</sup>。研究室でも起業家の養成に取り組み、研究室の元メンバが中心となって起業された実績も挙げている<sup>2-2)、2-3)</sup>。
- ・「東大准教授に教わる「人工知能って、そんなことまでできるんですか?」(2014 年出版)や、「人工知能は人間を超えるか」(2015 年出版)等の書籍の執筆により、人工知能を世間にわかりやすく伝え、特に後者は大川出版賞(2015 年)、ビジネス本大賞審査員賞(2016 年)等の受賞により評価を受ける<sup>2-4)</sup>。

上記の取り組みなどが、ディープラーニングのその後の日本でのさまざまな啓蒙活動につながった。

<sup>1</sup> “異なるモダリティ間の双方向生成のための深層生成モデル”情報処理学会誌 Vol 59 No.3 59-873(Mar. 2018)、研究者 HP

<sup>2</sup> 1)IT メディア記事(<https://www.itmedia.co.jp/news/articles/1905/16/news073.html>)、

2)日経ビジネス記事(<https://business.nikkei.com/atcl/NBD/19/special/00448/>)、

3)日本経済新聞(<https://www.nikkei.com/article/DGKKZO61363820Z00C20A7TJ2000>)、4)中小企業庁スマート SME(中小企業)研究会(第 3 回)配布資料

### 3.2 2009 年度採択研究課題

#### 3.2.1 物語構造に基づく情報編纂基盤技術 (赤石美奈)

##### 物語構造に基づく情報編纂基盤技術

赤石 美奈(法政大学情報科学部大学院工学系研究科 教授)

研究期間 2009 年 10 月～2013 年 3 月

##### さきがけの成果:

大量に蓄積された情報を機械的に処理し、大量の情報の中に隠された潜在的な物語を紡ぎだすことを目的として、物語構造モデルの導入にて文書の意味を解釈せずに物語構造を抽出し、文書を分解・再構成する手法を開発した。また本手法で得られた文脈(物語)を情報検索してアクセスするための情報検索手法についても研究を実施した。

##### 発展:

##### 1. 歴史知識学における資料の知識化・構造化フレームワークの構築<sup>1</sup>

様々な分野において、時間属性を伴うテキストデータ(ニュース、会議録等)が大量に蓄積されており、それらを有効活用する需要が高まっている。しかしそれらの大量のデータを情報ごとに関連付けてまとめるなどの処理は容易ではない。本研究では特に歴史資料に着目し、歴史資料に記載された情報を一度分解し、そして関連する情報を纏め直して再統合することであらたな歴史に関する知識を生み出すフレームワークを提案した。

歴史というのは時間的な変化を記述した編年データであり、資料の情報を纏める際には時間的な変化を記述する必要がある。そこで本研究ではまず編年型データ解析プロセスモデルの構築を実施した。データツール CAT(Chronicle Analysis Tool)は様々な視点からデータを可視化し、データの分布やパターンの変化の発見を促すツールである。特に時系列分布状況を俯瞰し、歴史の時間的な変動をとらえることに役立つ。特に本研究では歴史資料の分野における解析に有用な情報の関係性やパターンに基づいてデータを分析・可視化するフレームワークに対して CAT の拡張を提案した。例えば歴史における人間関係の変化は重要な情報の 1 つだが、この人間関係の変化がどのようにして生じたのかは歴史の解析には重要な視点である。そこで人間関係を贈答品や戦などの観点からどのような変化が生じたかを上記のフレームワークを用いて解析した(図 1)。このようにある要素の関係の変化をデータに基づいて特定の観点から可視化することで歴史を俯瞰でき、歴史上の出来事がなぜ起こったのかの考察などがし易くなるため今後の歴史学の発展に貢献できると考えられる。

##### 2. Meme Media に基づいた文脈依存の情報検索技術の開発<sup>2</sup>

本研究では「いつ」、「どこで」といった観点に関連する情報検索を行う文脈依存の情報検索技術の開発を実施した。また文脈に埋め込まれた興味深い関連性の発見を支援するために視覚的な年表分析のためのフレームワークとし、このフレームワークを基に文脈に潜むパターンを表現する機能を提供するツール「CAT」を実装した。これにより、歴史などに対してユーザは様々な視点から新しい解釈をすることが簡単にできるようになる。またこの技術既存の可視化ツールなどと組み合わせることで文脈の可視化など新たな応用可能性を切り開くことができると考えられる。

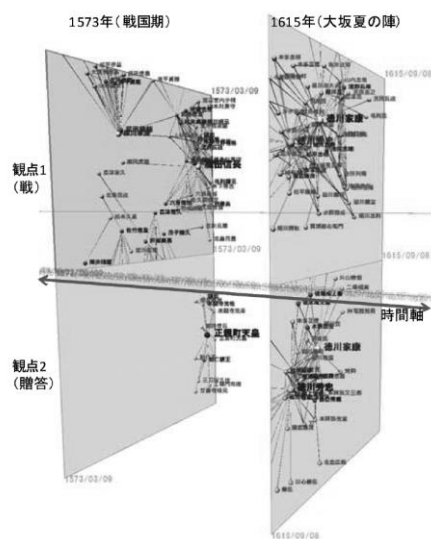


図 1. CAT を用いた贈答品や戦に基づく人物ネットワークの変化<sup>1</sup>

<sup>1</sup> 科研費「歴史知識学における資料の知識化・構造化フレームワークの構築」研究成果報告書

“歴史資料を解析する-歴史知識学の創成-,” 人工知能学会誌, vol.31, No.6, 2016

<sup>2</sup> “A Meme Media based Application Framework for Context-Driven Information Access,” Communications in Computer and Information Science, vol.372, 2013



### 3.2.2 マルチエージェントの交渉と協調に基づく集合的コラボレーション支援システムの開発 (伊藤孝行)

#### マルチエージェントの交渉と協調に基づく集合的コラボレーション支援システムの開発

展開している事業: CREST、JSPS NEXT、NICT 委託研究、科研費基盤(A)

伊藤 孝行(京都大学大学院情報学研究科 教授)

研究期間 2009 年 10 月～2011 年 3 月

#### さきがけの成果:

マルチエージェントシステムの自動交渉技術とその応用について、理論的アルゴリズムの検討を行い、エージェント間の自動交渉機構を用いた、デザイン支援システムを試作し、その評価を行った。

#### 発展:

##### 1. マルチエージェントシステムに関する研究<sup>1</sup>

マルチエージェントシステムの自動交渉技術とその応用について、理論的アルゴリズムの提案と、応用システムへの展開を行った。具体的な成果としてはグラフ理論とゲーム理論を組合せ、グラフに基づく新しい効用モデルを提案し、従来よりも複雑な効用空間の表現を可能とした。また、機械学習の技法を用いることができる新しいシミュレーション環境を実装し、より高速に大規模シミュレーションが実施できるようになった。(図 1)

##### 2. マルチエージェントシステムの自動交渉技術とその応用の研究<sup>2</sup>

人間の社会を成立させるための根源的社会活動の 1 つとして交渉や合意形成があげられる。しかしその負担は大きく、極めて多くの時間が交渉や合意形成に割かれている。また合意形成に至った結果が必ずしもよいとは限らない。人間の交渉や合意形成をマルチエージェントモデルでシミュレーションを用いて援助することを目的として、特に自動交渉アルゴリズムについてスケーラビリティの拡大と、議論における自然言語処理に基づく合意形成及びファシリテーション支援について研究を進め、理論及びモデルを構築した。

##### 3. 自動交渉機構に関する理論とその応用に関する研究<sup>3</sup>

マルチエージェントシステムで研究されている自動交渉や効用モデルに関する理論およびモデルを、大規模合意形成支援システムや議論過程と合意形成の支援などに応用した。評価のために名古屋市と協力して、名古屋市次期総合計画の中間案に対する市民の意見集約の場の一つとして、本システムを導入した。その結果、本システムは予想以上にうまく動作し、参加者からの投稿数が増加する傾向が見られた。

##### 4. マルチエージェントシステムにおける交渉・協調機構に関する研究<sup>4</sup>

合意最適化エージェントのプロトタイプのためのコア技術として、様々な機能や手法について探索的に研究を行った。特に、既存の手法やアーキテクチャに問題がある場合には再度設計をし直す等により、無駄のないシステムに改良した。また、これらのシステムを用いて様々な大規模議論支援実験を行うことで、多くの議論データを収集した。今後、同様の大規模議論支援実験を進めることで、更なる技術の発展が期待できる。

#### 特記事項

2019 年 3 月に、資本金 100 万円で Agreebit 株式会社を設立。

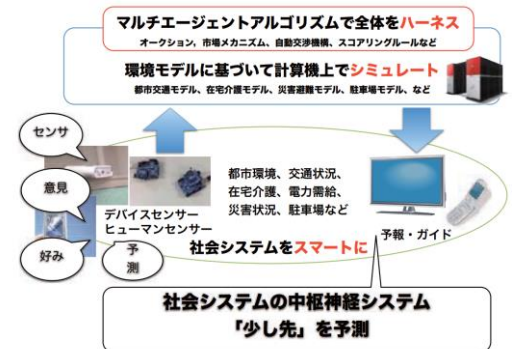


図 1. エージェントによる新しい社会システムの創生の創生

<sup>1</sup> 科研費「マルチエージェント自動交渉理論とその評価に関する研究」成果報告書、研究者 HP

<sup>2</sup> 科研費「マルチエージェント自動交渉理論とその評価に関する研究」成果報告書

<sup>3</sup> 人工知能学会全国大会論文集(第 33 回(2019))「エージェント技術に基づく大規模合意形成支援システムの創成: 自動ファシリテーションエージェントを用いた大規模社会実験」

<sup>4</sup> 科研費「マルチエージェント自動交渉理論とその評価に関する研究」成果報告書

**次世代データ同化：自動モデル化と情報フロー抽出技術開発**

展開している事業：

上野 玄太(情報・システム研究機構統計数理研究所モデリング研究系 准教授)

科研費基盤(A)、科研費基盤(B)

研究期間 2009 年 10 月～2013 年 3 月

**さきがけの成果：**

省コストのデータ同化手法(グラフベース・アンサンブルカルマンフィルタ)を開発した。モデルによる予測結果に入力変数のどの部分がどの程度寄与しているのかを対応づけて、結果に対する入力の感度分布を容易に求める方法を開発した。また、観測誤差の自動推定を含むデータ同化手法を開発した。



**発展：**

**1. アンサンブルフィルタにおける観測誤差のベイズ推定の開発<sup>1</sup>**

データ同化システムにおける正規化パラメータの最適化手法を開発し、大規模なシステムを対象とするパラメータ調整を反復アルゴリズムによって推定する方法を提示した。本手法は、観測ノイズの共分散行列( $R_t$ )をオンラインで推定でき、事前共分散( $S$ )とその重みを適切に選択することで、 $R_t$ の時間変化の滑らかさを制御することが出来る。これらの研究成果は地球物理学などの問題における非線形状態の推定に役に立つと期待される(図 1)。

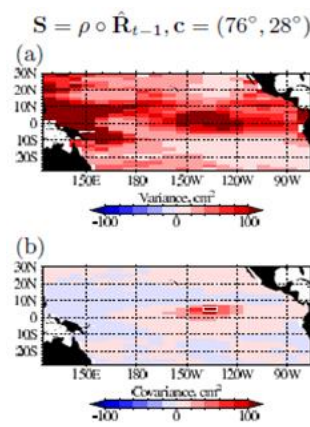


図 1. 海面高度変位の観測ノイズの分散および共分散較<sup>1</sup>

**2. 変分ベイズ法に基づく観測誤差推定機能を持つアンサンブルカルマンフィルタ<sup>2</sup>**

多くの場合、関心のある現象から必要なデータを全て観測することはできず、観測できた場合でも、データにノイズが含まれる場合が多い。このようなノイズを含んだ観測データから所望のデータを推定する手法としてカルマンフィルタが知られている。本研究では観測ノイズの性質が時間変化する場合を想定して、カルマンフィルタの一種であるアンサンブルカルマンフィルタ(EnKF)を拡張し、変分ベイズ法を用いて状態ベクトルと観測誤差共分散行列の同時推定を実現した。本手法は時間的に変化するノイズに対しても機能することを数値実験で確かめ、さらに外乱が存在する場合の長期同化も安定であることを示した。

**3. 外れ値を含む観測に対応した誤差切り替え機能を持つ非線形フィルタ<sup>3</sup>**

データに外れ値が存在する場合、データの誤差をガウス分布と仮定するフィルタリングの手法は、推定の精度を著しく低下させる可能性があることが知られている。そこで本研究では観測データに外れ値が入る場合を想定して、状態変数・外れ値の指標・外れ値の大きさの尺度の 3 変数をまとめ、同時分布に関する非線形フィルタ手法を提案した。本手法は外れ値の特性を考慮した追加のチューニングなしにデータセットに適用できる。また数値実験を行った結果、外れ値ありと外れ値なしの両データセットにおいて、本手法が既存の手法より優れた状態推定が可能であることが確認できた。

**4. 気象庁数値予報モデルと粒子フィルタによるデータ同化システムの開発<sup>4</sup>**

対流圏下層の積乱雲の発達過程における各種物理量(温度、水蒸気量、湿度、鉛直風、雲量、雨量)の変化を気象庁数値予報モデルと粒子フィルタにより推定した。積乱雲の発生において、最初に上向きの鉛直風(上昇気流)の確率分布に非ガウスの形状が見られ、続いて水蒸気量・温度の確率分布の形状の変化が現れることを明らかにした。またこの結果は観測なしのシミュレーションと比較して有意な改善を示した。

<sup>1</sup> “Bayesian estimation of observation error covariance matrix in ensemble-based filters,” Quarterly Journal of the Royal Meteorological Society, DOI: 10.1002/qj.2803, 2016., 研究者 HP

<sup>2</sup> “An Extension of the Ensemble Kalman Filter for Estimating the Observation Error Covariance Matrix Based on the Variational Bayes’ s Method,” Mon. Wea. Rev., 145, 199–213, 2017, <https://doi.org/10.1175/MWR-D-16-0139.1>.

<sup>3</sup> “Nonlinear Filtering Method Using a Switching Error Model for Outlier Contaminated Observations,” IEEE Transactions on Automatic Control. doi: 10.1109/TAC.2019.2947649.

<sup>4</sup> “Non-Gaussian Probability Densities of Convection Initiation and Development Investigated Using a Particle Filter with a Storm-Scale Numerical Weather Prediction Model,” Mon. Wea. Rev., 148, 2020, <https://doi.org/10.1175/MWR-D-18-0367.1>



## 大規模データに対する高速類似性解析手法の構築

宇野 毅明(国立情報学研究所情報学プリンシプル研究系 教授)

研究期間 2009 年 10 月～2013 年 3 月

展開している事業:

CREST、科研費基盤(A)

### さきがけの成果:

ビッグデータが内包する大規模データの類似性を超高速で抽出するアルゴリズムの構築と、その抽出された類似性を可視化するための多様でかつ新しい解析手法を作り出した。これを活用することで、大規模ゲノムデータや大規模 POS データなどの局所的・大域的な類似性やそれに基づくパターン抽出を可能とした。



### 発展:

#### 1. 質の高いクラスタリングアルゴリズムの開発<sup>1</sup>

多様性の高いデータからある程度明らかな特徴を持つ局所的なグループ構造を多く見つける問題に対して、局所グループの特徴や特徴を有するグループを見つけ出す様々な手法があり、決まった解決手法がない。既存の手法では複数のグループに属するようなソフトクラスタリングが難しいものも多い。一方で、データが大規模である場合には、全ての項目の組について類似度を調べると膨大なコストがかかるため、類似する物に対してのみ類似度を調べる問題設定で高速なアルゴリズムを開発した。また、そのアルゴリズムを発展させ、類似性を計算する高速なアルゴリズムを元に、マイクロクラスタリングと呼ばれる教師なしソフトクラスタリングの問題に対処した。この問題の目的は、互いに強く関連する記録で構成される全てのグループの列挙だが、標準的なクラスタリング手法では、各グループに含まれる記録が非常に少なくなってしまう。既存の手法では、無数のクラスタが生成される、各クラスタのデータサイズに偏りがある、安定性が欠けている、などの欠点がある。実現可能な仮説に基づき、データを摂動することにより、データのクラスタ構造を研磨する、新しいデータ研磨法を検討した。これにより、細かい構造を網羅的に抽出できるクラスタリングアルゴリズム、データ研磨アルゴリズムを開発した。さきがけの研究成果は、計算の高速性に活用されている。この類似性解析・クラスタリングアルゴリズムにより、顧客分析、経済動向把握、社会ネットワーク解析、センサデータ解析などで、表層には現れない深い意味が見えるようになる。具体的な応用先として、婚活のためのマッチングレコメンデーションエンジン、広告配信事業者のターゲティングアルゴリズムなどに用いられている。

#### 2. データ粒子化による高速高精度な次世代マイニング技術の創出<sup>2</sup>

列挙アルゴリズムは、理論計算機科学分野における、最近のホットトピックの 1 つである。列挙には興味深い特徴があり、計算の高速化や償却計算量解析の効率化に結び付く可能性があるが、基盤的な研究は限定的である。そこで、複雑性解析に対する新しいアプローチに取り組み、反復の計算時間をグラフ内の全ての親子関係に基づき反復を使用して償却分析を適用する新しい列挙アルゴリズムを提案した。償却分析が機能するように、列挙アルゴリズムの十分条件を明確にした。また、グラフ中の消去順序やマッチング、グラフ内の頂点連結している部分グラフ、およびスパニングツリーが、償却解析により、単純なアルゴリズムによって、 $O(1)$ 時間で列挙できることを示した。

更に、さきがけ研究成果を発展させ、CREST の研究領域「ビッグデータ統合利活用のための次世代基盤技術の創出・体系化」において、2 つのアプローチにより、アルゴリズムの高速化を達成した。まず、バイクラスタリングアルゴリズムの開発により高速化を行った。多数の小さなクラスタに分割するよりも少ないクラスタ数で分割できる現実的なデータを対象にした時に高速になるよう、データの質を活かしたアルゴリズムを設計したことで達成された。また、処理の並列化も成功し、4 コア使用した場合に 3 倍程度の高速化が達成できた。次に、研磨アルゴリズムを用いた名寄せアルゴリズムの開発により、高速化を行った。論文タイトルや広告文句などのセンテンスを単語に分解し、類似性に基づいて抽象クラスタ化することで、条件によっては、1,000 倍を超える高速化を実現した。

<sup>1</sup> Uno, Takeaki, et al. "Micro-clustering by data polishing." 2017 IEEE International Conference on Big Data (Big Data). IEEE, 2017. JST CREST 事業に採択された。(課題番号 JPMJCR1401)

<sup>2</sup> Uno, Takeaki, et al. "Constant Time Enumeration by Amortization" LNCS 9214, WADS 2015, 593-605, CREST「データ粒子化による高速高精度な次世代マイニング技術の創出」平成 28 年度実績報告書

## 映像分析による知識の抽出と、その利用による新たな映像合成

岡部 誠(静岡大学大学院総合科学技術研究科工学専攻数理システム工学コ

ース 准教授)

研究期間 2009 年 10 月～2013 年 3 月

### さきがけの成果:

大量の映像データから抽出した知識を利用することで映像制作を支援する技術の研究を行った。代表的な研究成果として、流体画像のデータベースに基づくアニメーション生成、料理画像のアニメーションによる魅力的な料理動画生成システム、動画内オブジェクトを容易に編集可能な 2.5 次元動画編集システム検討を実施し、国際学会などで発表した。



### 発展:

#### 1. 動画検索を用いた水シーン画像のアニメーション手法の開発<sup>1</sup>

自然な水のアニメーションを実現には、高い計算コストなどの理由から単調な動きのみの実現に留まっている。そこで本研究では写真や絵画における自然な水(水面、滝、河川など)のアニメーション手法を開発した。本手法では計算コストを下げつつ水の流れの大局的な情報でアニメーション生成を行うために、数百程度の水の動画データベースを作成し、オプティカルフローにより水の動きなどを予め解析した。そして映像合成の際には見た目と適合する動画をデータベースから検索し、合成時に背景か、流体かを示す指標に基づいたぼかしフィルタの適用で質の高いアニメーションを生成しつつ計算速度も改善した。一方でデータベースに適合する動画がない場合やノイズを含む動画を使う場合に生成画質が落ちるといった欠点の確認されたため、今後はこれらの問題点を解決しつつ炎や煙等他の流体にも手法を拡張する。

#### 2. インタラクティブな動画修復手法の開発<sup>2</sup>

本研究ではデジタル制作現場での実用化を目的とした対話型の映像修復手法を提案した。特徴は大きく 3 つあり、1) 既存手法と比べ、反復アルゴリズムによる動画の速度場と色を交互に最適化で修復結果のクオリティが高い、2) 計算が速く、インタラクティブな修復作業が可能、3) キーフレームを与え、結果を更に修正できる、である。本手法を用いて制作した映像は、他の最先端の手法を用いて制作された映像に比べて格段に見栄えがよくなることが確認できた。

#### 3. 疎な多視点画像からの流体ボリュームのモデリング手法の開発<sup>3</sup>

本研究では疎な多視点画像(例えば直交 2 視点の画像)から流体现象を 3 次元モデリングするための手法を提案した(図 1)。従来手法では視点を変えた画像を生成するとぼやけて不自然になるが、本手法では入力画像の見た目の情報を変換したモデリングであらゆる方向から見たときの立体的かつ自然な見た目を構築できる。また市販の流体シミュレータを用いたレンダリングや編集の手法も提案した。



図 1. 単一視点の画像を用いて 3 次元流体モデルに基づきアニメーションを生成する。

#### 4. SizzTass: 静止画に動きを付与する動画作成支援システム<sup>4</sup>

ショッピングサイトの商品イメージなど Web 上のサムネイルの多くは静止画像である。これらの静止画を低コストで動画化できれば効果的に人々を惹きつけることができる。本研究では特に炭酸の泡、熱や光の揺らぎに着目し、わずかな入力でも静止画を動画化できるシステムを提案した。ユーザーテストにて動画化に関する専門的な知識がなくとも簡単に、かつ短期間で動画化ができることを確認した。

<sup>1</sup> “動画検索を用いた水シーン画像のアニメーション”, Visual Computing(VIC)研究会/グラフィクスと CAD(GCAD)研究会 合同シンポジウム 2016 予稿集, pp. 18:1-18:5, “Animating pictures of water scenes using video retrieval”, The Visual Computer, doi:10.1007/s00371-016-1337-6, pp. 1-12, December, 2016 (volume 34, issue 3, pp. 347-358).

<sup>2</sup> “Interactive Video Completion”, IEEE Computer Graphics and Applications (CG&A), volume 40, issue 1, pp. 127-139.

### 3.2.6 ベイジアンテレビ：取材・配信・編集を自動化した緊急情報メディア（北本朝展）

ベイジアンテレビ：取材・配信・編集を自動化した緊急情報メディア

展開している事業：

北本 朝展(データサイエンス共同利用基盤施設人文学オープンデータ共同  
利用センター センター長/国立情報学研究所 教授)

SICORP、科研費基盤(A)、科研費基盤(B)

研究期間 2009年10月～2013年3月

#### さきがけの成果：

社会の危機(クライシス)時における情報量の爆発的な増大に対応するため、インターネット上の様々なデータを取  
材・整理し、個人に重要な情報を推薦するメカニズムに基づくプッシュ型メディアの基礎研究を推進するとともに、台風  
(気象)データや東日本大震災データなどを対象とした実践的なサービスの運用を通して有効性を確かめた。

#### 発展：

##### 1. 東日本大震災デジタルアーカイブを約10年にわたって拡充<sup>1</sup>

東日本大震災を契機に開始したデジタルアーカイブは、ニュース  
記事や電力データを対象に、10年にわたって拡充を続けており、関  
連分野の研究者が基礎データとして活用する例も増えている。

##### 2. 気象データのプル型・プッシュ型ハイブリッド基盤の構築<sup>2</sup>

防災情報がリアルタイムに配信される気象庁防災情報 XML を活  
用し、気象データへのアクセスのためのハイブリッド型基盤を開発し  
た。気象庁防災情報 XML をリアルタイムにデータベース化し、SNS  
で配信するプル型・プッシュ型の情報提供システムを構築し、利用  
者が延べ約4万人に達するなど多くの人々が利用する情報基盤と  
なっている。

##### 3. 大規模台風画像データセットの構築と機械学習の活用<sup>3</sup>

台風画像データを長年にわたってデータベース化してきた成果を  
基に、台風勢力の推定や危険な現象の発生の予測などの課題をデ  
ィープラーニングで解決する可能性について研究した。本研究では、熱帯低気圧と温帯熱帯低気圧の分類、および  
熱帯低気圧のカテゴリ分類を行い、前者で94.89%の精度、後者で65.90%の精度を達成した。また回帰を用いて中心  
気圧の時間変化を予測した。また「デジタル台風」のwebサイトは引き続き運用を続けており、年間2000万ページビ  
ュー以上に達するなど災害情報発信の基盤となっている。

##### 4. テキストを自動的に地図化する地名情報処理システム(GeoNLP)の開発<sup>4</sup>

ウェブサイトやSNS等に存在する膨大な自然言語テキストから地名情報を抽出し、曖昧性を解消した上で、地名ご  
との情報に統合・再編成して示すことは、実世界をセンシングする上で重要かつ基本的な処理である。しかしそのた  
めには、ある文字列が地名と非地名の両方として解釈できる場合、また同じ地名が複数の場所を指している場合な  
どを解決する必要がある。そこで自然言語処理(NLP: Natural Language Processing)と地名に関するヒューリスティクス  
を組み合わせ、自動的に地名情報の抽出と曖昧性解消を行い、テキストを自動的に地図化するためのソフトウェア  
GeoNLPを開発し公開した。また地名のリソースとなる地名辞書の共有サイトや周辺ソフトウェアなども開発した。  
GeoNLPの利用は自ら収集した様々なデータの可視化(図1)にとどまらず、オープンソースとして民間企業においても  
利用が確認されている。

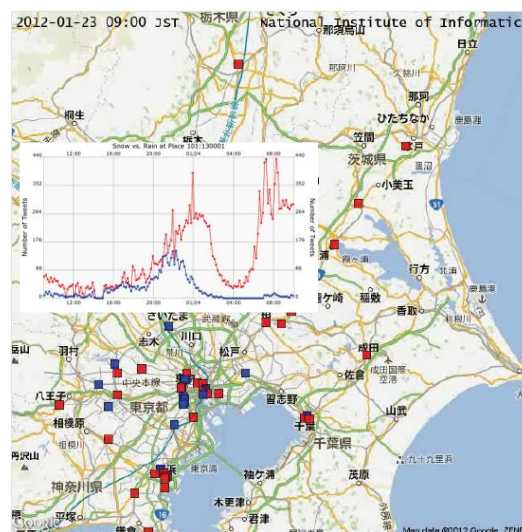


図1. Twitterに基づく天気推定の事例。青が  
雨ツイート、赤は雪ツイートに対応し、図中の  
グラフは当日のツイート数の変化を示す。

<sup>1</sup> <http://agora.ex.nii.ac.jp/earthquake/201103-eastjapan/>

<sup>2</sup> <http://agora.ex.nii.ac.jp/cps/jmaxml/>

<sup>3</sup> <http://agora.ex.nii.ac.jp/digital-typhoon/>

<sup>4</sup> <https://geonlp.ex.nii.ac.jp/>



### 3.2.7 圧縮データ索引に基づく巨大文書集合からの関連性マイニング (坂本比呂志)

#### 圧縮データ索引に基づく巨大文書集合からの関連性マイニング

坂本 比呂志(九州工業大学大学院情報工学研究院 教授)

研究期間 2009 年 10 月～2013 年 3 月

展開している事業:

科研費若手(A)、科研費基盤(B)2 件

#### さきがけの成果:

文字列圧縮の手法である「文脈圧縮」の可能性を広げ、大規模データストリームマイニングに対する基盤技術を開発した。具体的には、文脈圧縮に関して、全文検索や部分文書展開の機構を組み込んだコンパクトな索引構造を構築した。また、このデータ構造とアルゴリズムをさらに拡張し、ストリーム型データに適用可能とした。

#### 発展:

##### 1. 可逆圧縮アルゴリズムのハードウェア超高速化<sup>1</sup>

同じ「さきがけ」研究者である筑波大学山際伸一と協力して、ハードウェア化が容易で、さらに、データストリームを切れ目なく圧縮、復号が可能なロスレスデータ圧縮技術 LCA-DLT(Lowest Common Ancestor-Dynamic Lookup Table)を開発した。LCA-DLTでは、動的ヒストグラム作成により、実装時に利用できるメモリ量に見合った圧縮率を実現することができる。さらにそのヒストグラム情報は圧縮されたデータに暗に埋め込まれるため、復号側にデータが到着するとすぐに復号化が可能となる。2つのデータペアから1つのデータに圧縮するモジュールと、同様の復号化モジュールを複数段組み合わせることで、圧縮率とハードウェア量を自由に調整でき、資源コストが選べるという特徴が生まれた。(図1)

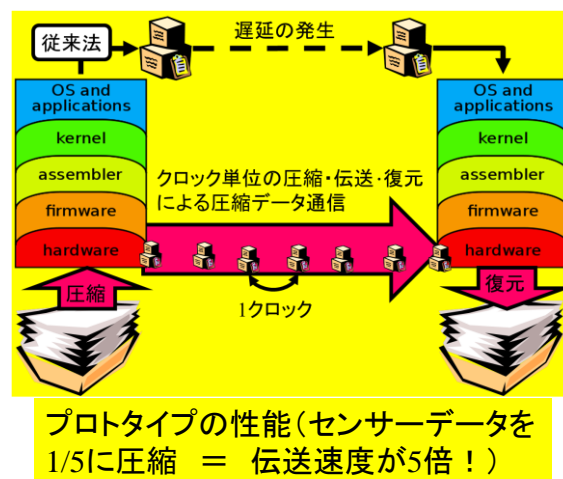


図1. LCA-DLT の特徴

##### 2. 圧縮索引による大規模データベースからの剽窃検出<sup>2</sup>

編集距離は、編集操作に関する指標の一つであり、計算することが難しいが、剽窃検出を含め、幅広い活用が期待される指標である。従来の編集距離は文字単位であったが、これを文字列単位で人間の感覚に似た距離とした。その上で、圧縮したデータから情報検索を可能とする手法の中で、理論的に最速でかつ省スペースで動作するアルゴリズム (siEDM)を実装し、論文の電子ファイルをアーカイブしているプレプリントサーバである arXiv 上の全論文に対する剽窃検出を可能にした。頻繁に更新が行われるプレプリントサーバ上のデータに対する剽窃検出は世界初と認識している。

##### 3. 圧縮索引のプライバシー保護マイニングへの応用<sup>3</sup>

編集距離の計算については、従来、準同型暗号を基とした手法が開発されていたが、長い文字列には活用できない等、この手法の活用には限界があった。坂本は、これを改善したアルゴリズムを開発し、プライバシー情報を含むものであったとしても、特徴量に着目して暗号化したままのデータの比較を可能にする技術を開発した。また、本研究によって拡張された編集距離の暗号検索を初めて実現した。この技術によって、遺伝子データの比較や生態認証による複雑な比較が可能となる。本研究の成果は、特許出願(特願 2019-93908)として出願されている。

#### 特記事項

LCA-DLT については、共同研究者の筑波大学 山際伸一が 2015 年 8 月にストリームテクノロジー株式会社を設立し、実用化を進めている。

<sup>1</sup> “Lazy Management for Frequency Table on Hardware-Based Stream Lossless Data Compression,” Information 7(4), 16pages, 2016、筑波大学「山際研究室」HP

<sup>2</sup> “siEDM: An Efficient String Index and Search Algorithm for Edit Distance with Moves,” Algorithms 9(2), 18pages, 2016

**実社会情報ネットワークからのプライバシー保護データマイニング**

佐久間 淳(筑波大学大学院システム情報工学研究科 教授)

研究期間 2009 年 10 月～2013 年 3 月

展開している事業:

CREST、科研費若手(A)、科研費基盤(B)2 件

**さきがけの成果:**

グラフ構造のデータを用いて、例えばノードにランクをつけることで知識が抽出できるが、エッジの部分が秘密である場合はどのノードが互いにリンクしているかが分からないため、それをデータから計算する必要があった。そこで、①ノードランキング、ノードラベル予測、オンライン予測等のアルゴリズムの実現、②本アルゴリズムをアンドロイドで実現可能なシステムである Fairy Ring の開発、③ネットワーク構造情報を表現する行列を含む半正定値行列を、安全に開示(差分プライバシー)するための理論的枠組みとそれを実現するアルゴリズムの提案、の 3 つに取り組んだ。



**発展:**

**1. 完全準同型暗号による秘密計算手法の設計とその応用<sup>1</sup>**

情報を暗号化したまま加算や乗算が可能である完全準同型暗号を用いて、数値属性データ、順序属性データ、離散属性データなどを暗号化したまま統計解析を実現するための秘密計算手法を設計し、ライブラリを構築した。具体的には、多くの統計計算が行列演算と大小比較演算で記述されることに着目し、完全準同型暗号を用いた効率の良い行列演算と大小比較演算のためのアルゴリズムを開発した。あわせて、統計解析に必要な高精度の数値演算を暗号文上で実現するための平文拡張演算を開発した。これらの新たな手法を組み合わせることで、数値属性、順序属性、離散属性を含む数万レコードの暗号化データを対象として、標準的な記述統計(最大・

最小、平均、分散、共分散、カウント、ヒストグラム、分割表、中央値、最頻値、k パーセンタイル)や予測統計(線形回帰、主成分分析)、統計的検定(カイニ乗検定、正確ロジスティック回帰)などの評価を数秒から 10 分程度で実現することに成功した。例えば線形回帰では、従来の手法に比べて 200 倍以上の効率化を実現した。(図 1)

上記で開発した準同型暗号の利用技術は、個別化予防医療において、個人ゲノムのプライバシーを保護しつつ、個人ゲノム情報や医療情報を関連付けることができ、生活習慣病の罹患リスクをレポートするシステムの構築に繋がった。ゲノム検査機関および 2 つの医療機関を共同し、構築システムを用いて、参加者に遺伝子検査に基づく生活習慣病の予防に関するアドバイスを実施した。

**2. プライバシーを保護したゲノムワイド関連解析<sup>2</sup>**

ゲノムワイド関連解析(GWAS)による疾患に関連する Single Nucleotide Polymorphism (SNP)の同定は、個別化医療の発展に貢献するが、複数機関から情報を収集する場合、プライバシーの保護が解析の障害となる。佐久間は、準同型暗号を用いて、GWAS に用いられるカイニ乗検定量や相関係数などの統計量を、遺伝情報を暗号化したままクラウドに委託計算する手法を開発した。標準的な GWAS の規模である 10,000 サンプル、20 万 SNPs を用いたカイニ乗検定において、既存研究では 2,000 日以上以上の計算と 5,600TB の記憶容量を必要とするところ、提案法では 63 時間 (8core 並列で 8 時間)の計算時間で計算を完了することができることを示した。

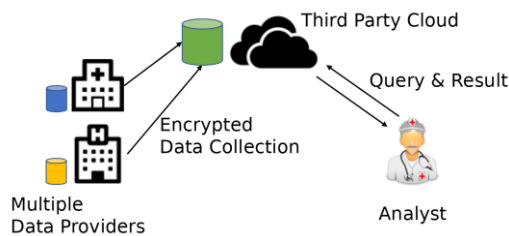


図 1. 暗号化されたデータに対して統計量をクラウド上で計算する

<sup>1</sup> "Privacy preserving Genome wide Association Studies on Cloud Environment using Fully Homomorphic Encryption," BMC Medical Informatics and Decision Making, 2015, 15 (Suppl 5), S1, 2015, CREST 研究終了報告書、研究室ホームページ

<sup>2</sup> "Using Fully Homomorphic Encryption for Statistical Analysis of Categorical, Ordinal and Numerical Data," The Network and Distributed System Security Symposium 2017 (NDSS 2017)、CREST 研究終了報告書

**密度比推定による大規模・高次元データの知的処理技術の創生**

杉山 将(理化学研究所 革新知能統合研究センター長/東京大学 教授)  
研究期間 2009年10月～2013年3月

展開している事業:  
科研費新学術、科研費基盤(A)、科研費若手(A)

**さきがけの成果:**

現在の AI ブームの基盤となっているのは、機械学習と呼ばれるコンピュータに人のような学習能力をもたせる技術である。本研究では、「密度比推定」と名付けた全く新しい独自の機械学習パラダイムを提案し、数学的な理論構築、実用的なアルゴリズム開発、実世界への応用を行った。



**発展:**

**1. 構造的変化の直接検知技術の開発<sup>1</sup>**

工業製品における複数部品の関係性や社会における人々の関係性など、様々な関係性において、変化を捉えることは産業や社会科学において重要である。本研究では、観測データから多数の関係性を示す高次相関の変化を直接推定する独自の技術を開発し、その有効性を理論的・実験的に示した。具体的には、半導体露光装置において、露光装置から得られたセンサログデータを機械学習することで露光装置の機械的な構造と正確な変化量が結果として得られるため、ナノメートルオーダーの露光精度の実現が見込まれる。

**2. 2セットの教師なしデータからの教師付き分類技術の開発<sup>2</sup>**

教師付き分類<sup>注1</sup>は機械学習の最も基礎的な課題の一つであるが、教師付きデータの収集にコストがかかる。本研究では、クラス比の異なる2セットの教師なしデータから教師付き分類の問題が解決可能であることを理論的に証明し、実用的なアルゴリズムを開発した(図1)。学習用の教師付きデータを収集するコストの大きな削減が見込まれる。

注1)教師付き分類は、正しいラベルを付与されたデータを教師データとして学習し分類器を生成し、それを用いて分類する手法

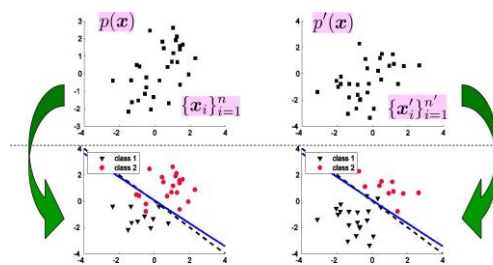


図1. 2セットの教師なしデータから教師付き分類技術の開発

**3. 確率密度微分の直接推定技術の開発<sup>3</sup>**

確率密度の微分に関わる量を、確率密度関数そのものを推定せず、直接求める独自のアルゴリズムを開発し、その有効性を理論的に示すとともに、その実用性をクラスタリングや密度リッジ推定<sup>注2</sup>等において実証した(図2)。これにより AI による分類精度の向上の効果が見込まれる。

注2)密度リッジ推定は、分布しているデータのピークを見つける手法

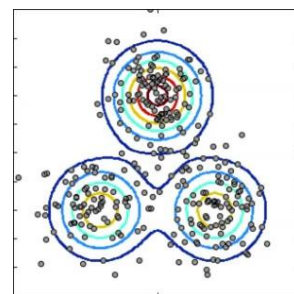


図2. 確率密度微分の直接推定技術の開発

**4. データ不足に対応できる強化学習技術の開発<sup>4</sup>**

近年注目されている強化学習に関する基礎的な技術をまとめるとともに、ロボットの制御などデータを取るのに時間やコストがかかるため、データが十分に取れない場面でも有効な強化学習技術を開発し、その有効性についてロボット制御などを通して実証した。これにより、自動運転など様々な環境下でのシミュレーションが難しい問題に対して効率的に学習できる効果が見込まれる。

<sup>1</sup> Liu, S., Sugiyama, M., Fukumizu, K., Suzuki, T., Relator, R., & Sese, J. Support consistency of direct sparse-change learning in Markov networks. The Annals of Statistics, vol.45, no.3, pp.959-990, 2017.

<sup>2</sup> Lu, N., Niu, G., Menon, A. K., & Sugiyama, M. On the minimal supervision for training any binary classifier from only unlabeled data. In Proceedings of Seventh International Conference on Learning Representations (ICLR2019), 18 pages, New Orleans, Louisiana, USA, May 6-9, 2019.

<sup>3</sup> Sasaki, H., Kanamori, T., Hyvärinen, A., Niu, G., & Sugiyama, M. Mode-seeking clustering and density ridge estimation via direct estimation of density-derivative-ratios. Journal of Machine Learning Research, vol.18, no.180, pp.1-47, 2018.

<sup>4</sup> Sugiyama, M. Statistical Reinforcement Learning: Modern Machine Learning Approaches, 206 pages, Chapman and Hall/CRC, Boca Raton, Florida, USA, 2015.



## インフルエンザ感染伝播のデータ同化モデルによる解析・予測技術

展開している事業:

鈴木 秀幸(大阪大学大学院情報科学研究科情報数理学専攻 教授)

CREST2 件

研究期間 2009 年 10 月～2013 年 3 月

## さきがけの成果:

インフルエンザ感染伝播の数理モデルにデータ同化技術を導入するとともに、インフルエンザ感染伝播の多様性を考慮した数理モデル研究により、モデルと現実との乖離の問題に対処するための数理的基盤技術の開発を行った。

## 発展:

さきがけ研究終了後も一貫して、非線形ダイナミクスに関する数理的基礎研究と実現象の数理モデリングによる応用研究を二本の柱として研究を進めてきた(図 1)。特に、電力システムの非線形数理モデル研究と複雑系計算原理の数理的基盤研究に関しては CREST 等によるプロジェクト研究を展開してきている。

1. 電力システムの非線形数理モデル研究<sup>1</sup>

再生可能エネルギーが大量導入される様々な状況の下で、電力システムを効率的・安定的に運用するためには、電力システムの数理モデルの構築・解析・最適化手法と再生可能エネルギー発電量の解析・予測手法が必要となる。CREST EMS 領域「再生可能エネルギーの大量導入を考慮した電力システムの複雑ネットワーク力学モデル構築とその最適化理論の創成」(2012 年 10 月～2015 年 3 月)および A-STEP「風力発電の高効率化・安定運用に資するリアルタイム予測技術の開発」(2011 年 10 月～2012 年 9 月)において非線形数理モデルの観点から研究を進め、電力システムモデル構築と安定性解析、サイバーセキュリティ指標の解析手法の提案、新しいオンライン発電量予測手法の提案などの成果を上げた。さらに、CREST EMS 領域「太陽光発電予測に基づく調和型電力システム制御のためのシステム理論構築」(2015 年 4 月～2020 年 3 月、研究代表者: 井村 順一 東京工業大学教授)に主たる共同研究者として参画し、数理モデル解析および再生可能エネルギー発電量解析・予測手法を提案してきた。

2. 複雑系計算原理の数理的基盤研究<sup>2</sup>

非線形ダイナミクスに基づく計算として、カオスポルツマンマシン、ハーディング系、ハミルトンモンテカルロ法、リザーバー計算等に関する研究を進めてきた。2013 年に提案したカオスポルツマンマシンは、機械学習等の分野において重要な確率的神経回路モデルであるボルツマンマシンの決定論的実装である。これはカオス・ハイブリッド力学系に基づく新しい複雑系計算原理を提案するものであり、確率的な神経回路モデルを効率的にハードウェア実装できる可能性がある。科研費(挑戦的萌芽)「深層学習ハードウェアのためのカオスポルツマンマシンの集積回路実装」(2015 年 4 月～2017 年 3 月、研究代表者: 森江 隆 九州工業大学教授)では、デジタル・アナログの 2 方式の集積回路によるカオスポルツマンマシンのハードウェア実装を提案している。さらに、これまでの複雑系計算原理に関する研究に基づき、CREST コンピューティング基盤領域「光ニューラルネットワークの時空間ダイナミクスに基づく計算基盤技術」(2018 年 10 月～2024 年 3 月)を実施している。これは、最先端のニューラルネットワーク計算技術と光計算技術を組み合わせることにより、時空間ダイナミクスの観点から新世代の光ニューラルネットワーク計算技術の開発を進めるものである。具体的には、光実装の特性を考慮したリカレントニューラルネットワークモデルを構築し、新しい光ニューラルネットワーク計算原理を提案するとともに、光ニューラルネットワークのハードウェア実装を提案するための研究を実施している。

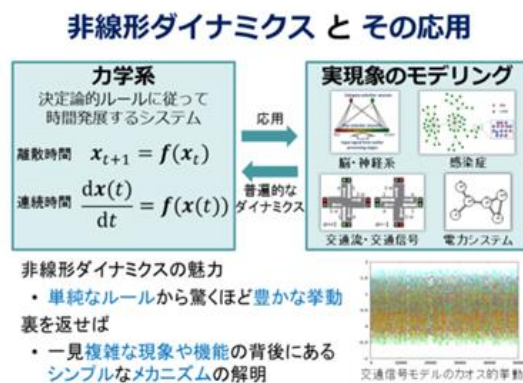


図 1. 非線形ダイナミクスとその応用

<sup>1</sup> “Online multi-step prediction for wind speeds and solar irradiation: Evaluation of prediction errors,” Renewable Energy, Vol. 67, 2014,  
“Enhancing synchronization stability in a multi-area power grid,” Scientific Reports, Vol.6, 2016, CREST 研究終了報告書

<sup>2</sup> “Chaotic Boltzmann machines,” Scientific Reports, Vol.3, 2013, “Hypergraph p-Laplacian: A differential geometry view,” AAAI 2018, 科研費挑戦的萌芽報告書, CREST 研究 HP: [https://www.jst.go.jp/kisoken/crest/project/1111102/1111102\\_2018.html](https://www.jst.go.jp/kisoken/crest/project/1111102/1111102_2018.html)

**金融市場における相転移の時空間構造の自動抽出と予測**

展開している事業:

高田 輝子(大阪市立大学大学院経営学研究科 准教授)

科研費若手(B)

研究期間 2009年10月～2013年3月

**さきがけの成果:**

金融バブルの予測や制御を目指し、まず、相転移現象解析のためのロバスト・効率的情報抽出を可能にする統計手法を開発し、次に、金融バブル生成・崩壊メカニズム解明において核心的な因子である投資家心理についての大規模データを用いた新事実発見を行った。



**発展:**

**1. IPO 時の株価変動に係る投資家行動の解析<sup>1</sup>**

新規株式公開(IPO)時につける高い初値と、その後の長期の株価変動は、コーポレートファイナンスにおいて難度の高い研究対象である。本研究では、これらへの投資家心理の影響を調べるため、Yahoo!Japan Finance のメッセージボードから得られる情報に対し、テキストマイニングとサポートベクターマシンによる分類を実施して、IPO 前の投資家感情を抽出した。その結果、投資家の注目度の高さや強気な投資家心理が、IPO の公募価格や初値に正の影響を与え、その後の価格下落(負の影響)をもたらすことがわかった。

**2. システム不安程度への影響因子探索とその影響パターンの解明<sup>2</sup>**

本研究では、アノマリーと呼ばれる効率的市場仮説で説明できない現象において、小型株効果の解明を目指して分析を実施した。具体的には、小型株効果のリスク要因及び流動性要因との関係について、相関分析やノンパラメトリック確率密度推定ベースの非線形特徴を分析した。その結果、小型株変動の特殊性が小型株効果の変動パターンを説明する主因になっていることを明らかにした。また、投資家の将来価格に対する期待の方向や確信度に応じて変化する「不安程度」に敏感に反応する指標として、小型株の変動パターンや投資家のリスク選好度に注目し、それらの暴落予測への有効性を示した。

**3. 相転移予測による暴落リスク回避性能の向上<sup>3</sup>**

本研究では、金融バブルの生成崩壊過程についての新事実を発見することを目的とし、特に投資家集団が生み出す株式バブルを相転移現象とみなすことで分析を実施した。はじめに、ニューヨーク証券取引所(NYSE)指数を用いて予測に有効な相転移現象特有の統計的パターンの抽出を目指した。具体的には、NYSE 市場全体指数と時価総額群別指数を作成し、それらの確率密度を推定した。その結果、相転移の前兆現象として推定した確率密度形状の非対称度が有効であることが分かった。また、この非対称性に影響を与える投資家のリスク回避度を重要因子として選別した。その後、本研究で発見された各指数の相転移現象特有の統計パターンを利用することで、データから高精度な相転移予測を実現することを目指して研究を進め、機械学習を用いた非対称度などのモーメント指標を使った株価指数のトレンド方向予測により、予測対象の株価時系列のみを入力として長期トレンド反転予測性能を向上させる手法を開発した。この手法は、従来の情報統計技術ベースの手法と比較して、投資ストラテジーにおいてより高い暴落回避性能を示した。

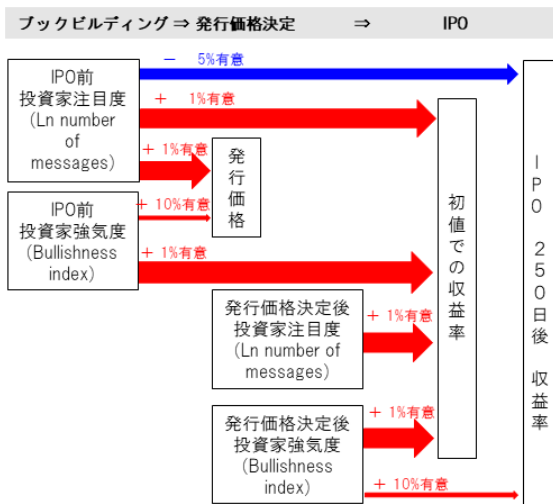


図1. 投資家の過剰期待がもたらすIPO後の収益低下

<sup>1</sup> Investor sentiment extracted from internet stock message boards and IPO puzzles,” International Review of Economics and Finance, 56, 2018.

<sup>2</sup> ”高頻度データを用いたノンパラメトリック小型株効果分析,” 経営研究, 68(4), 2018.

<sup>3</sup> KAKEN「株式市場の対称性の破れ度計測による完全データ駆動型相転移予測法の開発」



## 空間的な情報システムの設計開発支援システム

中西 泰人(慶應義塾大学環境情報学部 教授)

研究期間 2009 年 10 月～2013 年 3 月

### さきがけの成果:

空間的な情報システムを設計開発するにあたり、仮想空間におけるシミュレーションと実空間での模型空間におけるシミュレーションを行き来しながら設計開発を行う CityCompiler を構築し、その有効性を検証した。



### 発展:

#### 1. Processinglue の開発および Processing の書籍の出版<sup>1</sup>

CityCompiler では仮想空間内に配置できる情報出力装置として仮想ディスプレイや仮想プロジェクターを実装し、そのコンテンツとしてデザイナーやアーティストに多く用いられる開発環境である Processing によって開発されたソフトウェアを実行することができた。CityCompiler はゲームエンジンである JMonkeyEngine を下にしているため、シミュレーションを行うには Processing だけでなく JMonkeyEngine の知識が必要であるという課題があった。そこで Processing の 3D 空間内に仮想の情報入出力装置を配置でき、その中のコンテンツとしても Processing を実行することができる Processinglue を開発した。この実装方法を Processing の開発者である Casey REAS に解説する機会を持つことができ、それが Processing のバイブルとも呼ばれる書籍の日本語版の翻訳・出版をする契機となった。(図 1)

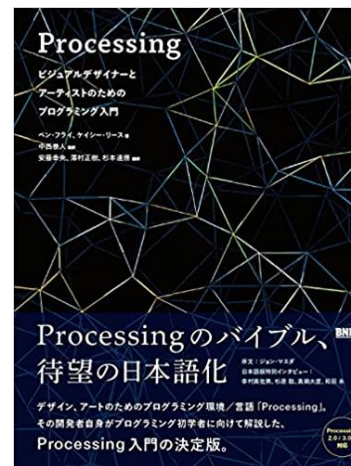


図 1. Processing の書籍

#### 2. デジタルスポーツのためのパラレルプロトタイピング(ゲーム、シミュレーション、ビジュアライゼーションの融合)<sup>2</sup>

空間的な情報システムの応用として、Augmented Reality や Mixed Reality 技術を用いたデジタルスポーツの領域がある。この領域は、実空間でのセンサやディスプレイの配置を仮想空間シミュレーションで検討しながらコンテンツを設計開発するだけでなく、プレイヤー達のデータの可視化や e スポーツとしてのゲームコンテンツの開発も同時並行して開発される可能性がある。CityCompiler では実空間・模型空間・仮想空間の行き来しながら設計開発を行うハイブリッドプロトタイピングを提案したが、本研究ではさらに、実空間・可視化・仮想空間・シミュレーションおよびゲームを行き来しながら設計開発を並行的に進めるパラレルプロトタイピングを提案した。

#### 3. 非人間型ロボットによる街全体のインタフェース化<sup>3</sup>

CityCompiler においては、実空間に配置される情報入出力装置は環境側に固定されることを想定していた。その一方で、実空間の形状や人々の動きを捉えるデプスカメラや Lidar を備えたロボットに情報出力装置を搭載することによっても空間的な情報システムを構築することができる。人間型のロボットではなく、家具型・遊具型のロボットを実空間の要素とすることで、動的でインタラクティブな空間を構築し、街全体をインタフェース化することができる。JST CREST 研究領域「人間と情報環境の共生インタラクション基盤技術の創出と展開」に採択されたプロジェクト「限定合理性を超越する共生インタラクション基盤」の一環として行っている。

<sup>1</sup> 研究室 HP、GitHub Processinglue、Processing: ビジュアルデザイナーとアーティストのためのプログラミング入門 <https://www.amazon.co.jp/dp/4861009502>

<sup>2</sup> "Prototyping for Digital Sports Integrating Game, Simulation and Visualization", Lecture Notes in Computer Science, 8512, 634 - 642 (Springer Verlark)

<sup>3</sup> <https://augmented-behavior-chance.com/>、中西泰人 他、スマートシティとポスト人間中心デザイン、サービソロジー論文誌 Vol.5, No.1 (2021 年 3 月公開予定)

### 3.3 2010 年度採択研究課題

#### 3.3.1 問題構造の解析に基づく組合せ最適化アルゴリズムの自動構成 (梅谷俊治)

#### 問題構造の解析に基づく組合せ最適化アルゴリズムの自動構成

展開している事業:

梅谷 俊治(大阪大学大学院情報科学研究科 准教授)

科研費基盤(B)

研究期間 2010 年 10 月～2014 年 3 月

#### さきがけの成果:

実行時に個別の入力データからアルゴリズムの性能向上に役立つ特徴的な構造を発見し、その構造に合わせてアルゴリズムの設定や構成を自動的に決定する手法を提案し、大規模かつ多様な現実問題に対応できる汎用的な組合せ最適化アルゴリズムを開発した。

#### 発展:

##### 1. 大規模な 0-1 整数計画問題に対する効率的な局所探索法の開発<sup>1</sup>

集合被覆問題や集合分割問題などの組合せ最適化問題を含む整数計画問題において、局所探索法(図 1)の探索空間を削減する手法を提案した。具体的には、局所探索法における近傍探索の走査範囲を「同じ制約条件に同時に現れる頻度の高い変数の組」に絞り込むことで、個別の入力データに合わせて探索を効率化する手法を提案した。また 200 万変数を超える大規模な整数計画問題に対して提案手法を適用し、計算効率の大幅な改善とそれに伴う求解性能の向上を確認した。本手法は人員計画、配送計画、広告配信などの応用事例で現在利用され始めている。また、本手法のアルゴリズムを商用ソフトウェアの新機能として実装する予定である。

##### 0-1 整数計画問題の近傍の例

- ✓ 1-flip 近傍:  $x_j = 0 \rightarrow 1$  (or  $1 \rightarrow 0$ )
- ✓ 2-flip 近傍:  $x_{j_1} = 1 \rightarrow 0, x_{j_2} = 0 \rightarrow 1$

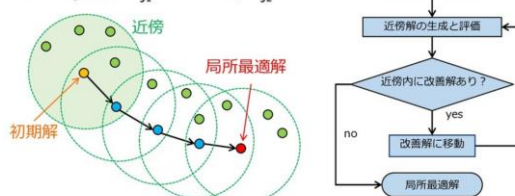


図 1. 局所近傍法

##### 2. 一般化上界制約付き多重被覆問題に対する発見的解法の開発<sup>2</sup>

集合被覆問題は人員計画や配送計画など多くの応用を持つ代表的な組合せ最適化問題である。本研究では、多様な応用事例に現れる追加制約に対応するため、多重被覆制約と一般化上界制約を含む集合被覆問題の効率的な発見的解法を開発した。従来の集合被覆問題ではプライシング法が導入され、大規模な問題を解くために効率的なヒューリスティックアルゴリズムが開発された。しかし、一般化上界制約は評価の高い変数を同時に含む解を妨げることがあり、プライシング法が有効に機能しなかった。そこで本研究では、次の 3 つの取組を通じて、この課題の解決を図った。①一般化上界制約を考慮した変数の新しい評価スキームを導入し、考慮すべきパターン数を小さくするヒューリスティックアルゴリズムを提案した。②解を落とすことなく近傍の候補数を減らす 2-flip 近傍探索アルゴリズムを開発した。③多様な良解への探索を進めるため、得られた複数の解を組合せて新たな解を生成する手法を導入した。これらの手法を既存の手法と比較すると、一般化上界制約のある大規模な組合せ最適化問題でも有効に機能することが確認できた。

##### 3. 劣モジュラ関数の最大化における効率的な分枝カット法<sup>3</sup>

劣モジュラ関数の最大化問題はゲーム理論や機械学習などの幅広い応用が期待される最適化モデルである。しかし合理的な計算時間内に最適解、または十分な近似最適解を見つけるには至っていない。そこで、本研究では 2 値整数計画法(BIP)の定式化に基づいて、劣モジュラ関数最大化問題のための効率的な分枝カットアルゴリズムを提案した。今回開発したアルゴリズムをベンチマークとなる最適化問題に適用してその性能を確認したところ最先端の他のアルゴリズムよりも優れた性能を達成することが確認できた。

<sup>1</sup> "Exploiting variable associations to configure efficient local search algorithms in large-scale binary integer programs," *European Journal of Operational Research*, 263 (2017), 72-81.

<sup>2</sup> "Relaxation heuristics for the set multicover problem with generalized upper bound constraints," *Computers and Operations Research*, 93 (2018), 90-100.

<sup>3</sup> "An Efficient Branch-and-Cut Algorithm for Submodular Function Maximization," *Journal of the Operations Research Society of Japan*, 63 (2020), 41-59.

## 大規模会話データに基づく個別適合型認知活動支援

展開している事業:

大武 美保子(理化学研究所革新知能統合研究センター認知行動支援技術

科研費新学術

チーム チームリーダー)

研究期間 2010年10月～2016年7月

### さきがけの成果:

認知活動支援を目的に、テーマに沿った話題と写真を使って話し手と聞き手が交互にグループ会話する共想法により、会話時認知活動計測評価手法とモデル化手法を開発した。またウェブデータベースと連携した会話支援システムの開発や介護施設等での実証実験を経て、人間の特性に合わせたフィードバックをかける会話支援ロボットを開発した。

### 発展:

#### 1. 会話支援技術と認知行動療法に基づく主体価値発展支援システムの開発<sup>1</sup>

さきがけ研究期間中に開発した共想法支援システムを発展させ、主体価値発展支援システムを開発した。本システムは、ウェブシステムと司会ロボット(図1)からなる会話支援システムと、スマートフォンやタブレットアプリケーションからなる話題発見共有システムで構成され、対面もしくはアプリケーションを介して写真と言葉により視点を共有することで、利用者の主体価値の発展を支援するシステムである。さらに、以前より提唱の高齢者の会話支援手法である「共想法」形式の会話データから主体価値を推定する手法を開発し、共想法における会話データの分析に適用した。



図1. 共想法形式の会話を支援する  
司会ロボット

#### 2. ベッドサイド見守り声掛けロボットの開発<sup>2</sup>

高齢者のベッドサイドにおける転倒の防止を目的として、マルチセンサを用いたベッド上の人の状況認識に基づく意図推定および行動予測手法を開発し、高齢者の離床の予兆推定でナースコールを発報し、高齢者本人への声かけにて会話を誘導することで、看護者や介護者が来るまでの時間を作る方法を提案した。従来の対話システムとは異なり、本研究ではベッドサイドでの転倒事故を防ぐための対話になるように開発した。また本研究では安全性、受容性、実用性の3つの観点について看護師や介護者などの協力も仰いでベッドサイド見守り声掛けロボットを開発した。安全性では、ロボットの外形や大きさを考慮したハードウェアを設計した。受容性では、音声対話の内容などロボットのソフトウェアを複数回の実験を通じて改良した。実用性では、介護施設や病院にロボットを導入して検証した。結果として、ロボットが転倒事故の減少等への寄与が確認できた。

#### 3. 共想法に立脚した対話支援システムの開発<sup>3</sup>

共想法支援システムを発展させ、認知機能の低下を防止するシステムの構築を目的に、ロボットと人間と一対一の対話形式で共想法を実施する対話支援システムを開発した。ロボットは高齢者との話から、質問と回答のペアのデータセットを生成する。そしてそのデータを基に対話型の会話を提供する。本研究では、ロボットに対する高齢者の態度、ロボットのファシリテータとしての機能、ロボットの話を高齢者の方がどれくらい記憶できるのかという点について調べた。結果としてほとんどの高齢者の方が問題なくシステムを利用できることが確認できた。事業化を踏まえた応用検討で、4件の特許を出願し、ライセンス先を開拓した。

<sup>1</sup> "Estimation of Personalized Value through the Analysis of Conversational Data Assisted by Co-imagination Method", The 2018 AAAI Spring Symposium Series, 2018, "Development of a group conversation support robot Bono-05 for cognitive health of older adults.", Gerontechnology, 19(0), 1-1, 2020

<sup>2</sup> "User-Oriented Design of Active Monitoring Bedside Agent for Older Adults to Prevent Falls", Journal of Intelligent & Robotic Systems, Vol. 98, 71-84, 2020

<sup>3</sup> "Cognitive Training for Older Adults with a Dialogue-Based, Robot-Facilitated Storytelling System", Interactive Storytelling, Lecture Notes in Computer Science, Springer-Verlag, Vol. 11869, pp. 405-409, 2019



### 3.3.3 高精度でスケーラブルな多項関係予測の実現 (鹿島久嗣)

#### 高精度でスケーラブルな多項関係予測の実現

鹿島 久嗣(京都大学 大学院情報学研究科知能情報学専攻 教授)  
研究期間 2010年10月～2014年3月

展開している事業:

科研費基盤(A)、科研費基盤(B)、科研費若手(A)

#### さきがけの成果:

物事の中に成り立つ多種多様な関係を表すデータ(人、もの、こと、時間等を結び付けたデータ)を対象に、主にテンソル分解と呼ばれる関係のモデリング技術に着目することで、大規模な関係データに対して高速かつ高精度で予測を行うことのできる手法を開発した。



#### 発展:

##### 1. 複数種類の多項関係の同時予測<sup>1</sup>

さきがけ研究で開発した多項関係予測技術を、さらに複雑な複数種類の関係が混在する場合に拡張した。具体的には、(1)3つ以上のデータ間で成り立つ関係性を捉える多項関係を扱え、(2)オブジェクトの属性情報を活用でき、(3)データ過疎に対して頑強な予測を行えるという3つの要求を満たしながら複数種類の関係を予測できる手法 Collective Incidence Matrix Embedding(CIME)を提案した。これにより、さきがけの研究で開発した手法の適用範囲が、さらに大きく広がった。

##### 2. 接続行列に基づく関係予測<sup>2</sup>

データの関係性の予測はデータの過疎性が大きな課題の1つとなっている。本研究では、疎なネットワークでは、接続行列分解(IMF)の方がこれまで多くの研究で使われてきた隣接行列分解(AMF)よりも優れた性能を発揮するという仮説を立て、それを検証した。また、その際 IMF を採用した最適化ベースの関係性予測手法を新たに開発した。様々なデータを用いた実験によって、仮説を裏付ける結果としてネットワークが疎になるほど IMF の方が AMF よりも優れた性能を発揮することが確認され、複数の AI 領域で重要な役割を持つ疎なネットワークでの関係性予測手法の性能が向上した。

##### 3. ビッグデータ時代の複雑構造データを扱う機械学習法の研究<sup>3</sup>

従来の機械学習では、グラフ構造などの複雑な構造を持つデータの取り扱いは困難であった(図1)。そこで本研究ではグラフ構造データに対する機械学習法を開発した。また近年注目を集める深層学習は、その応用範囲をさらに拡大させ、グラフ等の複雑な構造を持ったデータも扱えるようになりつつある。一方で、モデルが行う判断の根拠が人間には分からないという欠点が存在する為、予測結果に基づく意思決定が

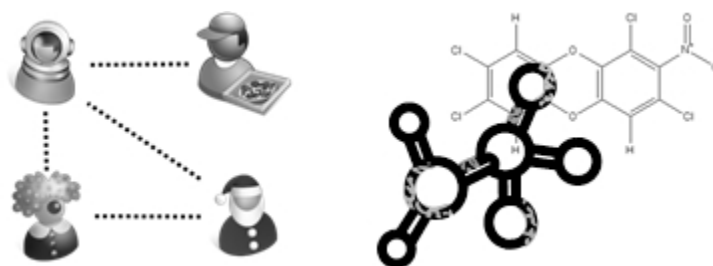


図1. 従来の機械学習では取り扱いが困難なグラフ構造データ

困難になっているという問題が存在する。そこでデータのグラフ構造のどの部分構造が判断の根拠となっているかを提示する頑強な手法を開発し、その有効性を示した。その他にも複数の人間の意見と機械学習を融合することによってよりよい意思決定を行うための統計的手法の開発、機械学習の実応用問題として、深層学習を用いた短期的な気象予測や、テンソル分解による関係分析手法を用いた材料科学における推薦システムへの応用にも取り組む等、人工知能と機械学習に関する幅広い研究成果を得ている。

<sup>1</sup> “接続行列埋め込みに基づく複数種類の多項関係の同時予測,” 人工知能学会論文誌, Vol.30, No.2, pp.459-465, 2015.

<sup>2</sup> 1) “Link Prediction by Incidence Matrix Factorization,” In Proceedings of the 22nd ECAI, The Hague, Holland, 2016, 2) “Link Prediction in Sparse Networks by Incidence Matrix Factorization,” Journal of Information Processing, Vol.25, 2017.

<sup>3</sup> KAKEN「ビッグデータ時代の複雑構造データを扱う機械学習法の研究」

**組合せ論的計算に基づく超高次元データからの知識発見**

展開している事業:

河原 吉伸(九州大学マス・フォア・インダストリ研究所 教授/理化学研究所 CREST、科研費基盤(B)2 件  
革新知能統合研究センター構造的学習チーム チームリーダー)

研究期間 2010 年 10 月～2014 年 3 月

**さきがけの成果:**

効率的に解ける問題(クラス P(多項式クラス)に属する最適化問題)に対して劣モジュラ性を利用して大域的な最適解を見つける効率的なアルゴリズムを開発し、遺伝子データ解析や脳画像データ解析など多くの応用事例へ展開した。また一般に計算困難な問題(クラス NP(非多項式クラス)に属する最適化問題)においては全域最適解のアルゴリズムに加え、多項式時間内に局所解ではなく大域解の一部の情報を計算可能なアルゴリズムを開発した。



**発展:**

**1.疎性モデリングへの組合せ論的アプローチと最適化<sup>1</sup>**

本研究は離散凸性に基づいた疎性モデリングのための組合せ論的な理論的解析方法と最適化を中心とした計算の方法論的基盤の構築を目的として実施した。データ構造情報を用いた疎性モデリングの組合せ論的解析方法において、離散凸性に基づいた理論解析を導入するため、既存の構造正則化を劣モジュラ関数の性質を用いて汎用的に扱う計算手法(高階結合正則化など)を導出し、その高速な最適化方法も開発した。また、得られた手法を、地理情報データや遺伝子データ等に適用することで、実際の応用性という観点においても有用性を確認することができた。

**2.機械学習に基づく汎用的な組合せ最適化アルゴリズムの自動生成<sup>2</sup>**

組合せ最適化の専門知識を持たない利用者が、産業や学術の幅広い分野において日々新たに現実に生じる問題を短期間で最適化するには、これらの問題を整数計画問題などの汎用的な組合せ最適化問題に定式化して高性能なアルゴリズムを適用することが望ましい。しかし、問題の汎用性が高いほどアルゴリズムの性能向上に利用できる特徴的な構造が失われるため、汎用的な組合せ最適化問題に対して常に高速に機能するアルゴリズムの開発は困難である。また、今解くべき問題の構造から適切なアルゴリズムを判断することも困難である。そこで本研究では、個別のデータ入力を基に、アルゴリズムの性能向上に役立つ構造を発見し、アルゴリズムの設定や構成を自動的に決定する手法を提案した。開発した手法での性能を整数計画ソルバ及び、各問題専用のアルゴリズムとの性能を比較したところ、整数計画ソルバの性能を大きく上回り、専用アルゴリズムにも十分に対抗できる性能を達成した。

**3.離散凸解析に基づく機械学習アルゴリズム体系の構築とその応用<sup>3</sup>**

本研究では、劣モジュラ関数(図 1)の最適化を中心とした組合せ的な計算に基づく機械学習のための基礎理論を構築し、これに基づいて厳密性や解釈性の高い機械学習アルゴリズムの体系的開発を実施した。本研究で開発した手法によって従来よりも効率的に厳密解に近い近似化が得られるようになった。また開発した手法の応用的な研究成果として遺伝子データや交通データへの適用、また画像データセットから代表的なサンプルを抽出する問題などの複数の事例に適用し、各事例において十分な有効性を発揮することが確認できた。

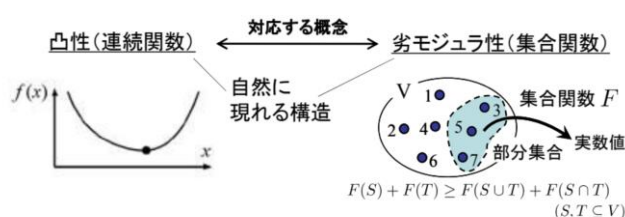


図 1. 劣モジュラ性の概略図

<sup>1</sup> KAKEN「疎性モデリングへの組合せ論的アプローチと最適化」、九州大学 HP

<sup>2</sup> 科研費「機械学習に基づく汎用的な組合せ最適化アルゴリズムの自動構成」研究成果報告書

<sup>3</sup> KAKEN「離散凸解析に基づく機械学習アルゴリズム体系の構築とその応用」、科研費「離散凸解析に基づく機械学習アルゴリズム体系の構築とその応用」研究成果報告書

## Building and Utilising High Precision Semantic Analyses of Unrestricted Natural Language

Alastair Butler (Hirosaki University, Faculty of Humanities and Social Sciences, Associate Professor)

Research Period: 10/1/2010~3/31/2014

### Research achievement of the PRESTO:

The research of natural language processing was conducted for two purposes: to provide computers with the ability to understand the core meaning of text; linking to world knowledge for reasoning. The main results are: 1) Developed a general method for transforming unrestricted natural language texts (Japanese and English) into formal semantic representations, 2) Built a full pipeline for parsing Japanese (HARUNIWA parser), 3) Built a parsed corpus of Japanese (Keyaki Treebank) with full constituent structure (analysis for 20,000 sentences).



### Further research progress based on the PRESTO achievement:

#### 1. Publication of Linguistic Expressions and Semantic Processing<sup>1</sup>

A 172-page book in 2015 based on his research findings from PRESTO was published. The book systematically introduces formal semantics techniques for a natural language processing audience. In particular, it focuses on techniques for judging the appropriateness of linguistic expressions generated by natural language processing from the perspective of whether they are what they should be, as well as so-called "interpretation" techniques, which are used to evaluate the meaning of a sentence with respect to contexts.

#### 2. The extraction and application of visual and semantic information from Japanese treebanks with syntactic analysis annotation<sup>2</sup>

The research aim has been to develop methods of visualizing semantic information from analyses of Japanese and English, notably: connective, predicate, and so on. The key result of this work has been the development of a visualization tool for semantic relationships derivable from corpus data with syntactic parsing information (treebank). This tool has aided to present many relationships found internally to sentences and across discourse providing a way to visualize semantic roles, and anaphoric dependencies, as well as to assess whether their interpretations of sentences have been adequately captured by the annotation. The techniques developed in this research have been used in the construction of corpus such as the Treebank Semantics Parsed Corpus (TSPC) (Figure 1).

Filename	tree count	word count	content description
1 texts_ANC_110CYL067	50	499	Open letter; philanthropic-fundraising, December 1998
2 texts_ANC_HistoryOfGreece	161	2707	Travel-guides: Berlitz History of Greece
3 texts_description	15	333	Miscellaneous examples
4 texts_driver_handbook	53	995	Parts of California Driver's Handbook; Sections: Right-of-way on mountain roads, Sp
5 texts_Duplex	55	1032	Press: reportage: BBC, Technology Reportage
6 texts_FACTBOOK_1-2010	86	1248	THE CIA WORLD FACTBOOK 2010
7 texts_FACTBOOK_2-2010	119	2670	THE CIA WORLD FACTBOOK 2010
8 texts_freedom	46	763	Instruction manual
9 texts_Hedgehog_Conservation	189	4079	Parliamentary language: Hansard
10 texts_IBM-1401	97	1176	IBM 1401 Programming Systems manual
11 texts_Johnson_letter_30_March_2020	30	587	Open letter, March 2020

Figure1. Treebank Semantics Parsed Corpus

#### 3. Developing a program for language teaching with parsed corpora<sup>3</sup>

This research will apply a grammar "toolkit" developed from parsed corpora of English and Japanese to real textual materials for language teaching purposes. This will provide students with grammatical analysis skills for criticizing and improving their own language use.

#### 4. Developing the Keyaki Treebank Parsed Corpus<sup>4</sup>

The Keyaki Treebank is a parsed corpus tagged with logical semantic indications aimed at constructing a coherent description grammar of the Japanese language, allowing for syntactic analysis (parsing). This corpus is utilized in the research of natural language processing in Japanese.

<sup>1</sup> "Linguistic Expressions and Semantic Processing: A Practical Approach," Springer International Publishing

<sup>2</sup> TSPC HP(<http://www.compling.jp/ajb129/tspc.html>), KAKEN "The Extraction and Application of Visual and Semantic Information from Japanese Treebanks with Syntactic and Semantic Analysis Annotation"

<sup>3</sup> KAKEN "Developing a program for language teaching with parsed corpora"

<sup>4</sup> The Keyaki Treebank Parsed Corpus (Version 1.1, 2018) <http://www.compling.jp/keyaki/download.html>



## 計算論的メディア操作の形式化

浜中 雅俊(理化学研究所革新知能統合研究センター音楽情報知能チーム チームリーダー) 科研費基盤(B)、科研費若手(A)

研究期間 2010年10月～2014年3月

### さきがけの成果:

音楽を題材に、メロディを要素とする束(楽曲構造束)を構成し、メディアデザインの操作を束演算で表現することで、専門家の操作事例の蓄積・再利用を可能にするシステムを構築した。また、音楽において音楽理論を拡張し、構造木からなるモデルで音楽構造の解析データを作成する分析ツールを開発した。



### 発展:

#### 1. 音楽音響信号に対する音楽操作の蓄積と再利用の実現に関する研究<sup>1</sup>

音楽は認識や表現が曖昧な為、音楽知識が乏しいユーザが思い通りに計算機に作曲させたり演奏させたりすることは困難である。そこで本研究では音楽家の音楽操作を一般ユーザでも可能とすることを目的として、音楽家がこれまで楽譜に対して行ってきた高次の音楽操作を一般ユーザにも可能とする為に、音楽理論 GTTM(A Generative Theory of Tonal Music)に着目した。GTTMによる分析の結果、タイムスパン木というデータが得られ、楽曲の要約、表情付け、生成、作曲支援などに応用できるが、一般の人には取り扱いが困難である。そこで、本研究では統計的な学習手法に基づくσGTTMIII というタイムスパン木分析手法を開発し、既存の分析手法と比較したところ、最も高い性能を示した。また、Deep Learning に基づく楽曲分析器では、さらに高精度の分析結果が得られた。

#### 2. メロディ操作体験<sup>2</sup>

GTTMに基づく音楽分析では、作曲の逆の過程、すなわち、メロディがどのような手順で作成されたかが解析される。これを用いるとメロディの生成や操作の支援が可能となる。図1はメロディを操作する体験を可能とする Melody Slot Machine である。指でダイヤルを回してメロディを選ぶこともできるし、スロットマシンのようにレバーを倒すとすべてのメロディが回転して、新たな組み合わせのメロディを生成することもできる。



図 1. Melody Slot Machine

#### 3. バンドネットワークを用いた関係性に基づく楽曲検索<sup>3</sup>

本研究ではミュージシャンとバンドの関係性を利用して楽曲を検索する新しい手法を提案した。バンドは、メンバーの変更、解散後の新バンド立ち上げ、他のバンドへのゲスト出演や楽曲提供等を通じて、相互関係がある。このバンドとミュージシャンの関係を線でつなぐことで、複雑なネットワーク構造(バンドネットワーク)が得られる。本研究では、web 情報を元にバンドネットワークを構築し、音楽の検索に応用した。具体的には、現在自分の好きなバンドと関係性が強いバンドをネットワークの数段階先まで推薦するシステムや、関係性に基づいてバンド同士を階層的にクラスタリングすることを実現した。また、収集データを基にバンドネットワークを辿ることで新しいバンドを発見するアプリ、BandNavi を開発、AppStore で公開した(図2)。



図 2. BandNavi

<sup>1</sup> 科研費「音楽音響信号に対する音楽操作の蓄積と再利用の実現に関する研究」研究成果報告書

<sup>2</sup> 研究者Webサイト <http://gttm.jp/hamanaka/music-analysis/>

<sup>3</sup> 研究者 Web サイト <http://gttm.jp/hamanaka/artists-discovery-system>

**高性能ストリーム・コンピューティング環境の構築**

山際 伸一(筑波大学システム情報系 准教授)

研究期間 2010年10月～2014年3月

展開している事業:

科研費基盤(B)3件

**さきがけの成果:**

アクセラレータのプログラムをコマンドラインから直接実行できるメカニズムを実現し、パイプライン処理を構成する、個々のモデルの実行順序と入出力関係を指示することで、自動的に処理を実行できる仕組み(CarSh)を実現した。また、アクセラレータが複数存在する環境において、構成したパイプラインから実行順序を特定し、さらに、パイプラインとしての並列性を抽出する手法(Scenario-based Execution)を開発した。

**発展:**

**1. 高性能ストリームデータ圧縮技術の研究開発<sup>1</sup>**

ハードウェア化が容易で、さらに、データストリームを切れ目なく圧縮、復号が可能な圧縮技術 LCA-DLT を開発した。LCA-DLT では、実装時に利用できるメモリ量を決めることで動的ヒストグラム作成でき、指定したメモリ量に見合った圧縮率を実現することができる。このヒストグラム情報は、圧縮されたデータに暗に埋め込まれて伝送されるため、復号側にデータが到着するとすぐに復号化が可能となる。さらに 2 つのデータペアから 1 つのデータに圧縮するモジュールと、同様の復号化モジュールを複数段組み合わせることで、圧縮率を調整することができ、ハードウェア量を調整することができる。

(図 1)

これにより、ジャイロ、加速度センサからのデータストリームを、2 段のモジュールで 50% まで圧縮できることが示されている。(図 2)

この技術は、ビッグデータを扱う場面でデータ伝送性能を劇的に向上することを可能にし、下記のケースを中心に広範な市場への展開が見込まれる。

- 【ケース 1】社内ネットワークの高速化
- 【ケース 2】LSI データバスの大容量化
- 【ケース 3】センサ・映像等のデータストリームのリアルタイム圧縮伝送

**特記事項**

・2015 年 8 月にストリームテクノロジー株式会社を設立し、LCA-DLT を実用化。

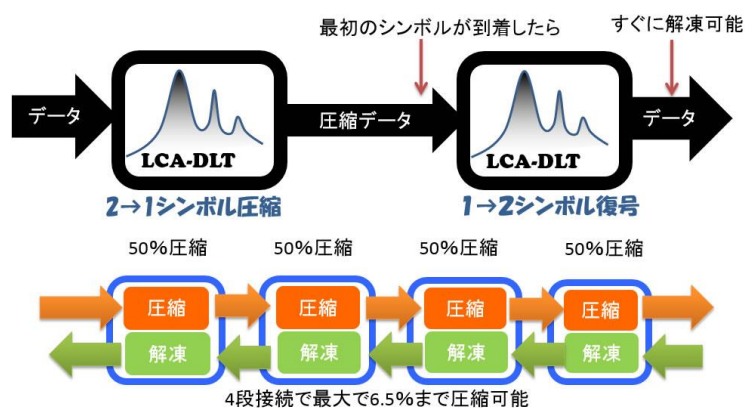


図 1. LCA-DLT の概要<sup>1</sup>

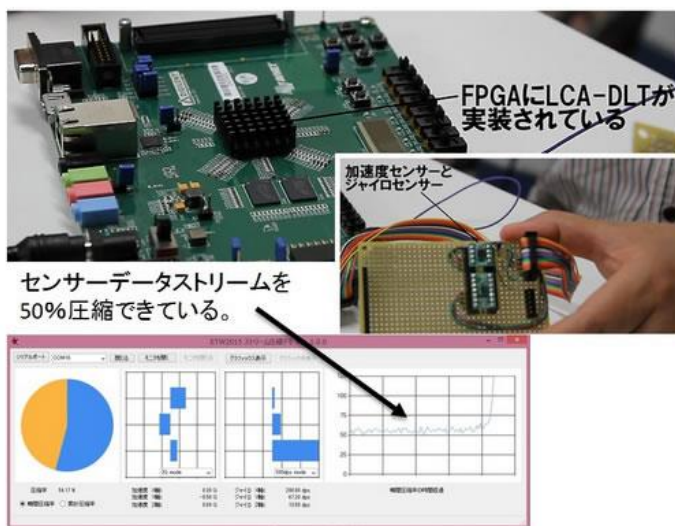


図 2. データストリームの圧縮に関する実演(センサデータ)<sup>1</sup>

<sup>1</sup> (1) "Lazy Management for Frequency Table on Hardware-Based Stream Lossless Data Compression," Information, 7(4), 63, 2016、(2) "Stream-based Lossless Data Compression Hardware using Adaptive Frequency Table Management," In Proceedings of VLDB 2015/ BPOE-6, pp. 133-146, Springer, Sep 2015、(3) 特許第 6256883 号、(4) 筑波大学「山際研究室」HP



**能動センシングによる日用柔軟物の情報知識化とその応用**

山崎 公俊(信州大学工学部機械システム工学科 准教授)

研究期間 2010年10月～2014年3月

展開している事業:

科研費若手(A)、科研費基盤(B)、

NEDO 次世代ロボット中核技術開発

**さきがけの成果:**

布製品の日常的な操作(畳む、着せる等)の自動化を目的として、センサデータの適切な処理方法・利用方法を主に研究し、新たな知識表現法、状態推定法、操作生成法など提案し、ロボットを用いた実験により効果を検証した。

**発展:**

**1. 布の操作計画手法 EMD net<sup>1</sup>**

布製品を所望の形状へ折り畳む作業(例えば、ハンガーにかけられたTシャツをタンスに入れられる形に折り畳む作業)を考えたとき、一度の畳み動作だけではその作業を達成できず、複数回の畳みが求められる場合がある。そのような「手順のある折り畳み作業」を自動化する手法として、EMD net を提案した。EMD net は、布製品の現在形状と目標形状を入力すると、その布製品がとるべき中間形状と、それを実現する畳み操作を同時に生成できる手法である。ニューラルネットワークの新たな構造を示すと

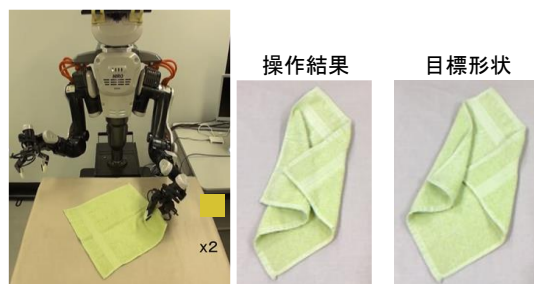


図 1. 手順を伴う折り畳み

ともに、シミュレーションデータを有効に活用する学習方法も提案し、図 1 のように折り畳み作業を実現した。

**2. 布製品の胸当て畳み作業能力の獲得<sup>2</sup>**

ロボットにより布製品を折り畳むことを目的とした従来研究では、水平なテーブルの上に布製品をおいた状態で作業を進めることを想定したものが多かった。一方で我々は、人間がよくおこなう畳み方として、布製品を胸部にあてながら畳んでいく行動に着目した。この畳み方は、テーブルなどの構造物を必要とせず、作業スペースが小さくて済むなどの利点がある。また、布が重力によって自然に垂れ下がるため、布の状態推定が容易である。

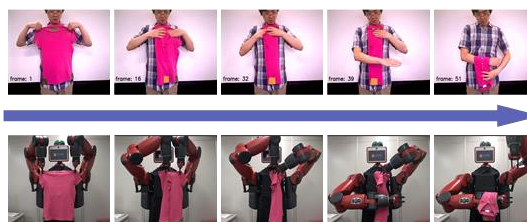


図 2. 双腕ロボットによる胸当て畳み操作

研究のアプローチとして、ロボット自身が試行錯誤することで折り畳み能力を獲得させる方法を採用した。提案手法は主に 2 つのフェーズから構成されている。一つは、作業全体をいくつかの折り畳み工程に分ける分節化フェーズである。もう一つは、各工程における布操作をロボットの試行錯誤によって獲得させる強化学習フェーズである。両フェーズともガウス過程を基礎とする手法で構成することにより、少ない学習回数により作業能力を獲得すること、学習結果を再利用することを可能にした。すなわち、ある布製品の折り畳み作業の経験を別の布製品の折り畳みに利用することで、効率的に行動学習ができる。実験では、人間が折り畳み作業をしている様子の動画を入力すると、Tシャツ、ズボンなどの布製品の胸当て畳みが可能になることを確認した。図 2 に実験の様子を示す。

**3. 不均一性・弾性・異方性を伴う紐状柔軟物の操作能力の確立<sup>3</sup>**

紐やケーブルに代表される紐状の柔軟物は、様々な場面に存在しており、様々な用途がある。もし、紐状柔軟物を揃える・束ねる・はめ込むなどの作業を自動化できれば、様々な役立つと思われる。そこで、自動機械に組み込むための紐状柔軟物の操作に関する知識表現の方法および動作計画方法の確立を目指している。柔軟物の形状を低次元空間での表現に落とし込む方法と、その低次元空間を利用して所望の形状変化を創造する方法を組み合わせ、軽量かつ高速に紐状柔軟物の操作を計画できる方法を提案している。今後は、弾性が高い、または異方性のある紐状柔軟物に対しても、生成された特徴空間を用いて適切な操作手順および操作方法を計画する方法を研究する。

<sup>1</sup> “EMD Net: An Encode-Manipulate-Decode Network for Cloth Manipulation,” IEEE Robotics and Automation Letters, vol. 3, no. 3., 2018.

<sup>2</sup> “人の実演教示の観察に基づく布製品の胸当て畳み作業能力の獲得,” 日本ロボット学会誌, Vol.37(2019)No.6.

<sup>3</sup> KAKEN「不均一性・弾性・異方性を伴う紐状柔軟物の操作能力の確立」