戦略的創造研究推進事業
―CREST(チーム型研究)―

研究領域
「信頼される AI システムを支える
基盤技術」

研究領域中間評価用資料

研究総括：相澤　彰子

2025 年 2 月

目　次

1. 研究領域の概要

(1)戦略目標

「信頼される AI」

(2)研究領域

「信頼される AI システムを支える基盤技術」(2020 年度発足)

(3)研究総括

相澤 彰子(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)

上記詳細は、以下 URL をご参照ください。

JST 公開資料「新規研究領域の事前評価」
https://www.jst.go.jp/kisoken/evaluation/before/index.html
上記 URL の以下
令和 2 年度新規研究領域の事前評価
https://www.jst.go.jp/kisoken/evaluation/before/hyouka_r2.pdf

(4)採択研究課題・研究費

表 1　採択研究課題と研究費　（百万円）

採択年度	研究代表者	所属･役職 採択時[2]	研究課題	研究費[1]
2020 年度	伊藤 孝行	京都大学・教授	ハイパーデモクラシー：ソーシャルマルチエージェントに基づく大規模合意形成プラットフォームの実現	387
	乾 健太郎	東北大学・教授	知識と推論に基づいて言語で説明できる AI システム	304
	越前 功	国立情報学研究所・教授	インフォデミックを克服するソーシャル情報基盤技術	321
	後藤 真孝	産業技術総合研究所・首席研究員（上級首席研究員）	信頼される Explorable 推薦基盤技術の実現	307
	森 健策	名古屋大学・教授	あいまい性を表現する Reliable Interventional AI Robotics	302
2021 年度	鹿島 久嗣	京都大学・教授	人と AI の協働ヒューマンコンピュテーション基盤	288
	高前田 伸也	東京大学・准教授	D3-AI: 多様性と環境変化に寄り添う分散機械学習基盤の創出	310
	竹内 一郎	名古屋大学・教授	AI 駆動仮説の静的・動的信頼性保証と医療への展開	307
	山田 誠二	国立情報学研究所・教授	納得感のある人間-AI 協調意思決定を目指す信頼インタラクションデザインの基盤構築と社会浸透	308
2022 年度	島田 敬士	九州大学・教授	教育大航海時代の羅針盤：学習分析の信頼基盤 ReLAX の創出	298
	清水 昌平	滋賀大学・教授	信頼される AI システムを実現するための因果探索基盤技術の確立と応用	296
	杉山 麿人	国立情報学研究所・准教授	記号推論に接続する機械学習	298
			総研究費	3,729

[1]各研究課題とも研究期間の総額、進行中の課題は予定を含む(2024 年 11 月 1 日現在)

[2]変更/移動のあった場合、下段に括弧つきで記載

2. 研究総括のねらい

(1) 領域設定の経緯、研究領域の位置づけや領域設定を踏まえて、研究総括はどのようにねらいを定めたか

本研究領域がスタートした 2020 年度は、様々な形で AI 技術が社会の中に浸透しつつある時期であった。その中で、AI システムの信頼性・安全性、データ自体の信ぴょう性にかかわる懸念が指摘されはじめ、今後の AI の進化と信頼性確保のための幅広い基盤技術の研究開発が必要という認識のもと、本研究領域の戦略目標である「信頼される高品質な AI(trusted quality AI)」が策定された。具体的には、信頼される AI の実現に向けた基盤技術の創出やそれらを活用した AI システムの構築に関する研究開発行うため、以下の 3 つの達成目標が設定された。

1)「信頼される AI」の実現に向けた発展的・革新的な AI 新技術

2) AI システムに社会が期待する信頼性・安全性を確保する技術

3) 人間中心の AI 社会に向けたデータの信頼性確保及び人間の主体的な意思決定支援技術

これを受けて本研究領域では、**「AI を信頼する主体としての人間」を AI 技術の研究開発の中核とする、新しい AI 研究の在り方を考究し、社会的課題の解決、新たなサイエンス、価値の創造につなげる**ことをねらいとして定めた。これは、内閣府の総合イノベーション戦略推進会議が定めるところの人間中心の AI 社会原則の理念に向けた、AI 技術開発の具体的な道筋を示すためのものと位置付けられる。

(2) 研究領域で実現をねらったこと、研究成果として目指したこと

本研究領域では、人社会の中で幅広く安心して利用できる「信頼される高品質な AI」の実現に向けて、人間中心の AI システムに関する**信頼性や安全性等の定義や評価法の検討への取り組みと、それに立脚した基盤技術の確立および社会実装**を**強く推奨した**。

公募説明会では、前出の 3 つの達成目標について、具体的な研究課題の例として以下を上げ、さらに新規な課題についても積極的に採択するとした(括弧内は対応する採択チーム)。

1)「信頼される AI」の実現に向けた発展的・革新的な AI 新技術

ア　深層学習のような帰納的な処理と知識・言語による推論・プランニング等の演繹的な処理を最適に融合させた AI 技術の研究(→乾チーム、杉山チーム)

イ　大量教師データが与えられなくても、実世界環境との相互作用を通して、知識獲得・成長する AI 技術の研究(→清水チーム)

ウ　人間の脳情報処理や認知発達過程に関する知見に基づく新しい AI 原理の研究(→後藤チーム、乾チーム、山田チーム)

2) AI システムに社会が期待する信頼性・安全性を確保する技術

ア　判断・推論の根拠を説明できる AI システムを実現するための技術の研究(→乾チーム、森チーム、竹内チーム、山田チーム、杉山チーム、島田チームなど)

イ　データ拡張やデータバイアス除去やデータ匿名化などデータを加工する技術の研究(→越前チーム)

ウ　未知・想定外ケースや環境変化にも頑健な AI システムを実現するための技術の研究(→森チーム、高前田チーム)

エ　AI システム全体の信頼性・安全性の確保、品質保証を可能とする技術の研究(→鹿島チーム、高前田チーム、島田チーム)

3) 人間中心の AI 社会に向けたデータの信頼性確保及び人間の主体的な意思決定支援技術

ア　データ改ざんやねつ造(フェイク)等を検知し対処する技術の研究(→越前チーム)

イ　人間が主体性・納得感を持って、適切かつ迅速に判断を下したり合意を形成したりすることを支援する技術の研究(→伊藤チーム、後藤チーム、山田チーム)

また、本研究領域は AI ネットワークラボ(文部科学省の人工知能/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(革新的な人工知能、ビッグデータ、IoT、サイバーセキュリティ等の先導的な基盤技術にかかわる戦略的創造研究推進事業)の研究領域が連携するラボ)の一環として運営されることから、**AIP ネットワークラボへの積極的な参加を通して、AI 技術と人間社会とのかかわり方について幅広い研究者ネットワークを構築すること**を目標とした。特に、同時期にスタートした、さきがけ「信頼される AI の基盤技術」とは、セミナーの合同開催等を通した連携の強化を目指した。

(3)科学技術イノベーション創出に向けて目指したこと、等

本研究領域の発足後に AI の技術が劇的に進展し、特に 2 年目となる 2022 年末に ChatGPT が登場して以降は、生成 AI が急速な勢いで社会に浸透して人々の日常生活から産業構造までを変革の渦に巻き込んでいる。これに伴い、AI が社会に浸透する中で生じる様々な課題についても広く認知されることとなり、AI のリスクに適切に対応して AI の社会受容性を高めることは、喫緊の課題として、教育からビジネス、個々の組織内での実践から国際協調までを含む、あらゆるレベルでの早急なルール作りや制度の確立が議論されるにいたっている。

このような動きは発足時には予想されていなかったことである。これによって、本研究領域が目標として掲げる「信頼される AI」の重要性がますます高まる一方で、いわゆるテックジャイアントによる莫大な投資を背景とした技術開発競争や、国家レベルでの AI ガバナンスに関する政策や国際社会における AI の法規制の動きなど、よりスケールの大きな動きを踏まえた研究計画の策定が必要となった。これを受けて本研究領域では、**AI を情報社会の中核として活用して行くために必須な技術基盤として「信頼される AI」を位置付け、各研究課題の毎年の研究計画においても、AI をめぐる最先端の動きを踏まえた検証と見直しを求める**こととした。

本研究領域では基礎から応用、教育・医療・科学まで、AI 信頼性にかかわる課題をバラ

ンスよく配置しており、各研究課題について、それぞれが設定した研究領域やドメインの中で、AI の社会受容性の向上に結び付く成果を上げることで、研究領域全体として横断的に信頼される AI に向けた道筋を示すと考えている。AI を通して人間社会や人類そのものの在り方が問われる中で、**「人間からの信頼」を評価指標とする AI 技術研究開発が、社会課題解決のための科学研究のパラダイムシフトに向けた先駆的な試みとなる**ことを目指している。

3. 研究課題の選考について

本研究領域の研究期間は 1 期 5.5 年間(1～3 期で 2020 年 10 月から 2028 年 3 月末まで)として、1 研究課題の研究期間全体における研究費は 3 億円(間接経費を除く)を上限とした。また、前掲の 3 つの達成目標について、1 つの課題の解決を目指す構成、複数の課題の解決を目指す構成、いずれも可とした。さらに、実績のある研究者のみならず、若手研究者によるチャレンジングな研究提案も推奨した。選考方針として特に考慮した点は以下の 3 つである。

① **マイルストーンの明確化**：安定したプロジェクト運営を重視して予算は 5 年間で配分した。ステージゲート方式をとらないことから、中間課題評価までのマイルストーンを明確に提示することを採択の条件とした。

② **信頼される AI の定義の明確化**：提案研究が人間中心の AI 社会に資する「信頼される AI」実現に向けて、想定する信頼性の定義および評価法、学術や社会にどのようなインパクトを与えようとするものであるかを明らかにすることを求めた。

③ **情報技術分野にとどまらない幅広いメンバー構成**：情報学分野の研究者だけでなく、倫理、法律および哲学を含む人文・社会科学系の研究者や AI システムのエンドユーザとなる産業界の関係者など、AI に関わる様々な分野・セクターを巻き込んだ幅広いメンバー構成を必要に応じて検討した上で、研究体制がベストチームであること及びその理由を求めた。

選考結果を以下に示す(表 2)。2020 年、2021 年、2022 年の 3 回の公募を通して、応募件数の総計は 128 件、採択数は 12 件、最終採択率 9%となった。なお、2021 年度、2022 年度は日仏共同研究提案をあわせて募集して 8 件の応募があったが、残念ながら採択にいたらなかった。

表 2　応募件数・採択件数

採択年度	応募件数	書類選考採択件数	面接選考採択件数
2020 年度	41	12	5
2021 年度	49	10	4
2022 年度	38	9	3
合計	128	31	12

採択課題は以下のポートフォリオに示す通り(図 1)、システム基盤から社会システム、基礎理論から社会応用を含むバランスのとれた構成となった。また、それぞれのチームが学際的なメンバーで構成されている点も特徴であり、これらによって、人社系に焦点をあてたセミナー企画や、信頼される AI に関するグループ別討議、AI ネットワークラボにおける連携やトラストセミナーへの参画など、戦略目標の達成に向けた領域マネジメントが可能になった。

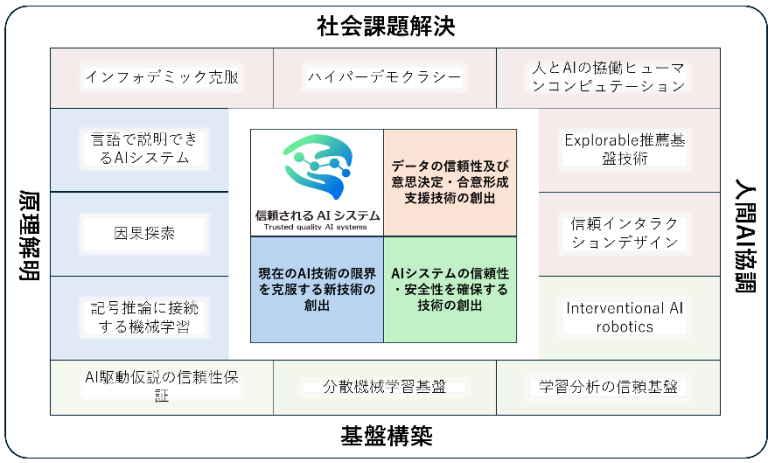


図 1　採択研究課題のポートフォリオ

4.　領域アドバイザーについて

「信頼される AI」の対象となる技術や領域が多岐にわたることを踏まえて、専門分野や経歴(産官学)のバランスを重視しつつ、AI に対する最先端かつ深い知見を持つ 12 名の専門家を領域アドバイザーとして迎えた。特に、技術分野からの 10 名に加えて、AI に関する法律や倫理の専門家として 2 名にご参加を頂いた。これにより採択した課題すべてについて、トピックが近いアドバイザーが配置され、的確なアドバイスが頂ける体制となった。各領域アドバイザーの一覧および専門分野は以下に示す通りである(表 3)。

表3　領域アドバイザー一覧

領域アドバイザー名	専門分野	着任時の所属[1]	役職
岡田 浩之	認知発達ロボティクス	玉川大学工学部 (東京情報デザイン専門職大学)	教授
奥村 学	自然言語処理	東京工業大学科学技術創成研究院 (東京科学大学 総合研究院)	教授
神嶌 敏弘	機械学習やデータマイニングの手法	産業技術総合研究所 人間情報研究部門 (横浜国立大学)	主任研究員 (非常勤講師)
佐藤 洋一	コンピュータビジョン	東京大学生産技術研究所	教授
辻 ゆかり	情報通信	NTT アドバンステクノロジ株式会社 IOWN 推進室 (日本電信電話株式会社)	取締役、室長 (研究開発担当役員 情報ネットワーク総合研究所所長)
福田 雅樹	情報通信法、 情報通信政策、 ELSI	大阪大学 社会技術共創研究センター	教授
福水 健次	知能情報学、統計科学	統計数理研究所数理・推論研究系	教授
村上 祐子	情報哲学	立教大学大学院人工知能科学研究科	教授
盛合 志帆	暗号技術、セキュリティ	情報通信研究機構 サイバーセキュリティ研究所 (情報通信研究機構)	上席研究員 (執行役・経営企画部長)
横尾 真	マルチエージェントシステム	九州大学大学院システム情報科学研究院	主幹教授
若宮 直紀	バイオ情報工学、 バイオシステム解析学	大阪大学大学院情報科学研究科	教授
鷲崎 弘宜	ソフトウェア、情報システム、情報・ソフトウェア・プログラミング教育	早稲田大学理工学術院	教授

[1]変更/移動のあった場合、下段に括弧つき記載

5. 研究領域のマネジメントについて

(1) 研究進捗状況の把握と評価、それに基づく指導

研究進捗状況の把握は、研究計画書および研究報告書の提出に加え、年 2 回(5 月と 11 月)の領域会議、サイトビジット、中間課題評価によって実施している。また年 1 回、秋の領域会議の後に、コメントをフィードバックするための研究総括の個別面談を行い、その結果を研究計画書に反映して頂いている。領域発足当時はコロナ禍のため領域のイベントもオンライン開催を余儀なくされたが、全体会議については 2022 年よりハイブリット形式に移行し、2023 年からは合宿形式となり、主要な参加者の大半がオンサイトで参加している。一方、サイトビジットや中間評価については、多忙な領域アドバイザーのスケジュールに配慮してオンラインで開催し、サイトビジットについては毎回ほぼ 2/3 以上、中間課題評価については 9 割以上の領域アドバイザーの先生方が参加し、十分なフィードバックを頂いている。

本研究領域で特に力を入れているのは、**領域会議中および終了後の領域アドバイザーや参加者に対するアンケートの実施であり、寄せられた意見を踏まえて次回領域会議の立案を行っている**。たとえば 2023 年の領域会議で頂いたコメントと、それに対するアクションを図 2 に示す。特に、「信頼される AI」という戦略目標に対する課題間での認識のばらつきに対する指摘への対応として、2024 年度は 2 回にわたり、研究代表者(PI)と領域アドバイザー混成の 4 グループでの討議および発表のセッションを企画し、後出の 7.(2)でまとめるように、領域全体としての横断的な議論の機会となった。AI システムの品質保証(reliability)のみを目標としていたチームについても、これらのセッションや領域アドバイザーからのフィードバックによって、AI システムの信頼性(trustworthiness)を中心に据えるよう目標設定に発想が転換されつつある。

また、**領域が発足した当時は予想されていなかった生成 AI の急速な発展を受けて、影響を受ける研究課題については積極的に研究計画にフィードバックをかけ、成果の強化もしくは研究計画の追加等の対応**を行った。たとえばエージェントによる対話性能の大幅な向上を受けた追加実装や、この 2 年で急速に進展した大規模言語モデル(LLM)内部構造の解析、クラウドソーシング型データ解析における生成 AI を用いたコンペティションの制度設計と実施などである。これによって最新技術動向を踏まえた機動的な対応が実現できたと考えている。

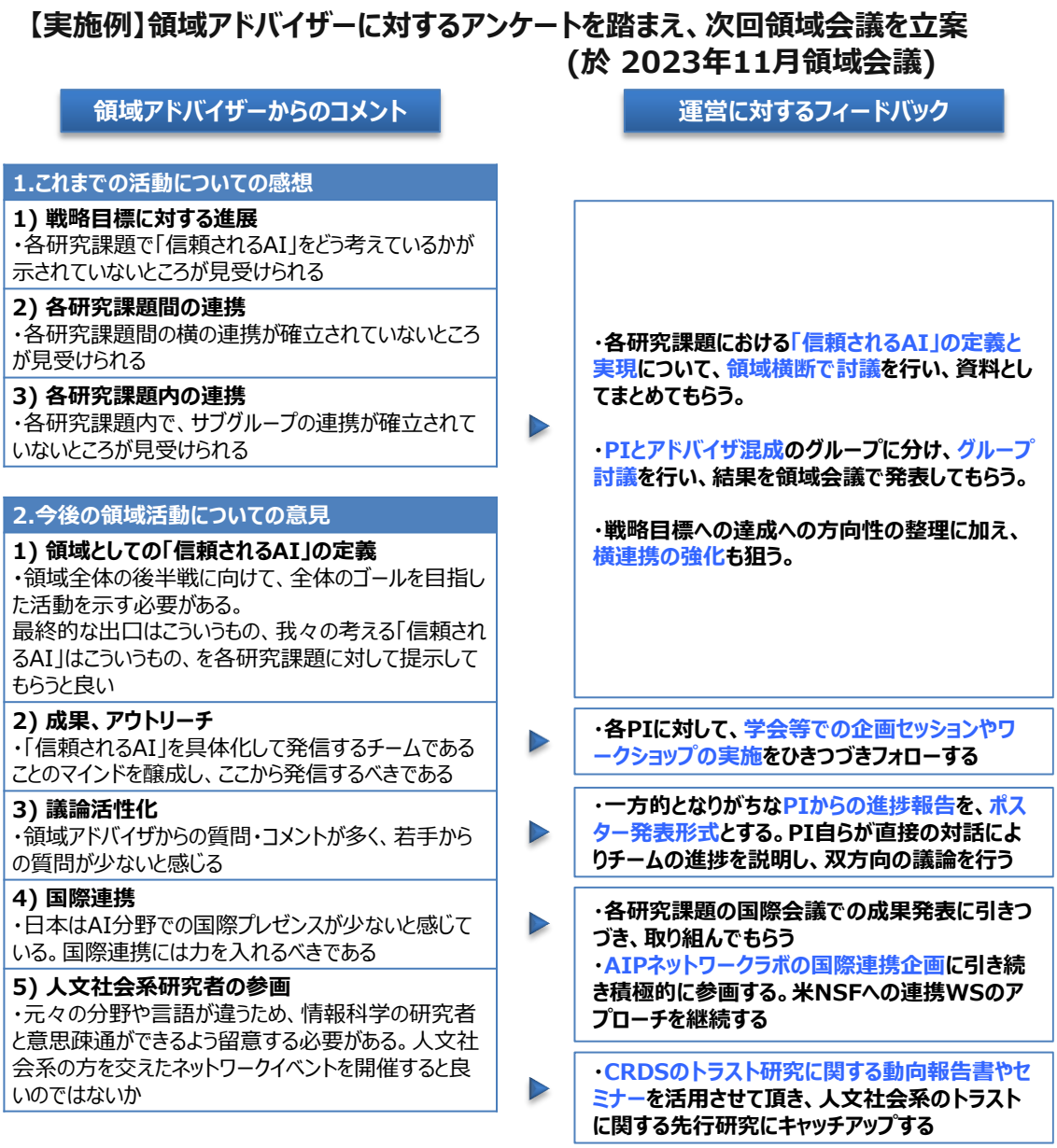


図 2　2023 年 11 月領域会議で領域アドバイザーから頂いたコメントと、
それに対するフィードバック事項

(2) チーム型のネットワーク型研究所として、研究課題間や他の研究領域、国内外の他の研究機関、異分野との融合・連携・協力の推進、新たな研究コミュニティの創成など

本研究領域では、情報技術分野にとどまらない幅広いメンバー構成を求めたことを踏まえ、新規に採択された各チームが持ち回りで、それぞれの研究課題の理解に必要となる基礎知識を領域の他の参加者に紹介するための領域セミナー(チュートリアル)を企画した。さきがけ［信頼される AI］メンバーも参加可能なクローズドセミナーとして、合計 12 回開催し、概ね 20 名から 60 名の参加者であった(表 4)。

表 4　領域セミナー開催実績　　＊内さきがけ参加者

No.	日付	研究課題 Team	タイトル	発表者	出席者数	＊
1	2021/3/19	伊藤 T	民主主義を実験する：信頼されない主体・信頼される決め方 ～市民参加による決定プロセスの含意～	大沼 進 Co-PI	37	4
2	2021/4/26	乾 T	新しい情報環境のための倫理とリテラシーとテクノロジー	久木田 水生 Co-PI	53	4
3	2021/5/27	越前 T	フェイクニュースと情報生態系の進化	笹原 和俊 Co-PI	35	2
4	2021/6/28	後藤 T	音楽の知覚・認知・情動処理に関わる神経生理学	古屋 晋一 Co-PI	46	2
5	2021/7/27	森 T	人の運動を支援するロボット	長谷川 泰久 Co-PI	21	0
6	2022/2/24	鹿島 T	ヒューマンコンピュテーション	鹿島 久嗣 PI	37	7
7	2022/3/22	高前田 T	機械学習のためのコンピュータアーキテクチャ	高前田 伸也 PI	22	2
8	2022/4/25	竹内 T	信頼される AI 診断法の開発に向けて：必要な検証規準	松井 茂之 Co-PI	52	0
9	2022/5/24	山田 T	コンピュータ支援診断から人間-AI 協調診断への展開	原 武史 Co-PI	65	3
10	20230224	島田 T	学習分析の信頼基盤の実現に向けて	島田 敬士 PI	19	1
11	2023/3/9	清水 T	統計的因果探索の概説	清水 昌平 PI	17	3
12	2024/4/11	杉山 T	行列・テンソル表現に基づく記号推論・学習	井上 克巳 Co-PI	27	5

また本研究領域では、AIP ネットワークラボ参加の他の研究領域と協力して国際連携の企画、運営、当日の討議等に積極的に参画し、AI・IoT・サイバーセキュリティ分野における国際ネットワーク拡大に貢献した。具体的には、2022 年の 2nd International Workshop on IAA(IAA+SoC)およびイノベーションジャパン 2022 における参加報告 JST セミナー(図 3)、2021 年から 2024 年まで毎年開催 JST-ERCIM(European Research Consortium on Informatics and Mathematics)ワークショップ(表 5)、2024 年の仏コート・ダジュール大学における特別企画ワークショップに研究総括、領域アドバイザー、PI、主たる共同研究者(Co-PI)や研究参加者などが積極的に参加している。

2nd International Workshop on IAA(IAA+SoC)開催概要
【日時】2022 年 7 月 18 日～20 日
【場所】社会科学高等研究院 コンドルセキャンパス(仏パリ北部オーベルヴィリエ)
【趣旨】AI 時代における人間−AI 共生社会実現の研究課題と方法論について、異なる文化的背景をもつ情報学、コンピュータ科学、脳科学等と社会人文科学(SSH)の研究者がテーマにそって議論をする場を提供する。
【参加者】
当領域:相澤 PO、盛合 AD、越前 PI、伊藤 PI、久木田 Co-PI
AIP: 江村ラボ長、間瀬 PO、栄藤 PO、AD2 名、PI 及び研究者 4 名
EU: 18 名
【内容】
- キーノート講演、招待講演(UNESCO による AI 倫理への取組紹介)(0.5 日)
- ワークショップ課題の話題提供トーク(0.5 日)
- 3 つのテーマに分かれてグループディスカッションおよびその結果発表(2 日)
  パネル 1: AI と脳
  パネル 2: AI と信頼
  パネル 3: サイバーフィジカル AI 社会

イノベーション・ジャパン 2022 JST セミナー開催概要
【タイトル】 コンピュータサイエンスと人文社会科学との交差点
-Intelligence Augmentation and Amplification plus Society (IAA+Soc)2022- 開催報告
【日時】2022 年 10 月 4 日～10 月 31 日(オンライン配信)
【Agenda】(敬称・役職略)
1. IAA+Soc 開催概要 間瀬健二(名大)
2. 挨拶 セバスチャン・ルシュバリエ(EHESS、日仏財団)
3. 参加報告
   パネル 1. AI と脳: 柳澤 琢史(阪大)
   パネル 2. AI と信頼: 盛合 志穂(NICT)
   パネル 3.サイバーフィジカル AI 社会: 住岡 英信(ATR)
4. パネルセッション
   江村克己、相澤 彰子、中野 有紀子、間瀬 健二(司会)
   ・IAA＋soc の全体評価・印象
   ・CS/SSH の異分野、日欧の異文化交流の困難と展望
   ・これからについて

図 3 2nd International Workshop on IAA 及び
イノベーションジャパン 2022 における開催報告

表 5　International JST-ERCIM Workshop 概要

年度	日付	場所	テーマ	参画領域
2021	12/8、9	Online	Accelerating Digital Transformation with Trust for a post-COVID19 Society	・信頼される AI システム ・IoT
2022	10/20、21	Rocquencourt、France	HUMANS AND INTERACTIONS ININTELLIGENT AND XR ENVIRONMENTS	・信頼される AI システム ・IoT ・ICT 基盤強化
2023	10/3～5	京都	Exploring New Research Challenges and Collaborations in AI/BD/HCI/IoT	・信頼される AI システム ・さきがけ信頼される AI ・共生インタラクション ・S5 基盤ソフト ・IoT ・ICT 基盤強化 ・社会変革基盤 ・AIP 加速
2024	10/17、18	Budapest、Hungary	Exploring New Research Challenges and Collaborations in AI/BD/HCI/IoT	・信頼される AI システム ・さきがけ信頼される AI ・S5 基盤ソフト ・IoT ・ICT 基盤強化

また、2023 年 6 月より米国国立科学財団 NSF(National Science Foundation)との連携を模索し、2024 年には PI と Co-PI の 2 名が米国カーネギーメロン大学で開催されたイベント(AI Institutes Expo Day 2024)に参加した。結果としては NSF で適切なパートナーを見つけることは困難との判断になったが、この経験を生かして国際連携の模索をさらに継続することとした。現在、JST 国際部よりカナダ先端研究機構 CIFAR(Canadian Institute For Advanced Research)が AI 分野で JST の研究プログラムとの協力関係を検討している旨の情報を頂いており、双方の関心領域が合えば、協業を検討したいと考えている。

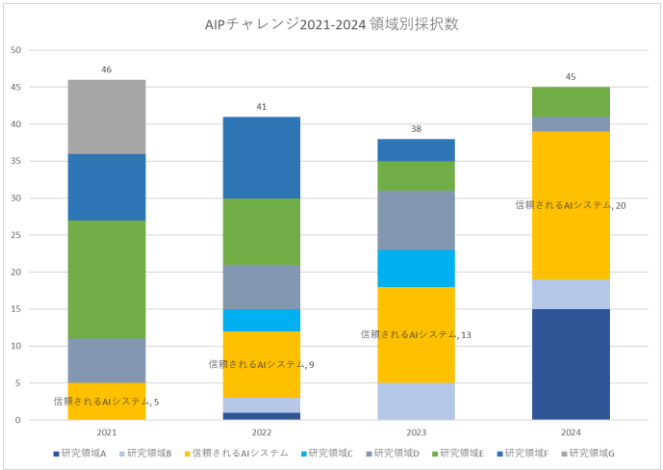
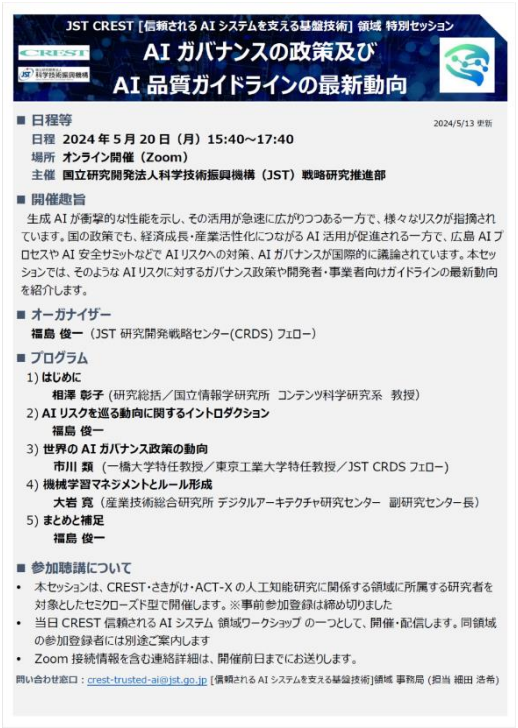


図 4　AIP チャレンジ 2021～2014 領域別採択数

AIP ネットワークラボで若手研究者対象に毎年募集がある AIP チャレンジついては、CREST に参加する若手研究者が自主的に研究に取り組む貴重な機会として、毎年積極的な応募を呼びかけ、この 4 年間で最多採択領域となっている(図 4)。


JST CREST［信頼される AI システムを支える基盤技術］領域 特別セッション

AI ガバナンスの政策及び
AI 品質ガイドラインの最新動向

■ 日程等

2024/5/13 更新

日程 2024 年 5 月 20 日（月）15:40～17:40
場所 オンライン開催（Zoom）
主催 国立研究開発法人科学技術振興機構（JST）戦略研究推進部

■ 開催趣旨

生成 AI が衝撃的な性能を示し、その活用が急速に広がりつつある一方で、様々なリスクが指摘されています。国の政策でも、経済成長・産業活性化につながる AI 活用が促進される一方で、広島 AI プロセスや AI 安全サミットなどで AI リスクへの対策、AI ガバナンスが国際的に議論されています。本セッションでは、そのような AI リスクに対するガバナンス政策や開発者・事業者向けガイドラインの最新動向を紹介します。

■ オーガナイザー

福島 俊一（JST 研究開発戦略センター(CRDS) フェロー）

■ プログラム

1) はじめに
相澤 彰子（研究総括／国立情報学研究所 コンテンツ科学研究系 教授）
2) AI リスクを巡る動向に関するイントロダクション
福島 俊一
3) 世界の AI ガバナンス政策の動向
市川 類（一橋大学特任教授／東京工業大学特任教授／JST CRDS フェロー）
4) 機械学習マネジメントとルール形成
大岩 寛（産業技術総合研究所 デジタルアーキテクチャ研究センター 副研究センター長）
5) まとめと補足
福島 俊一

■ 参加聴講について

- 本セッションは、CREST・さきがけ・ACT-X の人工知能研究に関係する領域に所属する研究者を対象としたセミクローズド型で開催します。※事前参加登録は締め切りました
- 当日 CREST 信頼される AI システム 領域ワークショップ の一つとして、開催・配信します。同領域の参加登録者には別途ご案内します
- Zoom 接続情報を含む連絡詳細は、開催前日までにお送りします。

問い合わせ窓口：crest-trusted-ai@jst.go.jp [信頼される AI システムを支える基盤技術]領域 事務局 (担当 細田 浩希)

また「信頼される AI の定義」討議にあたり、JST 研究開発戦略センター(CRDS)福島フェローの協力を得て、有識者による AI ガバナンスの政策および AI 品質ガイドラインの最新動向の特別セッションを企画し、さきがけ研究者参加可能のセミクローズドのセッションとして開催した。また、CRDS 福島フェローより、トラスト(信頼)に関わる JST の 3 プログラム(CREST［信頼される AI システム］、さきがけ［信頼される AI］、JST 社会技術研究開発センター(RISTEX)［デジタルソーシャルトラスト］)に向けて、横断的な議論を深め、研究開発をさらに発展させるために、広い分野で取り組まれてきた先行研究の概観を短期集中で俯瞰するセミナーの提案があり、領域にも広く案内した。本セミナーの資料及び録画閲覧について、特別にご承諾を頂き、領域会議におけるグループ討議のベースとして活用した。

JST CREST/さきがけ/RISTEX トラスト関連3プログラム向け

短期集中「トラスト研究俯瞰セミナー」全3回

回		講演者	トピック
第1回：トラストのモデル、異なる側面 **7/24(水)15:00～** 各講演は20分程度、質疑・議論を最後にまとめて30分＋(最長17:00)	1	**福島俊一** JST CRDS フェロー	トラストに関するCRDS での俯瞰的調査と戦略提言の紹介
	2	**中島震** 東京大学大学院情報学環 特任教授／国立情報学研究所 名誉教授	トラストを構成するシステムの品質指標
	3	**大屋雄裕** 慶應義塾大学法学部 教授	トラストの3 側面と信用・信頼・信託
第2回：トラスト研究の系譜 **7/30(火)15:00～** 各講演は20分程度、質疑・議論を最後にまとめて30分＋(最長17:00)	4	**小山虎** 山口大学時間学研究所 准教授	人文社会分野におけるトラスト研究の系譜
	5	**村山優子** 津田塾大学数学・計算機科学研究所 特任研究員／岩手県立大学 名誉教授	情報科学分野におけるトラスト研究の系譜
	6	**松本泰** NPO 日本ネットワークセキュリティ協会 フェロー	デジタルトラストの動向
第3回：ガバナンスおよび総合討論 **7/31(水)15:00～** (7)10分、(8)20分、(9)60分＋(最長17:00)	7	**福島俊一** JST CRDS フェロー	第1回・第2回の簡単な総括
	8	**工藤郁子** 大阪大学社会技術共創研究センター 特任准教授	トラストとガバナンス
	9		総合討論

CRDS 12

図 5　CRDS 福島フェローオーガナイズによる特別セッション(上図)及びセミナー(下図)

本研究領域では、各課題に積極的に学会での企画セッションや、ワークショップ開催、ジャーナルの特集号への投稿を通して、成果のアウトリーチを行うことを奨励して、年度末に報告シートの提出を求めている。2023 年度までの実績では、シンポジウム 77 件、学会誌等の特集 6 件、企画セッション・ワークショップ等 137 件、その他 35 件と活発に企画を行っている。また、2024 年度のトピックとして、日本で開催される人工知能研究者の最大級の集まりである、人工知能学会全国大会(5 月、浜松)において、4 つの研究課題がオーガナイズドセッションを行い(表 6)、存在感を高めた。

表 6　2024 年度 人工知能学会全国大会(2024/5/28～31 於 アクトシティ浜松)における本研究領域からのセッション

プログラム	番号	講演/セッション名	チーム/役割	講演者/オーガナイザ
オーガナイズドセッション	OS-1	計算社会科学	笹原 G（越前 T） ＜領域外研究者と共同＞	鳥海 不二夫（東京大学） 榊 剛史（株式会社ホットリンク） 笹原 和俊（東京工業大学） 瀧川 裕貴（東京大学） 吉田 光男（筑波大学）
	OS-5	ヒューマン・イン・ザ・ループ AI	鹿島 T	荒井 ひろみ（理研 AIP） 小山 聡（名古屋市立大学） 鹿島 久嗣（京都大学） 堤 瑛美子（東京大学） 森 純一郎（東京大学）
	OS-6	信頼と文脈のインタラクションデザイン	山田 T	寺田 和憲（岐阜大学） 今井 倫太（慶應義塾大学） 山田 誠二（国立情報学研究所）
	OS-8	AI とデモクラシー	伊藤 T	白松 俊（名古屋工業大学） 伊藤 孝行（京都大学） 大沼 進（北海道大学） 松尾 徳朗（産業技術大学院大学）

(3) 研究費配分上の工夫など

待遇面での課題などから人材確保の困難が予想されたことから、プロジェクトの安定運用を重視し、研究費については採択時に評価に基づき適切と思われる額を配分した。このとき採択後の予算は、各研究課題の成果を最大化するため、研究総括が管理する留保分は、合計 1,592 万円(直)と、成果展開や外部機関との連携に必要な最低限の金額としている。このため増額が必要となった課題に対しては、年 3 回程度実施される CREST の予算見直しに、研究総括による優先度付けをした上で申請している。過去 4 年間で総額 1 億 1,120 万円の増額を承認頂いたことは、各課題の大きなインセンティブとなっている。

(4) その他マネジメントに関する特記事項(人材育成等)

各 PI とも、チームの若手研究者に対し、キャリアパス支援のため、昇進、受賞、外部研究予算の獲得のための協力・指導に力を入れている。その結果、多くの若手研究員が CREST 研究の期間中に昇進や上位職への異動を実現している。

なお、本研究領域の領域アドバイザーや研究代表者は、竹内 PI がさきがけ領域研究総括を務めるなど JST の他のプログラムでも重要な役割を担って活躍しており(表 7)、戦略的国際共同研究プログラム(SICORP)、経済安全保障重要技術育成プログラム(K-Program)、ムーンショットの代表者や他の CREST 課題の共同研究者などをつとめる研究者も多く、さらに分野を横断する研究ネットワークが広がっている。

表 7　他の JST プログラムとの兼任を持つ研究代表

研究代表	関係する JST プログラム
後藤 真孝	・創発的研究支援事業 後藤パネル 創発 PO ・ACT-X[次世代 AI・数理情報]「次世代 AI を築く数理・情報科学の革新」領域アドバイザー
鹿島 久嗣	・ACT-X[次世代 AI・数理情報]「次世代 AI を築く数理・情報科学の革新」領域アドバイザー ・CREST[バイオ DX]「データ駆動・AI 駆動を中心としたデジタルトランスフォーメーションによる生命科学研究の革新」領域アドバイザー
竹内 一郎	・さきがけ[研究開発プロセス革新]「AI・ロボットによる研究開発プロセス革新のための基盤構築と実践活用」(2024 年度発足) 研究総括
山田 誠二	・CREST[海洋カーボン]「海洋と CO2 の関係性解明から拓く海のポテンシャル」「イメージングと AI で紐解く南大洋の炭素循環」主たる共同研究者

6.　研究領域としての戦略目標の達成に向けた状況について

(1)【研究課題 1】

ハイパーデモクラシー：ソーシャルマルチエージェントに基づく大規模合意形成プラットフォームの実現

研究代表者：伊藤孝行(京都大学・教授)

① 研究の概要

ソフトウェアエージェントと人間が一緒に参加するソーシャルネットワークでの民主主義(ハイパーデモクラシー)のための合意形成プラットフォームの実現に取り組んでいる(図 6)。具体的には、ソーシャルネットワークプラットフォームの中に複数のエージェントを常駐させ、これらが人間の代理として働き、意思決定やインタラクションを仲介し、より良い合意形成や集団意思決定を支援する基盤を構築する。また、現実的なフィールドで社会実装を行うことにより、AI を用いたシステムの社会的な受容性や信頼性の向上を実証する。

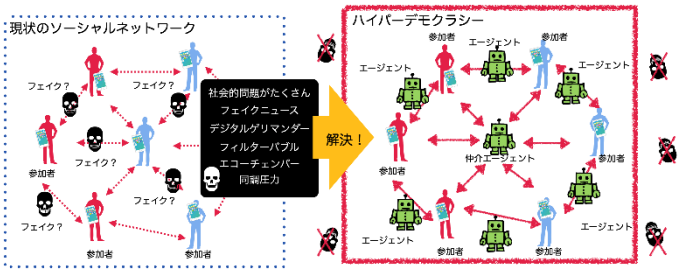


図 6　ハイパーデモクラシープラットフォームの実現

② 独創的で国際的に高い水準の研究成果

2021 年 8 月にアフガニスタンのカブールから米軍が撤退した際には、協力企業と力をあわせて、大規模議論支援システムの運用を続けた。ユーザの名前を秘匿とするなど、投稿する人への安全面を第一に対応し、市民の多くの実際の声(状況への批判、相談事、困っていること、など)を収集し、得られた情報を国連の組織(国際連合-住民居住計画 UN Habitat)にも提供した。これらの活動は新聞等でも報道されるなど話題となった(図 7)。

The Asahi Shimbun GLOBE+
World Now (2021/8/29)

「女性の絵消した」「タリバンを拒絶」…アフガニスタン人の本音　日本の IT 会社が公開

https://globe.asahi.com/article/14421938

図 7　The Asahi Shinbun Globe+ の記事

伊藤チームでは更に、構築したハイパーデモクラシープラットフォームを用いて、インドネシア、アフガニスタンにて 20 名～200 名規模の 4 つの社会実験を実施している。アンケート分析を通して複数エージェントが議論に参加することによる人間の議論の意見の変容を分析、ケーススタディとして民族間の偏見(バイアス)と集団間不安が、会話エージェントが議論を促進すると

減少し、促進しない場合は増加することを示した。これは、AI エージェントによるフェデレーションの効果を実証的に示す成果として重要である。また、集団討議における議論と討議の指標を可視化する指標の開発と社会実験による有効性検証にも取り組んでいることは、社会学分野からの新しい試みとして注目される。

ハイパーデモクラシーに関する上記の成果は人間と AI による意思決定支援は独創的かつ先駆的研究として注目され、英 Google DeepMind や米ハーバード大学などの研究チームの最新の論文でも引用されるなど、国際的にも高い評価を得ている。また、アウトリーチ活動として、国際ワークショップ「Democracy and AI」の 3 回連続開催や、COMPSOC 競技会(Computational Social Choice Competition)の設立(世界から 17 チーム参加)、電子情報通信学会研究会「合意と共創」設立などにより、ハイパーデモクラシーの新しいコミュニティ創出に向けた活動に取り組んでいる。

③ 新技術シーズへの展開

研究成果活用企業として伊藤らが設立したスタートアップ AGREEBIT 株式会社では、AI ファシリテーション・プラットフォームとして国内初・世界唯一の SaaS である「D-Agree」を活用し、自治体、教育機関、大手企業などで利用を拡大し成果事例を積み上げている(図 8)。さらに白松らは、行政や市民活動の効率化と質の向上を図るシステムの社会実装を目的とした株式会社ソシアノッターを設立している。

図 8　D-Agree の成果事例

(2)【研究課題 2】

知識と推論に基づいて言語で説明できる AI システム

研究代表者：乾健太郎（東北大学・教授）

① 研究の概要

現在の end-to-end の学習では解決が難しい「言語」の問題、とくに、人間であれば判断の過程や理由を言語で説明できる種類の問題に焦点を当て、人が説明するのと同様の仕方で判断を説明できる新しい計算パラダイムの設計・実現に取り組んでいる。さらに、人とのコミュニケーションを介して説明を更新しながら人の判断を支援する AI システムを構築し、それを通してコミュニケーションとしての説明の要件を明らかにすることにより、判断の過程を言語で説明できる AI システムの設計論の確立を目指している。(図 9)

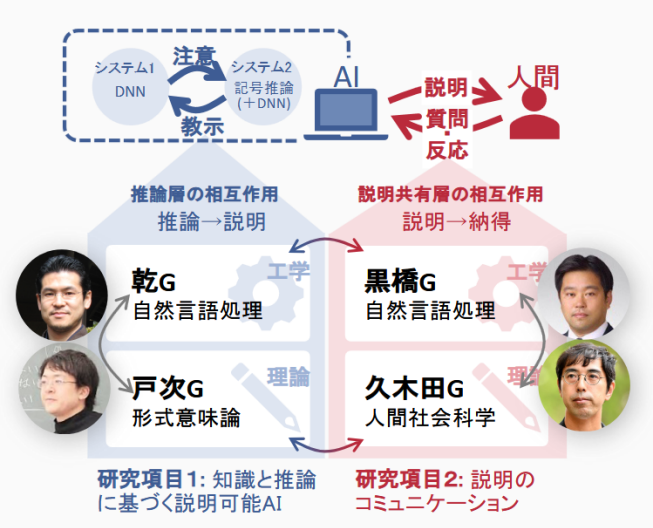


図 9　乾課題の研究テーマ構成と分担

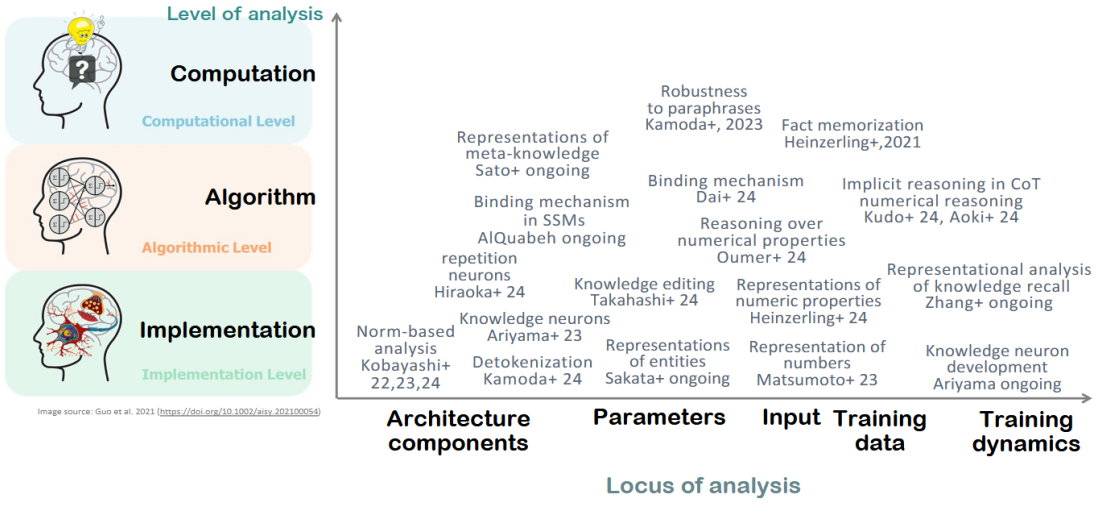


図 10　LLM の研究の体系

② 独創的で国際的に高い水準の研究成果

LLM の急速な進歩を受けて、LLM における知識の内部表象、推論の内部機序の解明に関する課題を新たに設定し、モデル内部で世界知識がどのように記憶されているか、プロンプト

により制御される推論とモデル内部の実際の推論の対応関係の分析、LLM のブロックごとの役割分析など、LLM の原理解明等に寄与する研究を体系的に推進し、学術的に顕著な成果を上げた。このことは自然言語処理・機械学習分野の中核的な国際会議における数多くの研究発表に裏付けられている(図 10)。また、高階論理と DNN の接続について、Neural DTS と呼ぶプロトタイプの実装を進めた。記号とニューラルネットワークの融合は、現在の AI における中核的なチャレンジの 1 つであり、記述力の高い高階論理の証明器を緻密に設計して組み合わせるアプローチは、独創性の高い学術的成果である。

さらに、社会学的なアプローチとして、誤情報とファクトチェックに関する心理と行動の研究に取り組み、一連のオンライン実験を通して、自分が信じている誤情報を打ち消すファクトチェック記事を選択的に回避する人が予想以上に多いことを明らかにするとともに、記事の提示順序や提示方法などの介入が誤情報の受容や信念更新に好影響を与える可能性を示した。AI による「説明」を人間が理解・納得して受け入れるには、どのような条件が重要であるかを、幅広い人間社会科学の記述的研究と規範的研究の両面から分析した重要な成果である。

③ 新技術シーズへの展開

企業との共同研究により、産業情報ドメインにおける知識ベース構築技術や、提案手法である Neural DTS の実装に取り組んでいる。また、市民のリテラシー向上のため、生成 AI の ELSI や誤情報等の課題に関する研究会や公開セミナーを定期的に実施している。

(3)【研究課題 3】
インフォデミックを克服するソーシャル情報基盤技術
研究代表者：越前功(国立情報学研究所・教授)

① 研究の概要

AI により生成されたフェイク映像、フェイク音声、フェイク文書などの多様なモダリティによるフェイクメディア(FM)を用いた高度な攻撃を検出・防御する一方で、信頼性の高い多様なメディアを積極的に取り込むことで人間の意思決定や合意形成を促し、サイバー空間における人間の免疫力を高めるソーシャル情報基盤技術を確立することを目的とする。具体的には、(1)多様なモダリティによる高度な FM 生成技術、(2)FM 検出・防御技術、(3)FM 無毒化技術、(4)インフォデミックを緩和し多様な意思決定を支援する情報技術について研究を推進している(図 11)。

② 独創的で国際的に高い水準の研究成果

研究面では、人間や AI が識別できない精巧な FM の生成とそれを検出し無毒化する要素技術を確立した。また、偽情報やヘイトの拡散が社会的分断へ与える影響を分析し、SNS ユ

ーザが誤情報を自発的にファクトチェックする行動の特徴を計算社会科学の手法を用いて解明した。査読付きジャーナル論文 49 件、査読付き国際会議論文 123 件など顕著な学術的成果に加え、メディア掲載 204 件、招待・依頼講演 58 件、書籍著者・解説記事 35 件など社会的にも大きな注目を集める成果を上げ、BTAS/IJCB 5-Year Highest Impact Award(IEEE Biometrics Council awards)を 2 年連続受賞、ドコモ・モバイル・サイエンス賞社会科学部門優秀賞受賞など 11 件を受賞した。

③ 新技術シーズへの展開

特に、FM 検出では、フェイク顔映像を対象とした自動検出プログラム(SYNTHETIQ VISION)を開発し、大手企業や公共団体を含むパートナー企業への事業ライセンスを開始している。また、生成 AI による偽情報への対策について、G7 内務・安全担当大臣会合での講演を含め各省庁から講演・意見交換を要請され、さらに 4 つの国プロへの採択が決定するなど、新たな研究開発の潮流を創出している。

図 11　越前課題の取り組み

(4)【研究課題 4】

信頼される Explorable 推薦基盤技術の実現

研究代表者：後藤真孝(産業技術総合研究所・上級首席研究員)

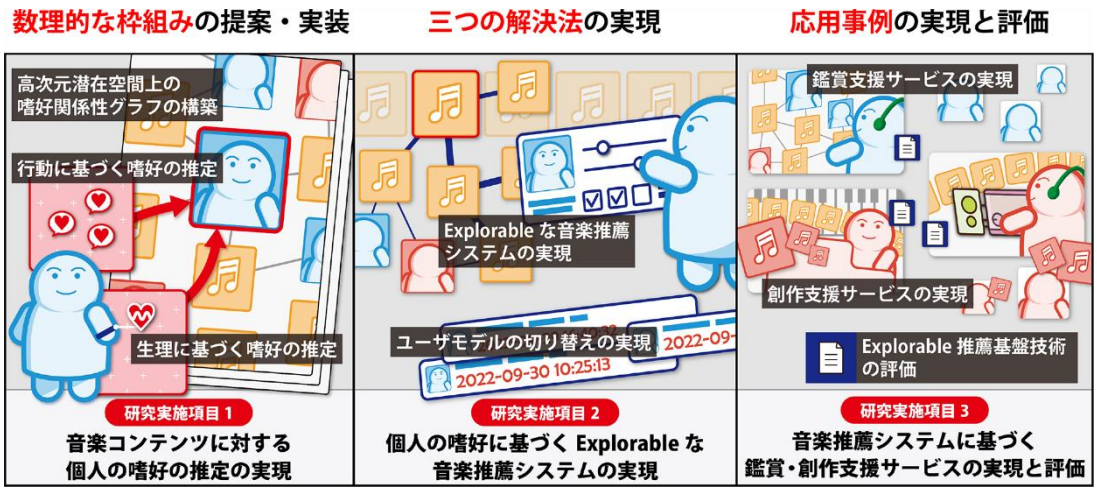


図 12　後藤課題の研究テーマ構成

① 研究の概要

AI システムによる個人に最適化された支援を人々が安心して受けられる未来社会の実現に向けて、推薦システムのユーザが推薦の挙動を探索できる基盤技術を研究開発し、音楽コンテンツを主な対象として、人間中心に制御できる透明性の高い推薦システムを提供可能にすることを目的としている。そのために、情報学、神経生理学、社会心理学の 3 グループが連携した学際的な研究を進め、基礎研究としての新技術開発と、応用研究としての社会実装を同時並行的に推進している点が特長である(図 12)。

② 独創的で国際的に高い水準の研究成果

独自のレイヤー構造に基づく推薦の数理的枠組みを提案して Explorable な音楽推薦システムを開発して実サービスとして運用している。また、音楽嗜好推定のための生理計測・脳機能計測技術を研究開発し、前者で音楽推薦が可能なプロトタイプシステムを実現した。さらに、推薦受容傾向の新たな心理尺度の開発にも取り組んでいる。一連の成果は学術的なインパクトが高い独創的なもので、市村賞 市村学術賞 貢献賞、電気通信普及財団賞(テレコム学際研究賞)入賞、文部科学省 NISTEP「ナイスステップな研究者」など多数の受賞、報道に結実した。

③ 新技術シーズへの展開

企業と連携して公開中の音楽サービス上で提案開発技術を実装し、ユーザが推薦の挙動を容易に変更可能な機能を実現した。それを発展させ、推薦の内部状態を可視化して透明性

を高くするインタフェースを備えた世界初の音楽発掘サービス「Kiite World」(https://world.kiite.jp)を立ち上げ、2023 年 7 月に企業と JST との共同プレス発表とともに公開し、実証実験を継続的に進めている。Kiite World はエンドユーザからの支持を集め、その上でユーザによるイベントも頻度高く開催されるなど、満足度の高いサービスとなっている。その他、音楽発掘カフェ「Kiite Cafe」、嗜好を自己分析できる「Kiite Report」、リリックビデオ制作支援サービス「TextAlive」などを継続的に研究開発して実証実験を進め、さらに他のサービスへの技術提供なども行っている。ユーザからのリアルなフィードバックに基づくシーズ発掘と開発のサイクルをまわしつつ、学術的に独創性の高い基礎研究に結びつた成功例として注目される(図 13)。

図 13　後藤課題の取り組みと成果

(5)【研究課題 5】

あいまい性を表現する Reliable Interventional AI Robotics

研究代表者：森健策(名古屋大学・教授)

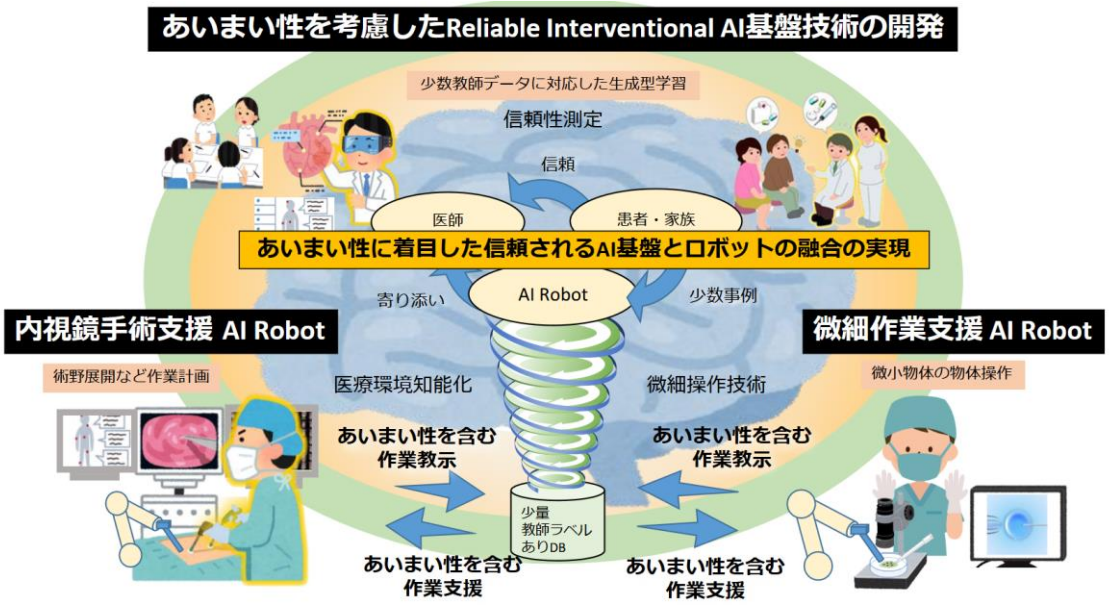


図 14　森課題の研究テーマと成果

① 研究の概要

あいまい性に着目した信頼される AI 基盤とロボットの融合の実現に向けて、あいまい性を考慮した Reliable Interventional AI 基盤技術の開発、内視鏡手術支援 AI Robot、微細操作支援 AI Robot、の 3 つの研究グループを設定して研究を推進している(図 14)。

② 独創的で国際的に高い水準の研究成果

基盤技術として、2 次元の腹腔鏡術中画像から 3 次元的な深度情報を高精度推定する手法や腹腔鏡画像のセグメンテーションにおける弱教師あり学習などで顕著な成果を上げるとともに、人間とロボットとのコミュニケーションにより信頼性の向上を目指した内視鏡手術ロボットシステムの開発と評価、あいまい性を含む環境における自律的なロボット支援提供に向けた手術助手の特性解析とモデリングや、理想軌道推定の信頼性に応じた「あいまい」な動作誘導による微細操作支援など、あいまい性に着目した医療ロボットの基盤実現に向けて成果を上げている(図 15)。

③ 新技術シーズへの展開

生殖補助医療を支援するシステムについて特許出願を行うとともに、顕微鏡メーカー、医科大学、研究所と名古屋大学の 4 機関合同の共同研究開発プロジェクトの準備を進めている。また、一般市民(日本医学会総会市民講座、学士会会員)や医師向けの講演、学会におけ

るオーガナイズドセッションの企画などで社会還元を行っている。

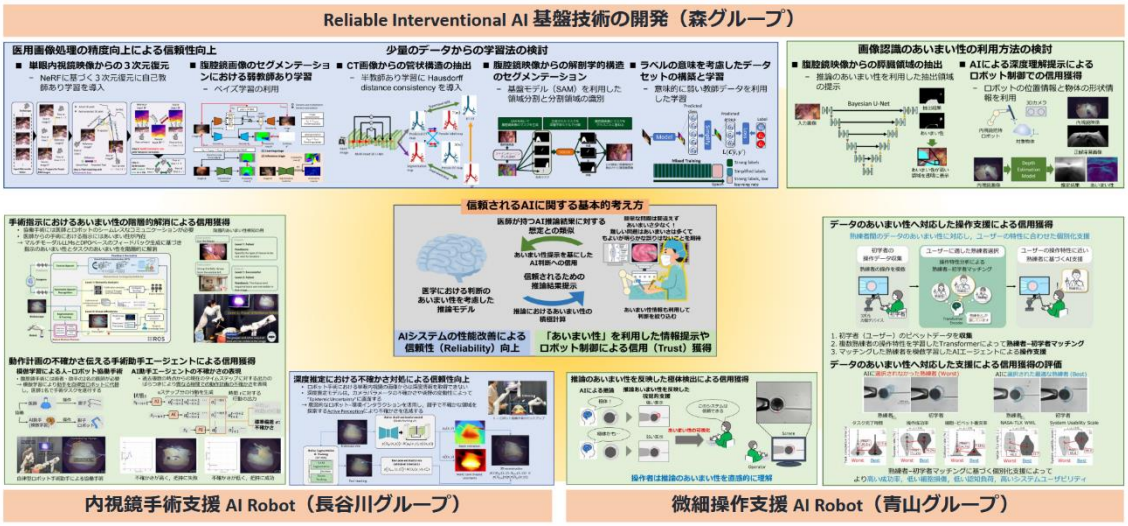


図 15　森課題の成果展開

(6)【研究課題 6】

人と AI の協働ヒューマンコンピュテーション基盤

研究代表者：鹿島久嗣(京都大学・教授)

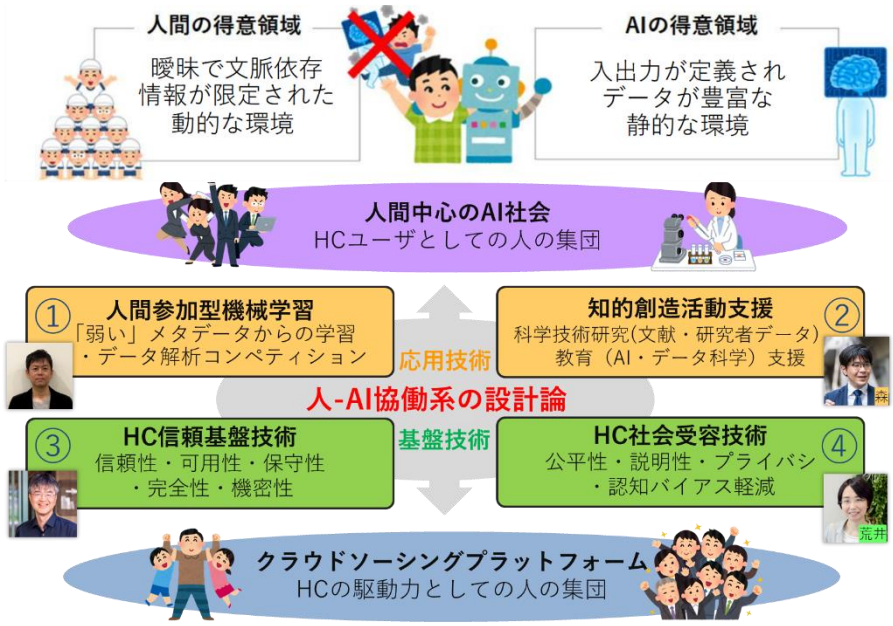


図 16　鹿島課題の研究テーマ

① 研究の概要

信頼される人間-AI 協働システムの基盤技術を確立することを目的に、人間参加型機械学習、ヒューマンコンピュテーション信頼基盤技術、AI システムの社会受容、そして知的創造活動支援という 4 つのテーマに取り組んでいる。また、当初の計画にはなかった大規模言語モデル(LLM)の急速な発展と普及に対応し、人間の協働相手の AI として LLM を想定することで、新たに LLM と人間の協働による問題解決を支援するための基盤技術の研究に取り組んでいる(図 16)。

② 独創的で国際的に高い水準の研究成果

ヒューマンコンピュテーションにおける信頼性にまつわる研究をまとめた包括的なサーベイを論文としてまとめたことは、人間と AI との協働フレームワークの研究分野の基盤となる成果である。また、人間参加型機械学習においては、専門家による「強い」メタデータだけでなく、クラウドソーシングによる「弱い」メタデータを統合した効率的な学習手法(例：図 17)、AI システムが社会で適切に受け入れられるための公平性とプライバシー保護を強化するための手法などで優れた基礎研究として優れた成果を上げている。これらの成果は、人間参加型機械学習をより難しい学習課題にまで広げ、信頼される高品質なデータを構築するための基盤となる技術として価値が高い。

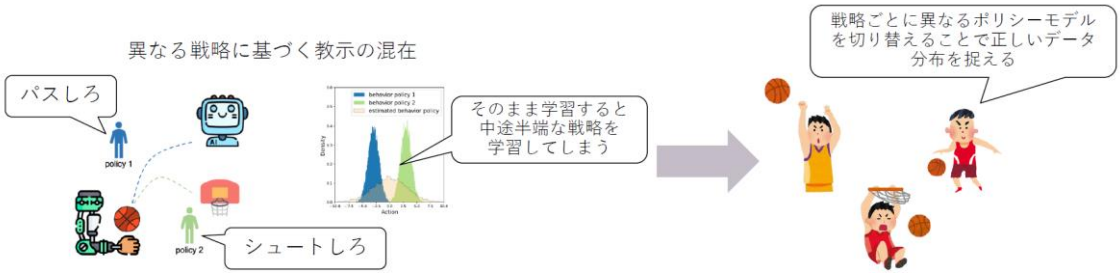


図 17　理論的な成果例：人間参加型機械学習における効率的な学習手法の提案(AAAI-2023)

③ 新技術シーズへの展開

データ解析コンペティションシステム「ビッグデータ大学」を開発・公開し(図 18)、地球惑星科学分野における課題解決を通じた実証へ向け取り組んでいる。また、自己教育・学習の支援のための深層学習モデリング技術の一部は、教育系企業において、e-learning システムにおける学習者の学修状況の把握や、課題の分析等に用いられている。

図 18　成果の一例「ビッグデータ大学」

(7)【研究課題 7】

D3-AI: 多様性と環境変化に寄り添う分散機械学習基盤の創出

研究代表者：高前田伸也(東京大学・准教授)

① 研究の概要

システムの利用者やデータの多様性を尊重し、時間的・空間的な性質の変動(環境変動)に適応できる、連合学習に基づく分散型・省エネルギーな AI システムの創出と、その社会システムでの利活用とする。本研究提案では、AI システムの「信頼」を、(1)データや機械学習モデルのプライバシーが利用者や設計者の意図通りに保護されること、(2)AI システムの健全な利用者が公平に AI の恩恵を受けられること、(3)利用環境の変動にシステムが自動的に適応し安心してシステムを利用できること、(4)省エネルギーで動作し環境負荷が小さいこと、と定義する。これらの「信頼」の要件を備える AI システムを本研究提案では D3-AI(Distributed AI for Dynamic and Diverse Environments)と呼び、その実現を目指して研究開発に取り組む(図 19)。

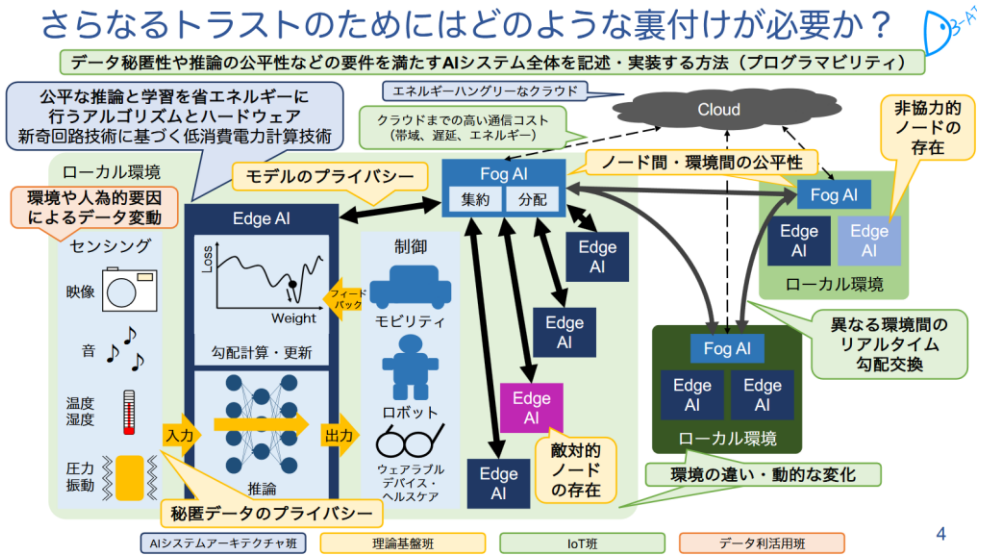


図 19　高前田課題の研究テーマ構成

② 独創的で国際的に高い水準の研究成果

基盤要素技術として、メモリ内計算(CIM)に基づく省エネルギーな CNN/Transformer 両対応 DNN 計算チップ、低計算コストなベイズ深層学習アルゴリズムとハードウェア、量子化ニューラルネットワークの学習理論、システムとして、省通信な連合学習アルゴリズム、連合学習における遅延バックドア攻撃、分散機械学習のための通信ミドルウェアと IoT 基盤技術、公平な連合学習のための学習理論とクライアントのクラスタリング手法、利活用について、地球観測データと高精度衛星測位データを用いた地すべりモニタリングシステムの実証など、各レベルの要素技術それぞれについて若手研究者がグループを率いて活躍し、分野

を代表する国際会議やジャーナルへの採択など、顕著な学術的成果を上げていることは注目に値する。

③ 新技術シーズへの展開

産業機械への連合学習の適用やロボット技術に基づくラボオートメーションに関するものを含め、複数の国内企業と機械学習、連合学習に関する共同研究を進めている。本プロジェクトは積極的なオープンソース戦略を採っており、研究開発の成果は早期に OSS として GitHub に公開している。これらのリポジトリでは、注目度を示す Star 数がすでに 100 を超えているものがあり、国内外の企業から製品展開に向けて技術相談や共同研究を打診・実施しているものもある。また CEATEC 2024 などにも出展している。

また、衛星による地球観測データと高精度衛星測位データを用いた地すべりモニタリングシステムについては、実証フェーズから実用化フェーズに向けて進んでおり、2022 年に公布され 2025 年からの本格的運用を控えた「宅地造成及び特定盛土等規制法」（通称「盛土規制法」）においては、当システムが有効であることが地元土木行政や土木関連企業からも確認され、実用化についての動きが加速している（図 20）。さらに、公共の社会基盤を支える土木フィールドを主なターゲットとして、静岡県や浜松市などの地域の土木行政および静岡県を中心とした土木産業界の事業者を呼び込んでコミュニティ形成を行っている。

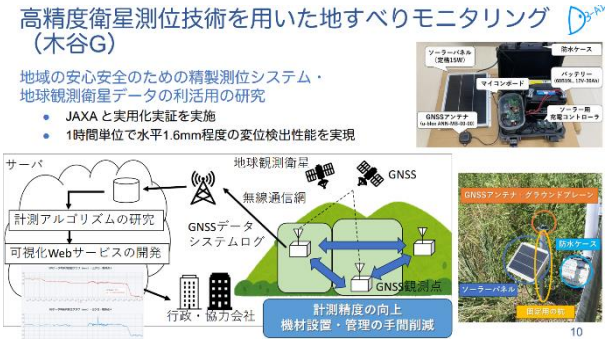


図 20　木谷グループの取り組む課題

【研究課題 8】

AI 駆動仮説の静的・動的信頼性保証と医療への展開

研究代表者：竹内一郎(名古屋大学・教授)

① 研究の概要

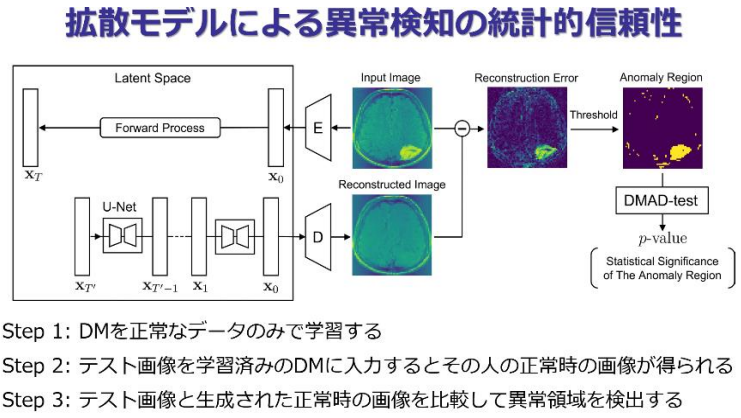
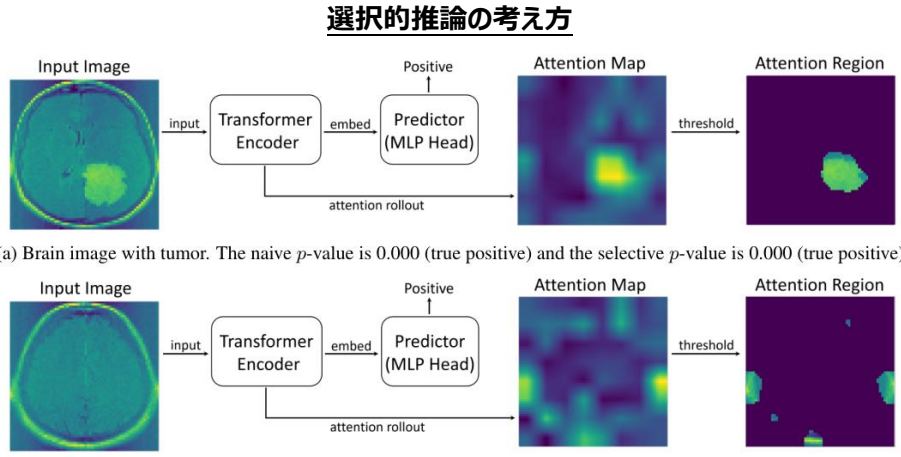


図 21　AI の信頼性指標

AI の信頼性指標として AI のアウトプットの統計的有意性(Statistical Significance)に着目し、選択的推論(Selective Inference)を中心に据え、深層学習モデルなど現在の AI で使用される複雑な機械学習モデルのアウトプットに対して妥当な p 値や信頼区間を付与するための方法を開発する(図 21)。統計的有意性は、研究開発で得られたアウトプットが真に意味のあるものなのか、それともノイズによる見かけ上のものなのかを判定するための指標であり、科学研究や技術開発の再現性を保証する上で欠かせない役割を果たす。また、薬剤治験や臨床試験などの生物統計(Bio-Statistics)分野で培われた品質保証の理論や方法を AI の分野へ発展・拡張する。最後に、病理診断(Pathological Diagnosis)分野において、日本最大規模の病理診断実績を持つ病理医を共同研究者として迎え、AI を用いた大規模病理診断システムの構築を進めるとともに、その信頼性を多面的に評価する。

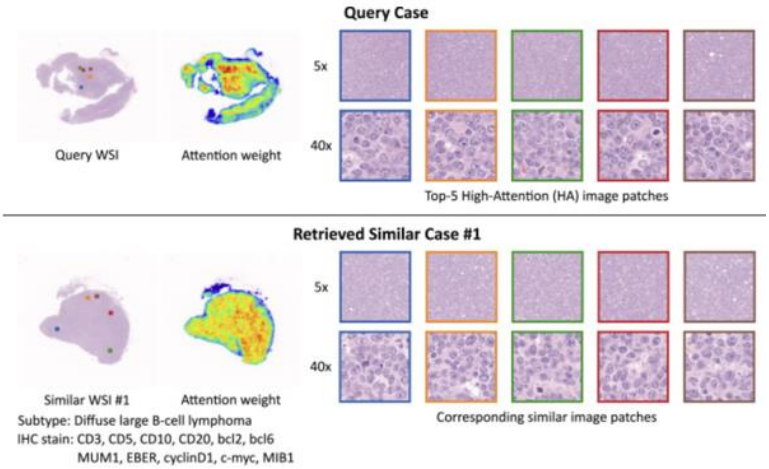
② 独創的で国際的に高い水準の研究成果

本研究における顕著な基礎研究の成果として、深層学習モデルに対する統計的検定法を構築したことがあげられる。深層学習モデルに対して有限サンプルで理論的な妥当性を持つ統計的検定法は他に存在しておらず、本成果は世界に先駆けてこれを実現したものである(図 22)。自動微分を模した基盤を開発したことにより、Transformer や Diffusion Model など様々な最先端の深層学習モデルに対する統計的検定が可能になった。これによって、本研究が最終的な目標として据える、科学技術分野における AI(AI for Science & Technology)の活用推進に向けて、科学の再現性の観点から大きな前進があったとみとめられる。また、説明指標のための統計的信頼性保証、個別改良の信頼性評価、脳画像解析の信頼性評価、AI を利用した病理診断の実証実験(図 23)などについても優れた学術的な業績を創出していることは、数多くの論文発表で裏付けされている。



出所）T. Shiraishi et al. Statistical Test for Attention Maps in Vision Transformers. Proceedings of The 41st International Conference on Machine Learning (ICML 2024)

図 22　深層学習モデルに対する統計的検定法



出所）N. Hashimoto et al. Case-based Similar Image Retrieval for Weakly Annotated Large Histopathological Images of Malignant Lymphoma Using Deep Metric Learning. Medical Image Analysis (2023)

図 23　AI を利用した病理診断

③ 新技術シーズへの展開

本研究で開発した AI の信頼性評価技術に関して民間企業 3 社と製造業 3 社と AI を活用した製品開発とその品質保証に関する共同研究を実施している。また、製造業向けの AI・機械学習に関するセミナーを 4 社 8 件実施、さらに 2022 年 4 月に東京ビッグサイトにて開催された第 6 回 AI・人工知能 EXPO アカデミックフォーラムに出展している。なお、本研究で開発したソフトウェア技術は、関連技術の開発者用のツールなどの基盤技術も含めて GitHub にて公開している。

(9)【研究課題 9】

納得感のある人間-AI 協調意思決定を目指す信頼インタラクションデザインの基盤構築と社会浸透

研究代表者：山田誠二（国立情報学研究所・教授）

① 研究の概要

AI が人の認知バイアス、価値観を基に信頼関係の崩れを検出し、適応的に較正キューを出して信頼較正を促すことで、信頼関係を最適化し納得感を向上させる信頼インタラクションデザインの理論構築と社会浸透を図る。具体的な社会浸透は、人間-AI 協調が必須な医療検診・健診における実稼働で実現する（図 24）。

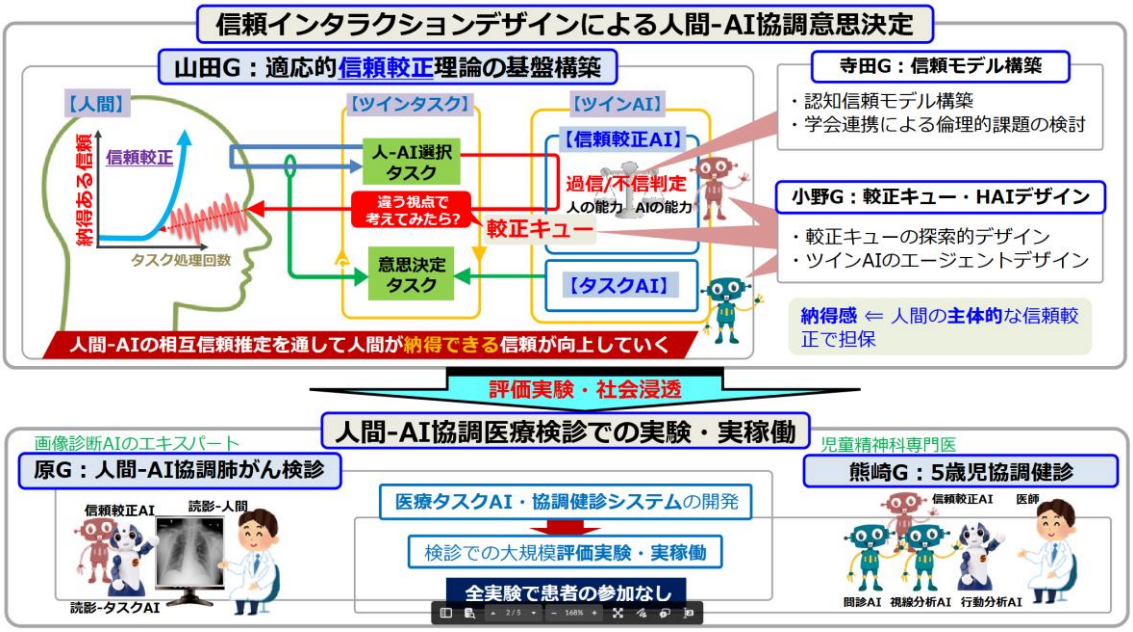


図 24　山田課題の研究テーマ構成

② 独創的で国際的に高い水準の研究成果

これまでの研究では、過不信方程式、リライアンス方程式、性能方程式の導入によって、信頼較正の枠組みを世界で初めて定式化し、信頼工学の基礎研究に大きく貢献した。また、較正キューという具体的な目的のために、ビデオ刺激や VR/MR 刺激によるナッジの工学応用を幅広く研究した。また、ナッジの基本概念である『リバタリアンパターナリズム』に注目し、効用の期待値に基づいた意思決定理論を基に利他性を効用としたナッジの定式化を行い、実験的に検証した。構造方程式モデリング SEM を信頼・過不信モデリングに応用し、時系列データ対応に拡張したダイナミック SEM を信頼ダイナミクス予測に応用した。さらに、医師と医療 AI による人間-AI 協調意思決定の実装と評価について、信頼較正付きの人間-AI 協調意思決定の枠組みを実際に胸部 X 線画像の人間-AI 協調読影に導入し、プロトタ

イプ実装がほぼ完成し、実稼働の簡単なデモが行える状態にまで達成した。信頼較正の促進が可能な人間-AI 協調読影システムは前例がなく、医療 AI の社会実装の観点からも、世界的にオリジナリティの高い試みであり、科学技術イノベーションに貢献する。

③ 新技術シーズへの展開

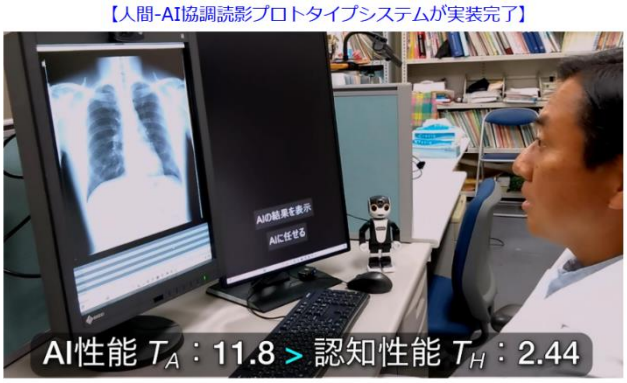


図 25　適応的信頼校較正システムのデモ

適応的信頼較正システム全体の実装は、胸部 X 線での人間-AI 協調意思決定システム全体のプロトタイプ実装が完成しており(図 25)、これから実用化に向けて大きく進展することが期待できる。一方、人間-AI 協調 ASD 児童のスクリーニングは、完成した 3 つの医療 AI の診断結果を統合するシンプルなアルゴリズムを開発し、実用化に結びつける予定である。また、適応的信頼較正理論については、アルゴリズムレベルでの特許申請の検討を始めている。

第 3 期採択の 3 課題については中間評価未実施であるため、進捗状況を簡単に記載する。

(10)【研究課題 10】

教育大航海時代の羅針盤：学習分析の信頼基盤 ReLAX の創出

研究代表者：島田敬士(九州大学・教授)

教育学習の文脈を深く理解し、解釈性の高い情報や根拠情報に基づいて適応的な教育学習を支援する AI の実現を目指している。具体的には、即時性、説得性、適応性の 3 つの観点で学習分析の信頼性を向上させる基盤技術の開発に取り組んでいる。

即時性の高い学習分析技術の実現に向けては、自然言語処理に用いられる fastText モデルを応用して、短い学習活動の特徴表現を獲得できる E2Vec モデルの開発や、授業期間の早い時期にドロップアウト予測を高精度に行える予測モデルを開発した。また、説得性の高い学習行動のプロセス解明に向けて、Attention からどのような行動が重要かを解析可能な Transformer による成績予測モデルを開発し、成績につながる一定時間における行動パターンを抽出できた。また、各講義回で実施した自由記述アンケートをもとにリスク予測モデルを開発し、リスクありの学生を推定する精度を向上させるとともに、リスク予測の説明性獲得を可能とした。さらに、分析モデルの適応性向上のために、ランキング学習のアイデアを取り込むことで学習データを増強する技術を開発した。高度な信頼性が求められる学校教育の場において、学習データの利活用による教育の質向上は切実な社会課題であり、AI を活用した分析基盤の実現に向けて、解釈可能性などにも配慮しつつ理論的な手法の開発に取り組んで成果を上げている。

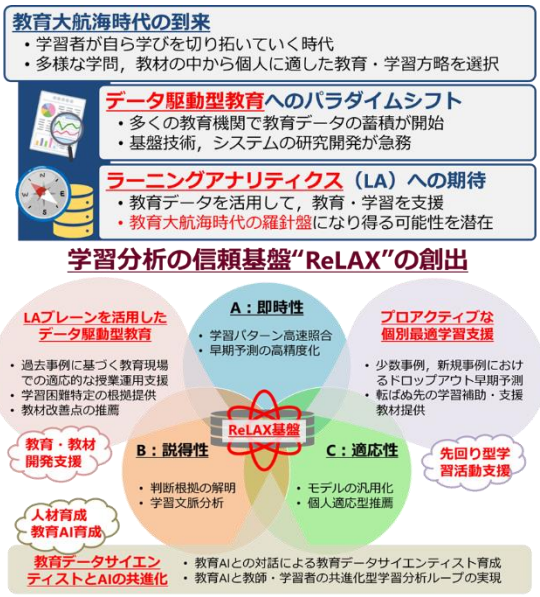


図 26　島田課題の研究テーマ

(11)【研究課題 11】

信頼される AI システムを実現するための因果探索基盤技術の確立と応用

研究代表者：清水昌平（滋賀大学・教授）

データを用いて因果グラフを推測する因果探索の方法論を研究する。信頼される AI に必要とされる説明性や公平性を達成したり向上させたりするために、因果グラフを用いた統計的因果推論が重要な役割を果たす。統計的因果推論を実行するためには因果グラフを分析者が描く必要があるが、それだけの領域知識があるとは限らない。そこで上述の方法論的課題とともに、政策、環境学、予防医学、臨床医学という高度な信頼性が求められる領域の研究課題にも取り組む。そして、各領域で必要とされるレベルで因果探索に関する方法論的課題の解決を目指す。

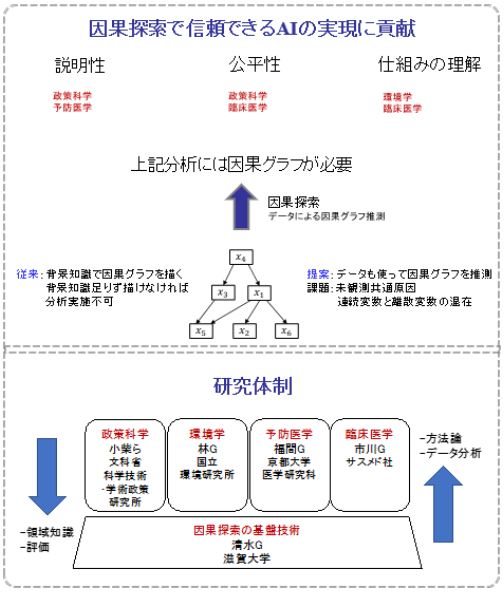


図 27　清水課題の研究テーマ

これまでに、離散変数と連続変数が混在する場

合の識別性や既存手法の実装の高速化、因果探索を用いた因果グラフの推定と反事実の確率計算などの基礎的研究で成果を上げるとともに、中枢神経系疾患に関するデータを用いた因果分析、日本の業種別健康保険組合が保有するヘルスデータ(年 40 万人規模)を対象として、保健指導介入から翌年度以降の健診結果への経時的な因果メカニズムの検討、メダカを用いた生態毒性試験における毒性発現の因果経路の探索などに取り組んでいる。理論的な成果についてユーザがすぐに使えるようなチュートリアル付きの code package を公開しており、応用研究の成果をもとに今後事例を作成する計画であり、社会実装への道筋も見えている。

(12)【研究課題 12】

記号推論に接続する機械学習

研究代表者：杉山麿人(国立情報学研究所・准教授)

本研究提案では、大量パラメータを利用する現代的な機械学習と、推論根拠の解釈性に優れた記号推論を融合することで、現在の機械学習がもつ信頼性についての課題と、記号推論が持つロバスト性の課題を同時に解決する基盤技術の開発を目的として、機械学習、モデリング、記号推論、アルゴリズムの 4 つのグループが協調して研究に取り組んでいる(図 28)。

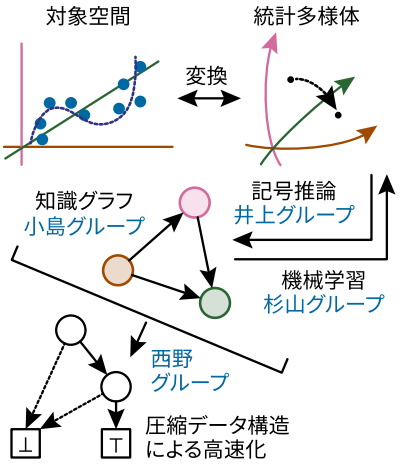


図 28　杉山課題の研究テーマ

これまでの成果として機械学習グループでは、テンソルを確率モデルとして扱い、情報幾何学の理論を援用して解釈可能かつ安定なテンソル分解を実現する技術であるテンソル多体近似を確立した。これは、本プロジェクトの核になる技術である。同時に、モデリンググループにおいて、記号推論をテンソル演算で実現する確率論理プログラミング言語 T-PRISM 上でこのテンソル多体近似を実装し、記号推論への接続を進めることに成功した。多体近似のための論理解釈について考察を進めるとともに、GPU 実装などの高速化についても取り組んでいる。この枠組みを発展させるために、記号推論グループでは、ベクトル空間における論理推論に関する研究を継続しており、連続ドメインにおける機械学習データと記号で表現される推論用の知識を接続する手法として、命題論理式の行列表現(ReLU NN)による DNF 式の学習や一階述語論理プログラムの微分可能学習方法を開発した。また記号推論と機械学習を接続した実応用として、自動運転等で用いられる事前訓練済み物体認識への論理制約の導入について検討し、試験開発したシステムにより ROAD-R Challenge for NeurIPS 2023 において優勝と 3 位入賞を果たした。さらに、本質的なアルゴリズム改善に向けた研究として、アルゴリズムグループでは、部分グラフの数を正確に数え上げる効率的なアルゴリズムを考案するとともに、考案したアルゴリズムをネットワークの信頼性解析へ応用した。以上のように学術分野で顕著な成果を上げており、記号推論と機械学習が接続した新たな計算原理の創出に向けて順調に研究が進展している。

7. 総合所見

(1) 研究領域のマネジメント

本研究領域で採択した 12 の研究課題は、「信頼される AI」の戦略目標に沿ったもので、達成目標の 3 本の柱である「現在の AI 技術を克服する新技術」、「AI システムの信頼性・安全性を確保する技術」、「データの信頼性及び意思決定・合意形成支援技術」を横断するポートフォリオを実現している。また 12 名の領域アドバイザーの専門分野もバランスがとれ、12 個の課題それぞれに専門的な立場からコメントや評価が頂ける体制である。それに加えて各領域アドバイザーの先生方が、信頼される AI の重要性を強く意識して、各研究課題の活動が領域全体の目標に沿っているかを常に確認し、不足している点について的確なフィードバックを頂いている点は、領域の運営の上で非常に有効に働いている。あわせて、領域が発足した当時は予想されていなかった生成 AI の急速な発展については、影響を受ける研究課題については積極的に研究計画にフィードバックをかけ、成果の強化もしくは研究計画の追加等の対応を行うことで、最新動向を踏まえた機動的な対応を行うことができた。

対外的な連携について、AIP ネットワークラボ等やセミナー等の活動を通して、CREST・さきがけ・ACT-X などの研究者間の交流を強化できたことは、特に今後を担う若手研究者育成の観点から効果があったと考えている。たとえば、AIP チャレンジの仕組みは、CREST に所属する若手研究者が自主的に研究に取り組み成果をアピールする貴重な機会となった。また国際連携やセミナー企画などでは、複数の研究領域が連携することで、単独では難しいイベントも実現することが可能になった。本研究領域の PI や Co-PI が、他の研究領域の研究総括や領域アドバイザーなどを兼任する事例も増えており、今後の AI 領域の中核を担う人材が育っている。また、課題中間評価などでは、若手研究者の育成について評価されるチームも多く、待遇面等で特任研究員等の雇用が難しいとの声もきかれる中で、チーム内の若手研究者や学生を大切に育成して行ける環境が提供できたと考えている。

(2) 研究領域としての戦略目標の達成に向けた状況

信頼される AI やトラストをめぐっては、国内外の様々な場で議論が活発化し、概念の体系化が進んでいる。本研究領域では、それを俯瞰するためのセミナーの企画や JST の CRDS 福島フェローの企画による一連のセミナーへの参加等を通して理解を共有した。また、2 回の領域会議でセッションを企画し、ブレーンストーミング的な議論を通して各課題の戦略目標のもとでの位置づけを議論した。

参考として、CRDS 福島フェローによる分野俯瞰を以下に示す。まず、トラストに関連する様々な研究活動の中で、本研究領域は図 29 のスライド中の「信頼される AI」に位置づけられ、デジタルトラスト、フェイク対策、AI ガバナンス、トラストの観察・理解などの研究と連携しながら、AI のふるまい予想・対応可能性の問題に重点をおく技術開発を担うこととされている。

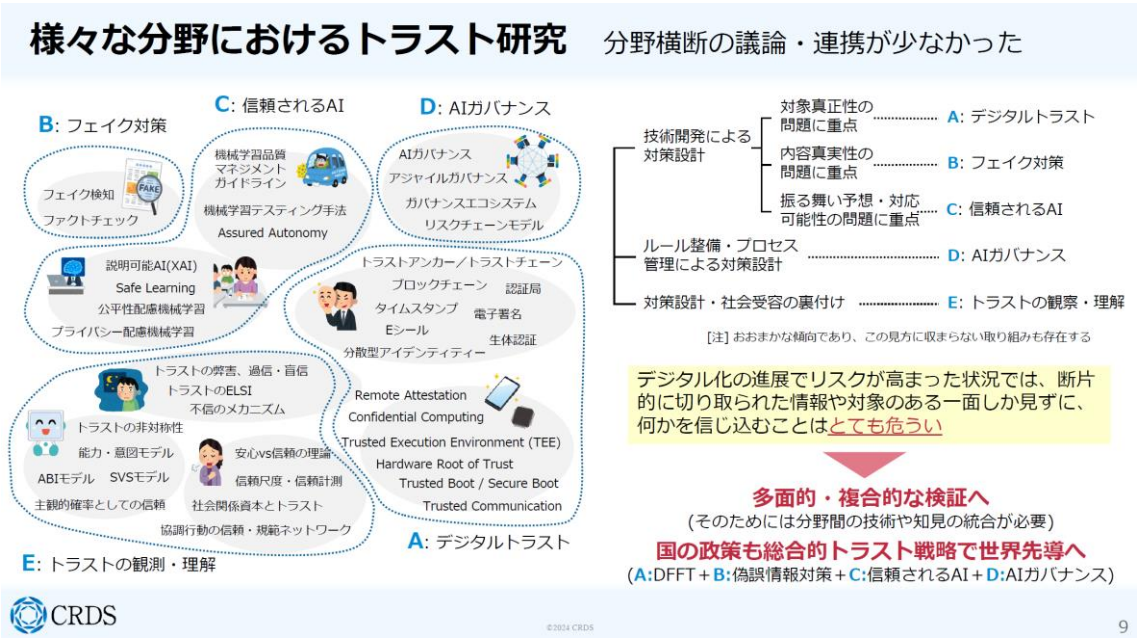


図 29　様々な分野におけるトラスト研究　(出所)CRDS「トラスト研究俯瞰セミナー」[1]

また、図 30 に示すトラストモデルは、人間中心 AI の実現に向けた概念整理として活用した。トラストや信頼の定義にはさまざまな解釈があることを踏まえつつも、第 1 回目の領域会議におけるセッションでは、スライド上に示された「Trustor＝信頼する側(人間)」と「Trustee＝信頼される側(AI)」を区別せず、AI の性能要件だけに言及するチームもあったことから議論がかみあわず、混乱を避けるために概念整理を行ったものである。特に、Trustworthiness を客観的にデータに裏付けられた AI の品質特性として捉え、裏付けのない場合でも信頼を寄せる人間の主観的判断とのずれがあるという定式化は、理解しやすいものとして受け入れたチームが多かった。

[1] 福島俊一：招待論文「SoK：デジタル社会におけるトラスト形成の課題と展望」，情報処理学会論文誌，Vol.65，No.12，pp.1620-1629（Dec. 2024）

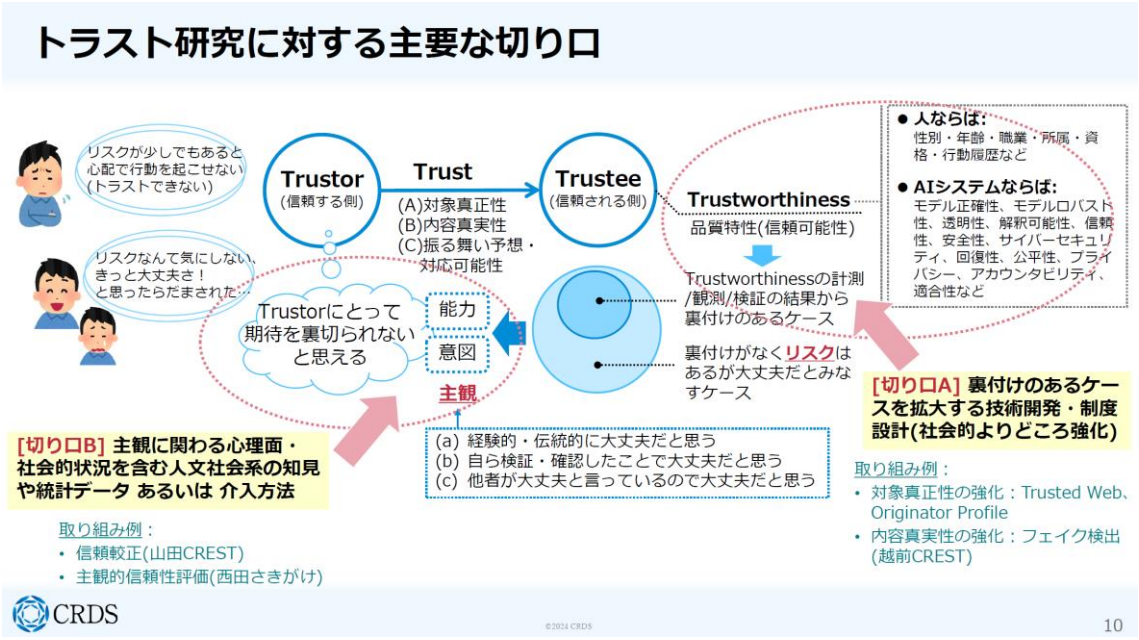


図 30　トラスト研究に対する主要な切り口　(出所)CRDS「トラスト研究俯瞰セミナー」[1]

以下では、12 のチームを①社会課題解決、②人間 AI 協調、③原理解明、④基盤構築の 4 つの観点にしたがって 3 つずつのグループに分けて、領域全体としてのゴール設定や状況をまとめた(表 8)。各チームが取り組んでいる「信頼される AI」の定義や位置づけ、実現について領域横断討議でまとめた。

表 8　「信頼される AI」に関する 4 つの観点と対応するチーム

観点	説明	対応するチーム
①　社会課題解決	信頼される AI の実現を通して社会課題の解決に取り組む	越前チーム：フェイクメディア 伊藤チーム：ハイパーデモクラシー 鹿島チーム：ヒューマンコンピュテーション
②　人間 AI 協調	AI と人間と協調して問題解決にあたることで、人間の能力の限界を突破することを狙う	後藤チーム：Explorable 推薦 山田チーム：信頼インタラクション 森チーム：あいまい性許容
③　原理解明	AI の動作原理の解明や、新たな AI モデルの創出を目指す	乾チーム：言葉で説明できる AI 杉山チーム：記号推論に接続する機械学習 清水チーム：因果推論
④　基盤構築	幅広い問題解決に資する AI 基盤の構築を目指す	竹内チーム：AI 駆動仮説信頼性保証 高前田チーム：分散機械学習基盤 島田チーム：学習分析信頼基盤

① 社会課題解決

「社会課題解決」では、社会科学的な観点も取り込みながら、信頼性の観点から人間と AI がどのようにかかわるかを体系化した上で、AI システムの社会実装に取り組む。伊藤チームでは、ソフトウェアエージェントと人間が一緒に参加するソーシャルネットワーク上での民主主義(ハイパーデモクラシー)のための合意形成プラットフォームの実現を目指している。特に、ソーシャルネットワークの中に常駐して、人間の代理として意思決定やインタラクションを仲介する AI エージェントの研究に取り組み、AI によるファシリテーション機能を備えたハイパーデモクラシープラットフォームの実装と合意形成実験で成果を上げている。越前チームでは、AI を悪用したサイバー攻撃の脅威から社会を守る手段として、AI により生成されたフェイク映像、フェイク音声、フェイク文書などの多様なモダリティによるフェイクメディアを用いた高度な攻撃を検出・防御する技術の研究開発に取り組んで、国際的にも大きな注目を集めて、社会実装や大型プロジェクトへの展開も順調に進んでいる。鹿島チームでは人間と AI が協調して問題解決を図る人-AI 協調システムの設計を体系化し、データ解析コンペ基盤での実証に着手している。このために、課題発見、データの収集や注釈、モデリング、評価を含む AI プロセス全体を多数の人間の参加を得て効率よく実行するための、データ分析・予測、アノテーションバイアス制御技術や人間参加強化学習の手法を提案している。

② 人間 AI 協調

「人間 AI 協調」では、AI システムが常に正しい出力をするわけではないという想定のもとで、人間と AI システムの協調作業をどのように支援すればよいのかという問題に焦点をあてて新たな AI システム設計手法の創出を目指す。後藤チームの音楽推薦基盤では、ユーザが推薦システムの挙動を探索できる基盤システムを実現し、透明性の高い推薦サービスを提供することで、パーソナライズ化されたサービスをユーザが安心して享受できる社会の実現を目指して、音楽発掘サービスを実証実験の場として、すでにサービス運用を開始している。森チームでは、医療画像の上であいまい性をリアルタイムに可視化して提示することで、医師の判断を支援するフレームワークの実現を目指して、提案手法をオープンライブラリの基盤として社会実装している。山田チームでは、人間が AI を過信したり過小評価したりする状況をモデル化して、人間の AI に対する信頼を適応的に較正する理論を構築し、信頼インタラクションデザインによる人間-AI 協調意思決定システムの実現と実証に取り組んでいる。

③ 原理解明

「原理解明」は、記号的な推論を取り込んだ新しい計算手法を確立し、AI システムの信頼性を確保する挑戦である。乾チームでは、知識グラフとテキストの混合グラフ埋め込み表現や、深層学習と相互作用する高階論理推論の研究で取り組んでいる。杉山チームでは、機

械学習と記号推論の接続を目指して情報幾何学に基づく理論研究に取り組むとともに、記号推論を行列・テンソル表現して論理推論をベクトル空間上で実現する方式や、記号推論をテンソル演算で実現する確率論理プログラミングなどの実践的な手法の開発にも取り組んでいる。清水チームでは、AI の信頼性評価・向上に有効な「因果グラフ」に焦点を当て、因果グラフをデータから推測するための因果探索の手法と因果グラフを用いた説明性・公平性の解析に関する研究を実施している。いずれのチームもインパクトの高い学術的成果を上げている。

④ 基盤構築

「基盤構築」は、社会システム基盤としての AI システムの信頼性にかかわる技術を研究開発し、基盤システムに必要となる高度な機能の実現を目指す。信頼性の要件は応用分野によって異なる。たとえば竹内チームでは、AI システムがデータから仮説を生成し評価する一連のプロセスの中で、発見した知識(AI 駆動仮説)の信頼性を保証するための数理基盤の確立と実証を通して、AI によって発見される知識の信頼性の保証を目指して、独創性の高い手法を世界に先駆けて提案している。島田チームでは、データ駆動型教育支援や個別最適学習支援基盤の構築を通して、データ駆動型教育データサイエンティストと AI の共進化を目指している。このような学習分析基盤では、即時性、説得性、適応性などが重視される。高前田チームでは、IoT に基づく分散 AI の信頼性の性能要件から出発して、連合学習に基づく分散型かつ省エネルギーな AI システムの創出と利活用に取り組んでいる。このために AI システム・アーキテクチャ、機械学習理論、IoT のための連合学習コンピューティング基盤などの異なる技術階層間の連携と、モビリティや土木等での実証を進めている。信頼される AI の実現においては、人が AI システムを信頼することと、AI システムが信頼するに足りる性能要件を満足することは、大きく観点が異なる。一方で、特に社会システム基盤の実装では、後者と前者が密接にかかわり、ときとして後者が前者の支配的な要因となるため切り離して考えることができない点に注意が必要である。

⑶本研究領域を設定したことの意義と妥当性

本研究領域の発足後に、AI をめぐる状況は劇的に変化した。当初の戦略目標では AI が社会に普及する時代を予見して、社会に受容される AI を実現するための基盤技術を「信頼される AI」として定義し、新たな研究コミュニティを創出することを目的の 1 つに掲げていた。しかしながら僅か 2 年のうちに AI のリスクへの対応や品質管理は国際社会の喫緊の課題となり、各国でルールや仕組み作りが急速に進められ、政策として実装が進んでいる。

このように信頼される AI をめぐる社会の動きは、本研究領域の当初の想定を超えて、本研究領域自のスケールをはるかに凌ぐ大きな広がりを持つものとなった。一方で、その技術基盤については未だ模索状態にあるといえる。すなわち、生成 AI の安全性や脅威、社会受容性などの課題は、そもそも公平性とは何であるか、偏見とは何か、倫理的であるとはどの

ようなことであるかなど、人間社会が従来から抱えていた問題そのものであり、AI が登場したからといってにわかに解決されるものではない。また、AI の解釈性や説明性を高めるためには、理解や納得といった人間の認知プロセスやその限界を知ることが必要であるが、この問題は現在の AI ではアプローチできない領域である。

本研究領域では、単に現在の AI を利活用するための技術開発ではなく、これまで解決できなかった社会的な問題にチャレンジするための新たな AI の研究開発に取り組んでいる点で、その目標設定は妥当なものであった。社会課題解決型の研究に加えて基礎的な研究もバランスよく配置されており、分野ごとの方向性や進捗度合いにはばらつきがあるものの、以上を踏まえて、研究は順調に進んでおり、本研究領域の目標に沿った役割を果たしていると考えている。

(4)科学技術イノベーション創出に向けた、今後への期待、展望、課題

これまでの科学では、モデルによる予測の根拠を人間が説明できることがその信頼性の基盤とされてきた。しかし、AI の登場によって、解釈は可能であっても説明が困難な「複雑なもの」を、複雑さを保ったままモデル化する手法が広く受け入れられるようになった。これは AI の信頼性が、必ずしも人間が結果を完全に説明できることに依存しないことを意味している。この変化は、AI がもたらす科学のパラダイムシフトであり、人間の認知的限界を超えた計算と予測を可能にする AI は、科学や社会基盤システムの発展に新たな方向性を示している。現実世界の複雑性をどのように受け入れ、折り合いをつけるのかは、我々が「理解すること」と「信頼すること」の境界をどのように再定義し、AI を信頼してその出力を受け入れるのかという、人間自身のあり方に深く関わる問題であるといえる。

本研究領域の課題の中で、信頼性保証や因果探索などは科学の方法論そのものにかかわるもの、ヒューマンコンピュテーションや信頼インタラクションデザインなどは人間の認知能力の限界突破を目指すものであり、上記の科学のパラダイムシフトに直接かかわっている。一方、インフォデミック克服、ハイパーデモクラシー、Explorable 推薦基盤などは産官へのライセンス提供、起業、企業との協業による実サービス運用などにより、すでに社会実装されていて、フィードバックに基づきさらに技術が展開するサイクルを確立しつつある。他の課題も、新しい理論やモデルを考究する高い学術性や医療や教育や土木など具体的なドメインにおける社会実装によって、科学技術イノベーションへの貢献が見込まれる。特に後者については、試行的な実証段階でのユーザを含む評価が課題となり得るが、人間中心の AI を実現する上でこの問題は避けて通ることはできないことから、中間課題評価(第 3 期チーム)や最終評価(第 1 期・第 2 期チーム)に向けて、研究総括・領域アドバイザーで、しっかりとフィードバックをして行きたい。

(5)所感、その他

巨大な生成AIは、その構築コストに加え、推論コストも非常に大きい。そのため、ハードウェアの改良、モデルアーキテクチャの最適化、圧縮技術や推論高速化の技術開発が急速に進んでいる。一方で、期待される経済的効果と運用コストのバランスや、社会基盤としてAI をどのように実装し、計算基盤などのコアファシリティをどう配置するのかといった問題もある。

このようにAIの社会実装における技術的課題が次々と明らかになる中で、現状では我が国はAI分野で世界を牽引する立場にはない。今後の研究開発を見据えたとき、まずはイノベーションを生み出せる基礎研究力と国際的な競争力を備えた人材を育成することが必要である。その際、成果のみを性急に求めることなく、忍耐強く研究に取り組む姿勢が求められる。また近年のAI研究は、多様化が進み、扱うべき課題も多岐に渡る。単一の研究領域ではカバーしきれない問題が増えているため、新たな戦略目標の設定では、これまで以上に研究領域間でのバランスへの配慮や連携を強化するための方策が必要となる。

以上