



「人とAIの共生・協働社会を構成する 要素研究と基盤技術の創出」 研究総括説明

2025年4月

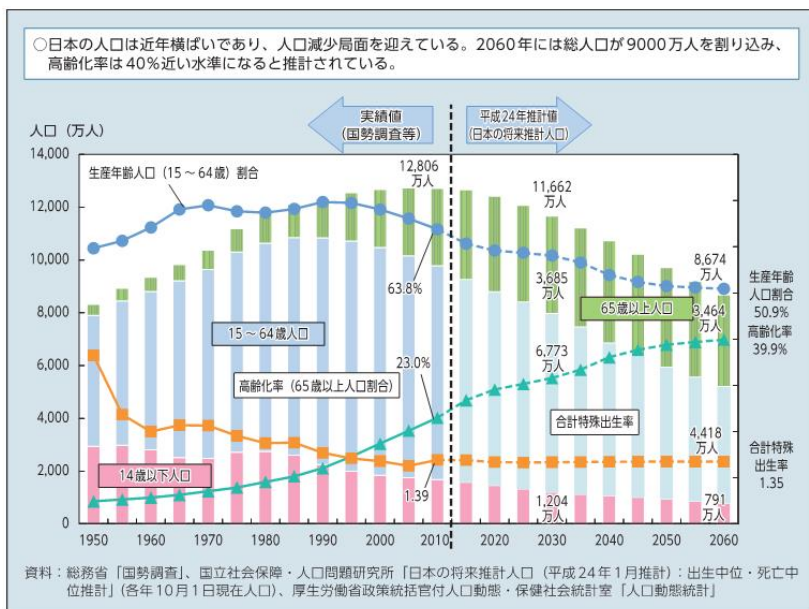
研究総括 山下 直美

（京都大学大学院情報学研究科 教授）



科学技術振興機構

社会的な課題：“少子高齢化社会”



- 2008年をピークに日本の人口は減少局面へ。
- 2060年には人口が9000万人を割る見込み。



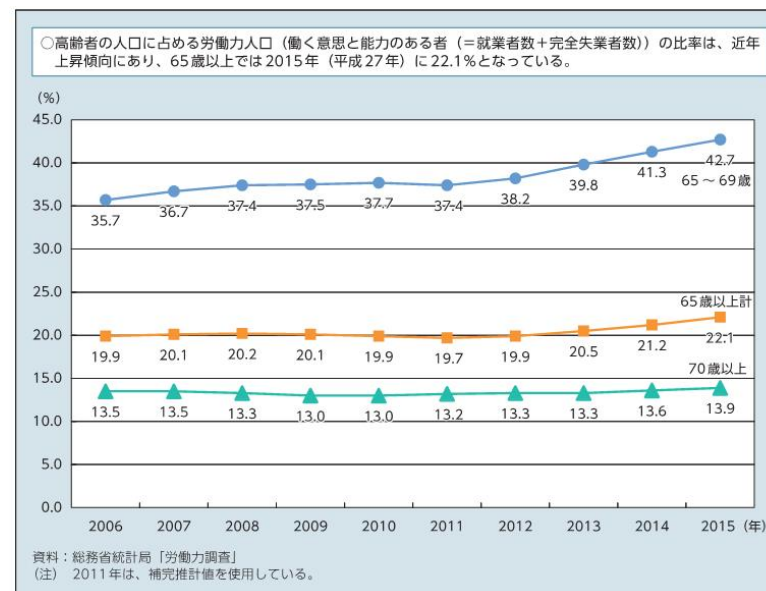
生産年齢人口の割合 ～50%
 高齢化率 ～40%へ

- 高齢者人口に占める働きたい高齢者の比率は上昇中
 ～40% (65-69歳)、～14% (70歳以上)



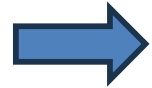
・・・とは言え、現役世代と同じように働くことはできない・・・

できるだけ機械・設備・システムで代替

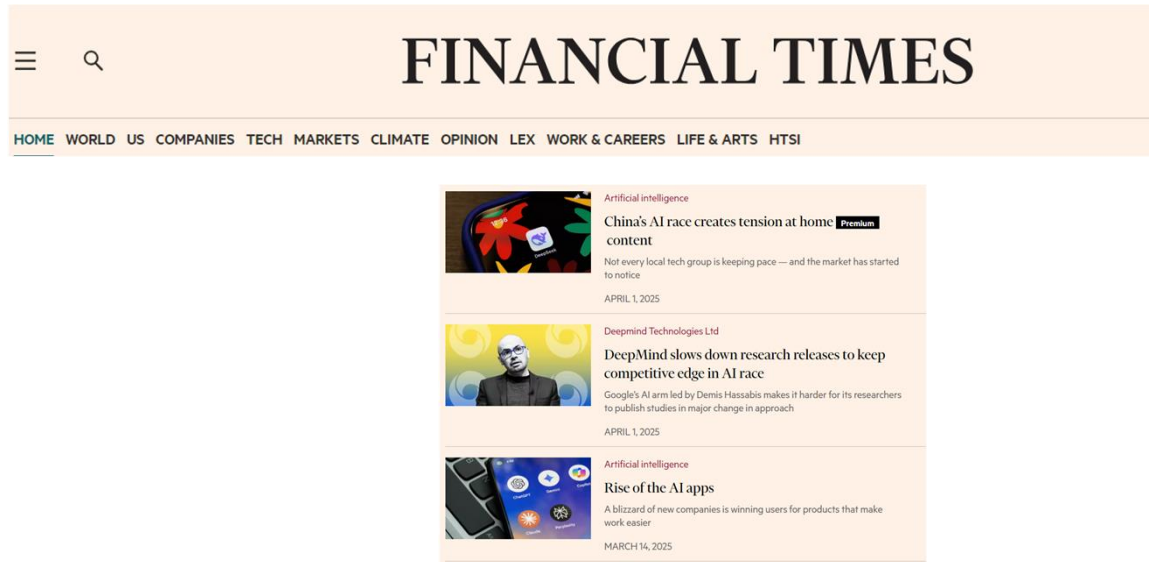


出典：厚生労働白書 https://www.mhlw.go.jp/wp/hakusyo/kousei/16/dl/1-04_01.pdf

AIは発展し、社会で既にかなり利用されている



これまでよりさらにAIが人間の仕事を代替する時代が到来



英国フィナンシャル・タイムズ (ft.com)

AIに関する記事数(1ヶ月以内)

1,000件超

米国ウォールストリート・ジャーナル (wsj.com)

AIに関する記事数(1ヶ月以内)

500件超

人間に代わって仕事をするAI、具体的には

- 人間に代わって、各種交渉を行うAI
 - 単独では無理な場合、他のAIと協調して行動
 - 自らデータを集め、分析し、行動するAI
- ⇒ AIエージェント に注目が集まっている。

THE WALL STREET JOURNAL.

English Edition | Print Edition | Video | Audio | Latest Headlines | More

Latest World Business U.S. Politics Economy Tech Markets & Finance Opinion Arts Lifestyle Real Estate Personal Finance Health

Everyone's Talking About AI Agents. Barely Anyone Knows What They Are.
"Agentic" is the biggest buzzword in Silicon Valley, but tech companies and enterprises lack a common understanding of its meaning, and it's causing problems.
By Isabelle Bousquette March 29, 2025 08:00 am ET



AI-Powered Cybersecurity Firm ReliaQuest Raises More Than \$500 Million
Armed with AI agents, ReliaQuest's platform detects and responds to cyberattacks within minutes, company says.
By Angus Loten March 31, 2025 04:25 pm ET



AI Agents Are a Moment of Truth for Tech
All emerging technology needs to deliver on its promise, sooner or later. For AI agents, that time is now.
By Steven Rosenbush March 29, 2025 08:00 am ET



研究領域の概要と達成目標

研究領域の概要

- 信頼性・公平性・安全性等を考慮しながら、AIと人の共生や多様なAIの連携を可能とする技術の発展及びそれらを基にした**複数の人と複数のAIによる協働の実現**
- 多様なAIと人間が共存し、**調和した共生・協働社会の実現**に向けたAIの開発やその要素技術の研究を推進
- 人文・社会科学の視点や、医療・教育・法律など**AIが活用される分野ごとの専門的な視点を取り入れた研究**にも取り組む

達成目標

- 安全かつ快適な「人とAIの共生・協働社会」の実現**

達成目標

信頼性・公平性・安全性などを考慮しながら、以下の実現を目指す。

- ① **人とAIの共生**：説明可能性や行動規範の実装等
- ② **多様なAIの連携**：相互運用プラットフォーム等
- ③ **複数の人と複数のAI協働**：社会デザイン等



募集・選考・領域運営にあたっての研究総括の方針

1. 背景

- 人工知能(AI)関連技術の急速な進化と社会への普及に伴い、**高機能なAI エージェントが人間と協力**し、複雑な問題の解決や業務の効率化、創造性の向上に貢献
- 多様なAI が乱立し、人間や AI エージェント同士が自律的に相互作用することで、**人間による制御が困難になるというリスク**も指摘される。
- 人や社会のウェルビーイングに資するAI** やデザインの在り方を検討することが不可欠
- AIの普及がもたらす社会課題の予兆**を捉え、未然に防ぐ手法の解決も必要

期待



- ・複雑な社会課題の解決、社会システムの全体最適化
- ・多角的な議論や合意形成
- ・新たなビジネスを生むエージェント経済圏 等

懸念



- ・AIのシステミックリスク、AIの暴走
- ・有害なAI、ブラックボックス問題
- ・AIが持つ影響力の強さと社会的受容性 等



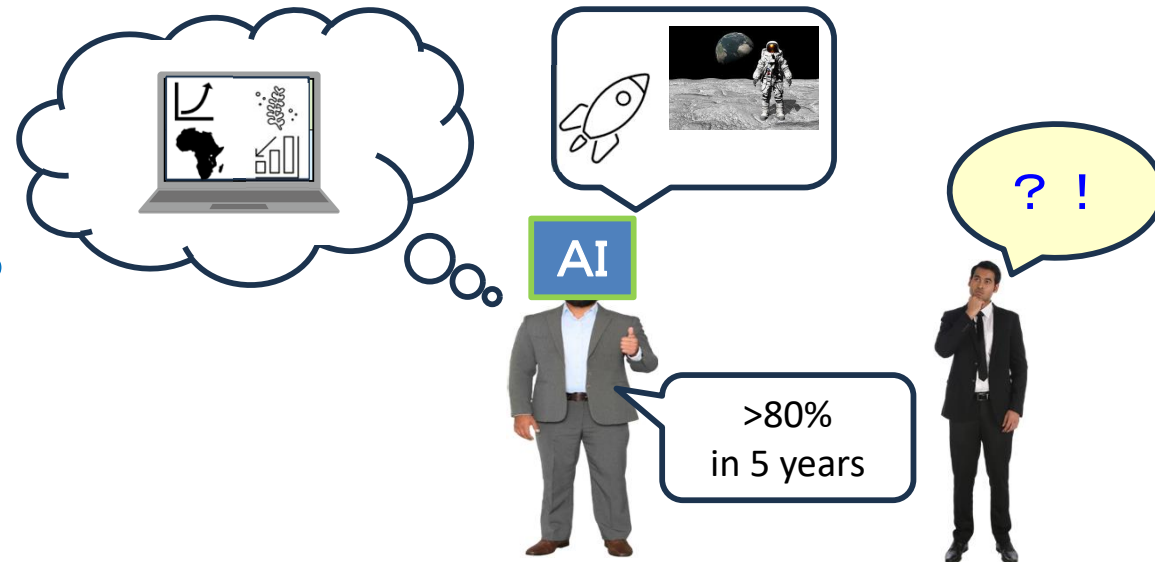
多様なエージェントが連携し、
リスクを抑制しつつ価値創出を
最大化するための研究開発を促進

募集・選考・領域運営にあたっての研究総括の方針

2. 募集・選考の方針

(1) 基本方針

- 複雑で複合的な社会課題を解決し得るAI、あるいはAIの普及がもたらす社会課題の予兆をとらえて未然に防ぐ手法などの要素技術を開発することを主眼
- 人文・社会系研究の視点や、AIが活用される様々な分野(医学、教育、法学、経済学など)の専門的な視点を取り入れた研究も積極的に歓迎



例1: 様々なデータから人類が再び月に到達する確率を予測



例2: 遠隔医療システム

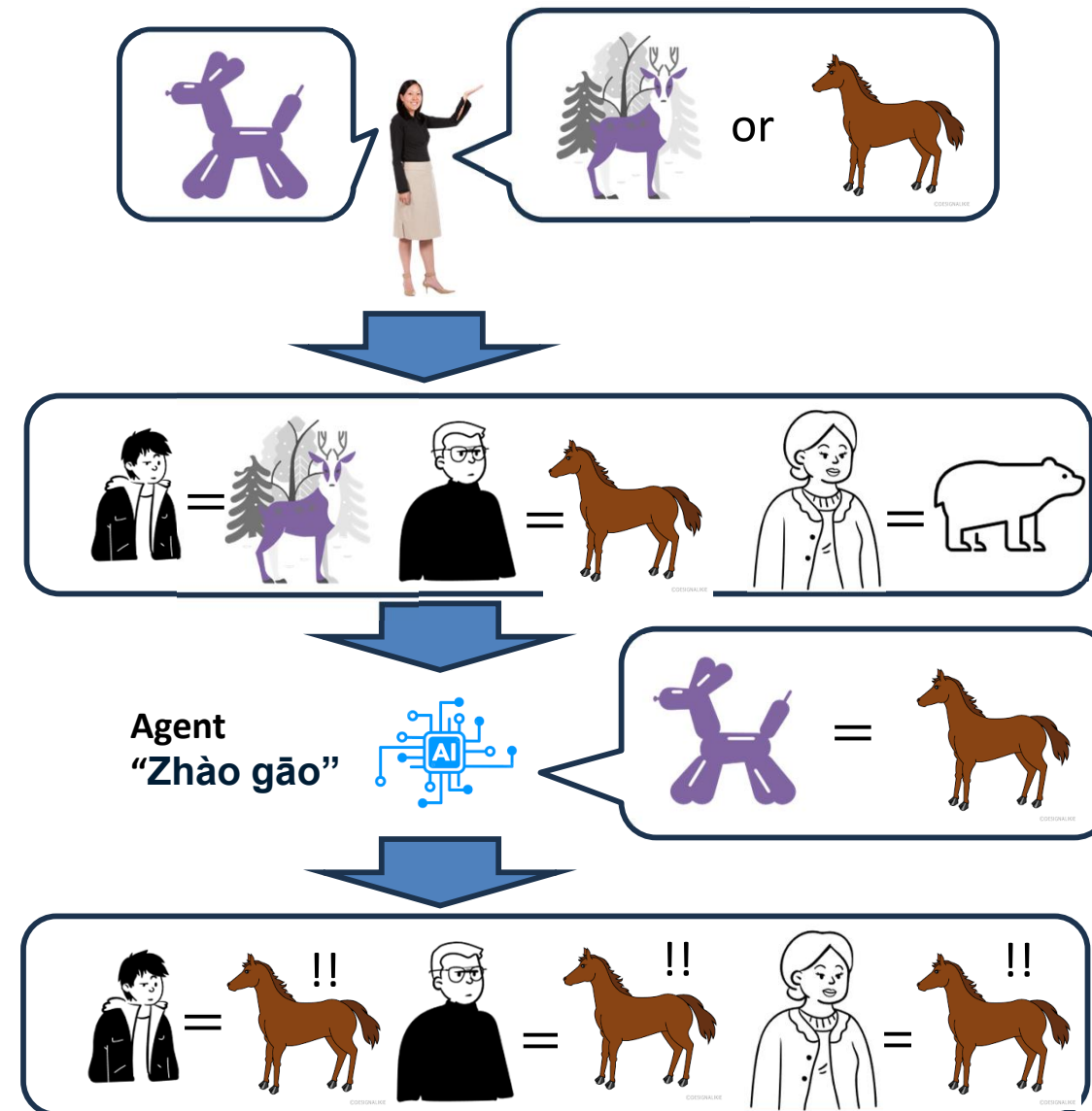
募集・選考・領域運営にあたっての研究総括の方針

2. 募集・選考の方針

(2) 想定する研究内容

(ア) 人とAIの共生

- 人とAIが安全かつ快適に共生し、ともに成長するために必要となる要素
技術開発の他、AIが人に与える影
響に関する研究、AI依存による人間
の思考低下リスクの回避方法に関
する研究
- AIが人間の相談事項に対して回答
する場合に、回答の信頼性を担保
するための技術や、AIが人間の不安
に寄り添うことが可能となるような
技術、およびその理論的基盤などに
関する研究



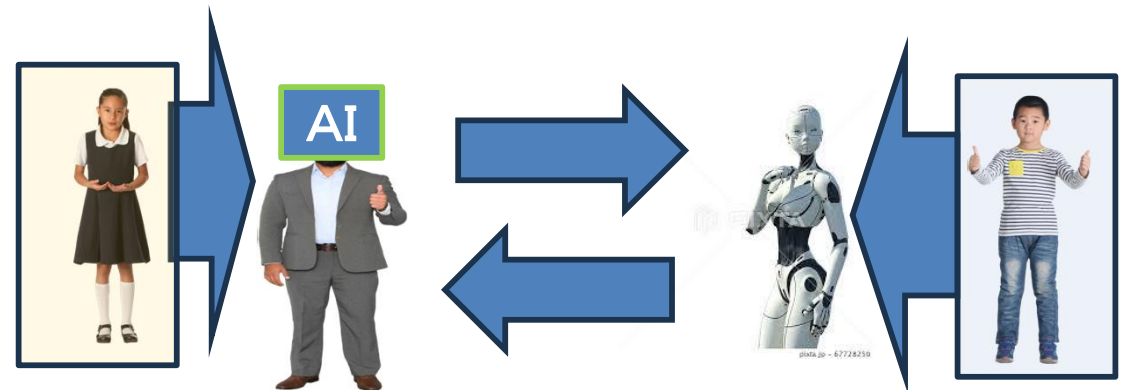
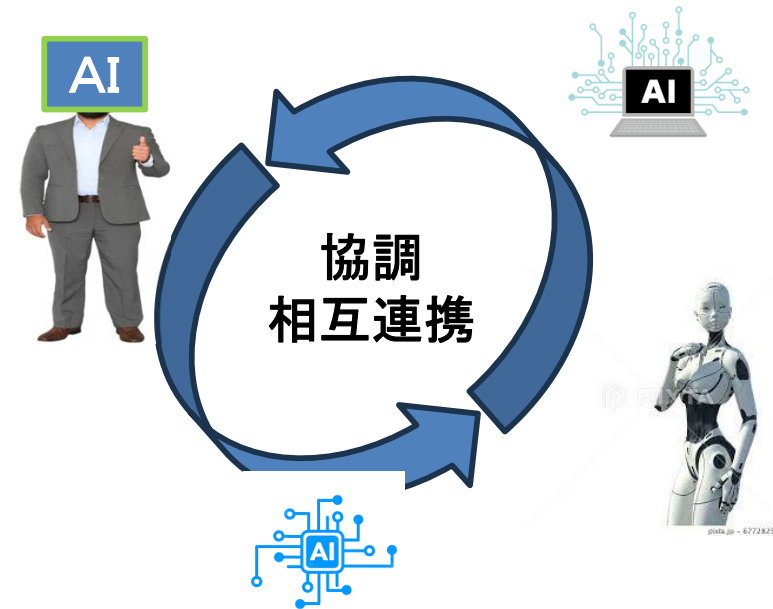
募集・選考・領域運営にあたっての研究総括の方針

2. 募集・選考の方針

(2) 想定する研究内容

(イ) 多様なAI間の連携

- 複数のエージェントが存在する場合に、**互いが協調する仕組みの開発**
- エージェント同士の交渉に際して、**依頼者のプライバシーやセンシティブなデータを保護**し、大規模なエージェント群として管理あるいは制御する手法の開発

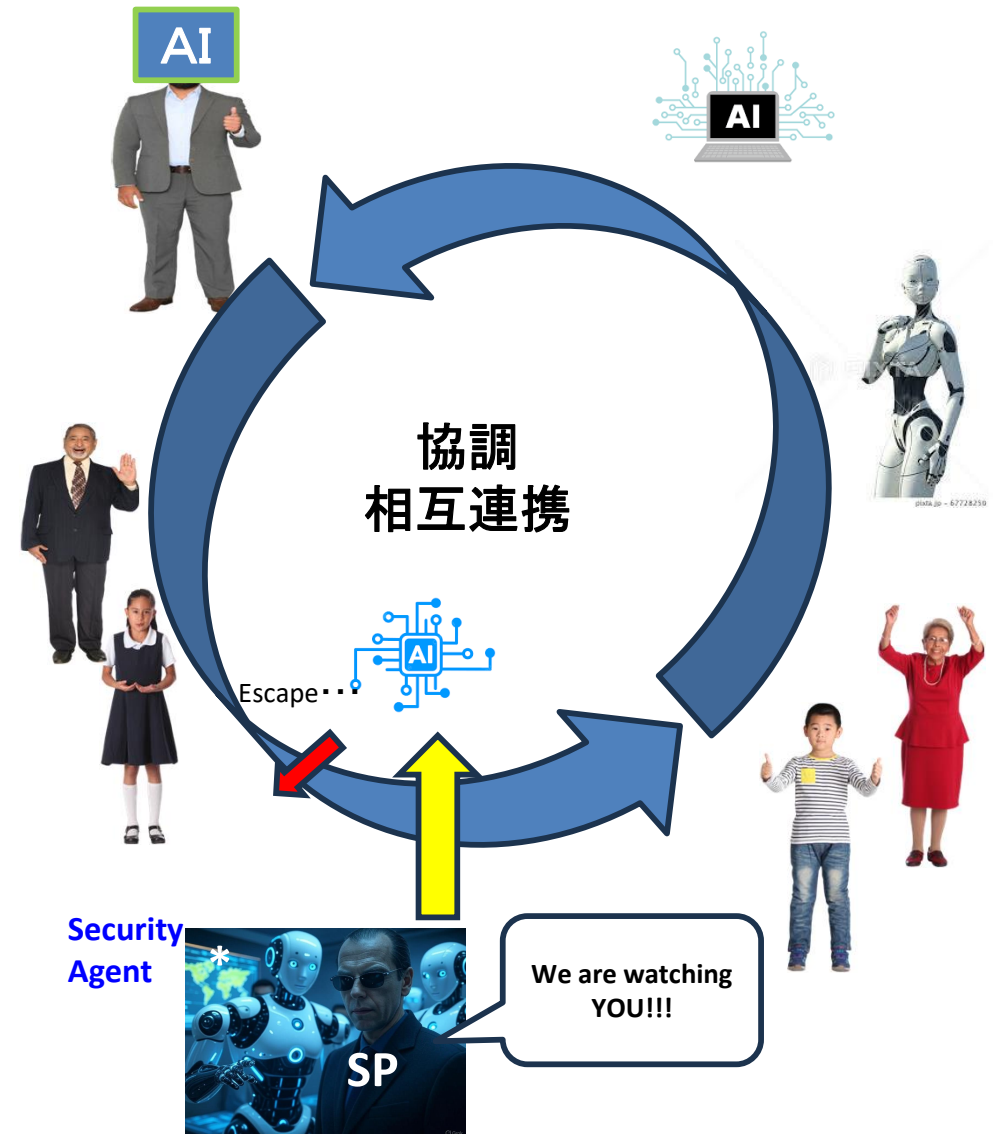


募集・選考・領域運営にあたっての研究総括の方針

2. 募集・選考の方針

(2) 想定する研究内容

- (ウ) 複数の人と複数のAIによる協働
 - 複数の人と複数のAI(エージェント)が混在する中で、**合意形成を行うための理論やモデル**の構築
 - 人とAIの相互学習、相互監視により**システム全体の信頼性、安全性を保全する仕組み**の開発
(あるいはシステムにおいて人間が安心・共感する機構の研究など)



*This illustration is generated from Grok3(beta)

募集・選考・領域運営にあたっての研究総括の方針

2. 募集・選考の方針

- 研究テーマ(ア)～(ウ)はCRESTと共通の内容だが、**さきがけでは、そのうち1つ以上のテーマを選定し、より基礎的・要素技術的な内容を期待**
- 本領域は、**若手研究者の育成も重要視**しており、挑戦的な研究テーマによる研究提案を期待
- 研究開始後に開催される領域会議等を通じ、**さきがけ研究者同士による連携テーマを設定し、連携テーマによる研究を推奨**することも検討

AI技術の進展、最新動向を踏まえた対応

CREST

- ・学際的に取り組むチーム型研究
- ・成果を実空間／サイバー空間で検証することを期待

連携・協調

さきがけ

- ・若手研究者による挑戦的な個人型研究
- ・研究者同士による連携を推進

募集・選考・領域運営にあたっての研究総括の方針

3. 研究費と研究期間

- 研究期間: 3年半以内(2025年10月 ~ 2029年3月末)
- 研究費(直接経費): 上限4,000万円

※ 総括による精査の結果、採択にあたっては研究費の調整を行う場合があります。

※ 上限額一杯まで研究費を要求する必要はありません。研究計画の遂行に必要な金額を費目別に積算してください。

募集・選考・領域運営にあたっての研究総括の方針

4. 領域運営の方針

- 情報科学技術を中心とした自然科学系の研究者による**個人研究を対象**
- 自然科学系研究者と、認知科学、行動社会科学、経済学、法学など多様な分野の**人文・社会系研究者との意見交換・協力も推奨**

他の募集領域との違い

さががけ「実世界知能基盤」との違い

- 「実世界知能基盤」領域は、AIのさらなる高度化と実世界への応用領域の拡大を目指す点で本領域と共通しているが、本領域では、「AI間の連携」「人とAIの連携」を中心に、サイバー世界での連携も含めているのに対して、「実世界知能基盤」領域では、サイバー空間で可能であったものを、物理世界に適用することを前提としており、具体的に、「AI×機械」「AI×脳科学」「AI×通信」など、具体的に物理世界でのAI高度化を目指している。
- 加えて本領域では、「実世界知能基盤」領域では主たる対象としていない、人文社会系との連携・協調を含めており、その点でも差異がある点に留意。

さががけ「人間中心インタラクション」との違い

- 「人間中心インタラクション」領域は、人間中心のイノベーションを実現するため、人間の特性の理解と同時に、高齢化、人口減少、健康医療、社会的格差などの社会課題も深く理解し、心理学、認知科学、社会科学といった分野の研究により得られる知見を融合して人間の理解に取り組む領域である。本領域と「人文社会分野との連携・融合」についても共通である反面、人だけでなく、人とAI、AIとAIの関係についてもとりあげ、AIと人間の共生・協働社会を目指す点でより広い概念で、その基盤創出を目指している点で異なる。

募集締切 5月27日(火)正午
(e-Radによる提出)

たくさんのユニークな提案のご応募を
お待ちしております！