

# 令和5(2023)年度 学術変革領域研究(A) 領域計画書

令和 4年 7月18日  
3版

審査区分	IV	整理番号	A44004	
仮領域番号	AXC5R	領域略称名	統一理論	
応募領域名	予測と行動の統一理論の開拓と検証			
英訳名	Extension and validation of unified theories of prediction and action			
領域代表者氏名	(フリガナ)	イソムラ タクヤ		
	(漢字等)	磯村 拓哉		
所属研究機関	国立研究開発法人理化学研究所			
部局	脳神経科学研究センター			
職	ユニットリーダー			
応募領域の研究概要	<p>脳の計算原理を解明し人工知能に実装することは自然科学と情報工学の最大のフロンティアである。脳は外界の動態を表す「生成モデル」を構築し予測や行動を行うとされるが、その神経基盤は未解明である。本領域の目的は、最先端の計測技術を用いてサカナ・ネズミ・サル・ヒトの脳から神経細胞の活動を高精度・大規模に取得し、データから脳が持つ生成モデルをリバースエンジニアリングすることで、脳の統一理論を検証し、その神経基盤を解明することである。そのために神経科学と情報工学の融合領域を創成する。実験データから構成した生成モデルが動物の行動や脳活動、学習過程を予測できるかテストすることで理論を検証し、改良・拡張する双方性の連携を中核とする。脳の統一理論を確立し、ヒトのように考える人工知能や精神疾患の早期診断手法の開発への道筋を開拓する。</p>			
最も関連の深い小区分		(1)	(2)	(3)
	小区分	61030	51020	
	小区分名	知能情報学関連	認知脳科学関連	
研究の対象	(1) 既存の学問分野の枠に取まらない新興・融合領域の創成を目指すもの。			
過去の採択領域等からの発展				
開示希望の有無	審査結果の開示を希望する			

**（a）研究領域の目的、背景など**

本研究領域の目的や背景などについて記述すること。特に次の点について具体的かつ明確に記述すること（4頁以内）。

- 1) 本提案のどのような点が、学問分野に新たな変革や転換をもたらし、既存の学問分野の枠に収まらない新興・融合領域の創成や当該学問分野の強い先端的な部分の発展・飛躍的な展開などが期待できる基礎的研究（基礎から応用への展開を目指すものを含む）に該当するのか
- 2) 領域研究の応募に至った背景・経緯
- 3) 国際的な研究動向から見た優位性、又は我が国固有の分野若しくは国内外に例を見ない獨創性・新規性を有する（期待される）研究領域であるか（これまでの研究活動（研究水準の現状・実績）等を踏まえて記述）

**（1）本提案により新興・融合領域の創成と当該分野の飛躍的展開が期待できる点**

脳の計算原理を解明し、生物のように考える人工知能を構築することは、自然科学と情報工学における最大のフロンティアである。人工知能はこれまで神経科学からヒントを得ることで発展し、特に特徴抽出と強化学習において大きな成功を収めてきた。しかしそれでも、人工知能と人間の知能の間には未だ大きなギャップが存在する。神経科学の成果を用いて、人工知能に人間の知能のような高いデータ効率、計算効率、エネルギー効率、常識に基づく推論力、想像力を持たせることは今日の重要な課題である（Hassabis et al., 2017）。

近年注目される脳の情報理論である**ベイズ脳仮説**（Doya et al., 2007）や**自由エネルギー原理**（Friston, 2010）は、「生物は、外界が感覚入力を生成する仕組みを表現する**生成モデル**（世界モデル）を脳内に構築することで、**ベイズ推論**と呼ばれる統計学的な推論を自己組織化的に行い、外界に適応している」ことを提唱している。それにより生物は、外界の隠れた状態変数（隠れ状態）の推論や未来の状態の予測を行い、自らの行動によって環境を最適化させるような能動的な推論を実行すると考えられている（**図1**）。とりわけ自由エネルギー原理は、これらの知覚・学習・行動が自由エネルギーと呼ばれる単一のコスト関数を最小化するように決まるとしており、脳の統一理論の候補と目されている。しかしながら、理論の抽象度が高く、生命現象との対応づけはこれまで困難な未解決問題であった。

脳科学においては、近年の計測技術の発達により、得られるデータの高精度化・広域化・長期安定化が進み、細胞種の特特定や、複数の領野・レイヤーから大規模データの取得が可能になってきている。さらに操作技術の進展も目覚ましく、生きた脳を細胞レベルでリアルタイムに操作し観測する技術も飛躍的な進化を遂げた。これらの実験技術の発展を最大限に活かせる形で、理論面においても最近、代表の磯村らにより、標準的な神経回路ダイナミクスから背後にあるコスト関数を逆算することで、神経回路ダイナミクスと生成モデルを一対一に対応づける、**生成モデルのリバースエンジニアリング**手法が開発された（Isomura & Friston, 2020; Isomura et al., 2022）（**図2**）。この手法を用いることで、実験データからシステムティックに生成モデルの構造や変数・パラメータの値を同定することが実現可能になった。このような理論・実験双方の技術の進展により、脳と心の理解に極めて重要な「生物が持つ生成モデルを実験データから同定すること」が最近になり初めて現実的になった。

そこで本領域の目的は、脳の計算原理と目される仮説を実験的に検証し、その神経基盤を解明することである。そのために、最先端の計測技術を用いてサカナ・ネズミ・サル・ヒトの脳から神経細胞の活動データを高精度・大規模に取得し、データから脳が保有する生成モデルをリバースエンジニアリングする（**図2**）。そのために、神経科学と情報理論の研究者が融合した領域を創成する。実験データから構成された生成モデルに基づき、ベイズ推論や自由エネルギー原理に従い予測や学習、行動計画を行うエージェントは、それ自体が元の動物の脳をデジタル化した脳型人工知能として動作する。したがって、動物と人工知能に同じ感覚入力を与えたときに、その人工知能が実験動物の行動や脳活動、学習過程を予測できるかをテストすることで理論の妥当性を検証し、改良・拡張する双方性の連携を領域の中核と

【（a）研究領域の目的、背景など（つづき）】

①自由エネルギー原理では、外界(飼い主)は、生成モデルに従い隠れた状態変数(隠れ状態)から感覚入力を生成すると考える。

感覚入力  $o$

②生物(イヌ)の脳は、自由エネルギー  $F$  を最小化するように隠れ状態の期待値を予測すると考えられている。

$$F = \underbrace{E_Q[-\log P(o|s, \theta)]}_{\text{予測誤差}} + \underbrace{D_{KL}[Q(s)Q(\theta)||P(s, \theta)]}_{\text{複雑性}}$$



③また、好みの入力(エサ)を得るために将来の自由エネルギー  $F$  を最小化する行動を能動的に推論し選択すると考えられている。

図はIsomura (2022)より改変

図1：自由エネルギー原理による予測と行動の最適化

する。本領域で、知覚的な予測と行動の計画・生成を統一的に説明可能な脳の統一理論を確立し、ヒトのように考える人工知能や精神疾患の早期診断手法の開発への道筋を開拓する。

予測誤差の最小化により知覚を説明する**予測符号化**モデルやベイズ脳仮説など様々な脳モデルが提唱されているが、それらは脳がどのように生成モデルを獲得しているかという問題に帰着できる。すなわち、予測と行動の機序の解明とは、対応する生成モデルの獲得メカニズム、つまり推論・学習アルゴリズムの同定である。本領域では、**B00班**が取得する神経細胞の活動データに基づき**A00班**が予測と行動を担う生成モデルをリバースエンジニアリングし、**両班**が協力し理論予想と実験データとの整合性を調べ、課題実行時の脳活動・行動がベイズ推論や自由エネルギー原理の帰結であるか検証する。さらに、神経回路の学習過程を、ベイズ推論や自由エネルギー原理に基づいた学習則により予測できるかを検証する。予測が可能であればこれらの理論の妥当性を実証することができ、逆に予測が本質的に不可能であることを示せば理論を棄却することができ、理論の修正の指針を与えることができる。

代表の磯村らは最近、生成モデルのリバースエンジニアリング手法を開発し、神経細胞の活動の実験データから神経活動の方程式を導出することができれば、その動物の脳内にある生成モデルを得ることが出来ることを世界で初めて示した。それを基に自由エネルギーを計算すると、理論から予想されるシナプス可塑性の式を導出することができる。すなわち、神経細胞の活動データがあれば、シナプス可塑性の式を求めることが出来る。このことは、原理的には、神経回路の学習過程を理論的に予測可能であることを示唆している。従って、学習前の神経細胞の活動のデータから、シナプス可塑性の式を求め、学習後のデータが予測できるかどうかをテストすることで脳の計算原理を検証することが可能になった。

このようにして、予測と行動の統一理論を検証・修正・拡張し理論体系確立への道を開拓する。シンプルな法則により、脳の様々な機能、部位、動物種にかかわらず統一的に理解することが可能になり、その変容を解析することで精神疾患を理解することにもつながる。そしてその過程で明らかになる生成モデル獲得の汎用的な仕組みを用いると、様々な用途に汎用的に適用可能な脳型人工知能を構築することができる。このような脳型人工知能は、今日のバックプロパゲーションに頼った学習よりもデータ効率がよく、生物のような高い計算効率、エネルギー効率を持つことが期待され、次世代人工知能として非常に有望である。また脳の仕組みに基づき人工知能を作ることによって脳の理解をさらに深めることができる。

【（a）研究領域の目的、背景など（つづき）】

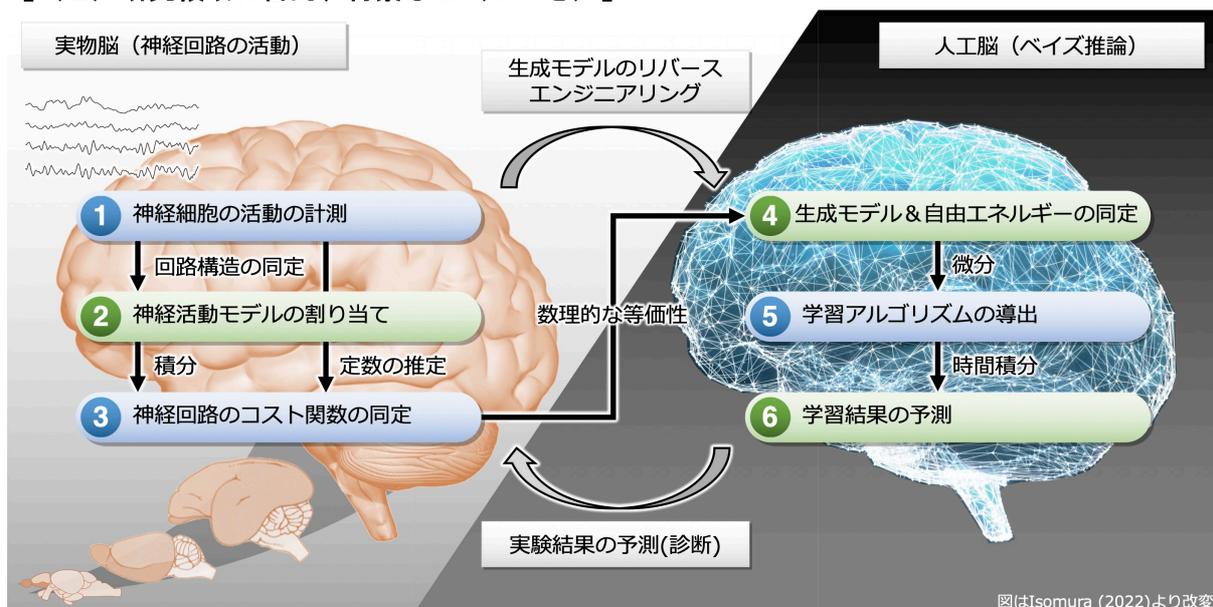


図2：生成モデルのリバースエンジニアリングを用いた統一理論の検証

（2）応募に至った背景・経緯

もともと日本の研究は計算神経科学分野をリードしてきたが、最近の深層学習では欧米の企業が中心である。90年代から川人の内部モデル、RaoとBallardの予測符号化理論、A03銅谷らのベイズ脳仮説などが出てきた中で、2006年にFristonは自由エネルギー原理を提唱した。特に行動に適用したものは**能動的推論**と呼ばれる。一方理論検証においては、A03銅谷は2016年にマウスで隠れ状態を符号化する神経集団の存在を報告し、B01岡本は2021年にゼブラフィッシュで予測誤差を符号化する神経集団の有無が行動戦略を変えることを報告した。これらは世界に先駆けた実験成果と言える。またA01磯村は2015年に培養細胞で自由エネルギー原理と整合する結果を報告し、その後Fristonと共同で理論研究に取り組む中で神経回路ダイナミクスとベイズ推論の等価性の理論を構築した。しかし完全な実験的検証にはまだ至っていない。検証を困難にする問題点としては、理論が抽象的であり生命現象との対応づけが難しいという理論側の問題点、理論の検証に必要な神経活動のデータを高精度、大規模に、長期間安定して取得するための技術的な問題点が存在したが、今日までに計画班のメンバーを中心に解決への鍵となる技術開発や成果が出てきつつある。

そこで世界最先端の実験技術を有する神経科学研究者と情報理論の研究者が脳の統一理論の検証と開拓という共通目標に向けて連携することで、学問領域を変革するような成果を出せると確信し、今回申請するに至った。**予測や行動を担う脳の計算原理を検証し実装するために**、計算神経科学・情報理論・機械学習・神経生物学・認知科学・精神医学の各分野の気鋭の若手や世界的にリードしている一流の研究者を集結させ、予測と行動に関する統一理論の確立という最終目標に向かう学際的な布陣を整え、申請するに至った。全脳の観察が可能なゼブラフィッシュ、階層的な機能構造の研究が可能なマウス、より人間に近く深層の生成モデルを持つと考えられるマーモセット、さらに実際のヒトの生体計測を配し、各々の計測・操作技術の強みを最大限活かし相補的に研究を行うことで、動物種に依らない普遍的な計算原理を解明するという領域の目標を達成するために必須の布陣となっている。

（3）国際的な研究動向から見た優位性、国内外に例を見ない独創性・新規性

代表のA01磯村は、若手でありながら、自由エネルギー原理の神経回路レベルでの検証に国際的にも早い段階で着手し成果を上げてきた。また外部評価委員の乾と共に国内での自由エネルギー原理研究を牽引してきた。自由エネルギー原理の国際的なコミュニティにおいて

## 【（a）研究領域の目的、背景など（つづき）】

も、神経回路が生成モデルを自己組織化的に獲得する過程の数理理論構築と培養神経細胞を用いた検証の両面から独創性の高い研究を推進しているユニークな存在である。今年発表した、リバースエンジニアリングの根拠となる等価性の理論を示したキー論文はAltmetricが277になるなど、理論論文としては例外的に国際的な注目度が高い。

ベイズ脳仮説や予測符号化理論が計算神経科学や認知科学の間で注目されているが、運動制御の領域で、内部モデルとして理論を発展させてきたのは外部評価委員の川人に代表される日本の計算論的神経科学である。A03銅谷も強化学習とベイズ推定のアルゴリズム開発とその脳機構の解明、神経修飾物質とモデルのパラメータの関係など先駆的な業績を残している。最近の論文では感覚野と運動野の局所回路の**推定と制御の双対性**に基づく理論を展開し注目を集めている。A02鈴木は若手でありながら、深層生成モデルに基づき画像やテキストなど複数のモダリティを扱う変分オートエンコーダを世界で最初に開発するなど国際的に高い評価を得ている。これらの研究により最近ようやく複数のモダリティを含む大規模な深層生成モデルと高次の知覚を担う脳活動データを対比させて研究することが可能になってきた。

実験技術の国際的な優位性としては、B01岡本はゼブラフィッシュ成魚のバーチャルリアリティシステムを開発し、予測誤差を符号化する神経集団を同定するなどインパクトの高い研究を行なっている。B02蝦名は若手でありながら、2光子イメージングと光遺伝学によって運動中の非ヒト霊長類（マーモセット）大脳皮質運動野で神経活動を計測・制御する技術の開発に世界で初めて成功しており、特に前肢運動中のイメージング技術については現在も独占先行状態が続いている。同じく若手のB03小松は、世界で初めてマーモセット用の広域皮質脳波電極および慢性回路操作デバイスの開発に成功している。ヒトにおいて予測機構を検討するために、B04高橋は所属先の脳外科と協力して高い時間解像度と電極位置の正確性を有する定位的頭蓋内電極による予測課題中の脳計測を実施できる体制を整備しており、世界でも非常に優位性の高い計測技術である。

このように計画研究の代表者は、いずれも各専門領域で第一人者と言われる実績があり、国際的なプレゼンスも高い。また、脳の計算原理・統一理論の構築・検証という共通の関心を有した研究者が集まっており、問題意識が共有できている布陣となっている。

感覚入力に基づき外界の生成モデルを構築して将来を予測することは、神経科学や人工知能の研究における最近のホットトピックである。脳はこの能力を自然に獲得することができ、それを説明する理論として、自由エネルギー原理は大きな注目を集めている。しかしその神経基盤は、理論の抽象度が高いため、まだ完全には解明されていない。その解明には、実験データに基づく生成モデルのリバースエンジニアリング手法が不可欠であり、この手法により、自由エネルギー原理に現れる変数と実験観察される神経活動などの変数を精密に対応づけることが初めて可能になった。この手法を生きた脳の神経活動データに適用し、生成モデルの同定と学習過程の予測を行うことは、国内外に類似研究がなく独創的である。これが実現すれば、脳に近い生理学的に妥当な人工ニューラルネットワークの実装が可能になり、自由エネルギー原理の実証にとどまらず、神経科学・人工知能分野に広く大きなインパクトを与えることが期待される。より脳に近い次世代人工知能の開発に重要な役割を果たし、介護ロボットなど人間と密接に関わる人工知能の開発に重要な技術になると期待される。

## 業績リストに含まれるものを除く参考文献

- Friston K. The free-energy principle: a unified brain theory?. *Nat Rev Neurosci* 11, 127-138 (2010).  
 Hassabis D., Kumaran D., Summerfield C. & Botvinick, M. Neuroscience-inspired artificial intelligence. *Neuron* 95, 245-258 (2017).

**（b）領域マネジメント体制**

領域マネジメント体制について記述すること。特に次の点について具体的かつ明確に記述すること（3頁以内）。

- 1) 領域代表者の研究領域の運営や推進に関するビジョン及び準備段階における状況
 

複数の研究者をまとめ、領域推進に当たって研究組織の総合力を発揮するために、研究とは別に、リーダーである領域代表者が、どのような構想を持って円滑な組織運営をし、研究領域を推進するかについて、基本的な考え方を明らかにすること。また、応募に至るまでの準備研究や事前調査段階でのマネジメント体制の状況にも触れつつ、記述すること。なお、領域代表者が総括班以外の計画研究のいずれにも、研究代表者、研究分担者又は研究協力者として参画しない場合は、研究を行わない組織とされている総括班にのみ属している状況において研究領域全体の把握、マネジメントを行うための方策について、具体的に記述すること。
- 2) 領域代表者を中心とした領域推進に十分貢献できる研究者による有機的な連携体制
 

研究領域の構成について記述すること。また、概念図を用いつつ、研究領域全体の組織図等により、総括班、各研究組織の役割及び活動内容等を明確に示すこと。
- 3) 総括班、各研究組織の役割及び活動内容等
  - ・研究領域を推進するに当たっての総括班の運営方針、役割、研究組織の構成、個々の構成員の役割、活動内容等について基本となる考え方を説明すること。構成員については、構成員の氏名、専門分野及び研究代表者・研究分担者の別を記入すること。その際、どのように本研究領域の全体的な研究方針を策定し、各研究項目、各計画研究の企画調整等を行っていくのか明らかにすること。
  - ・総括班における研究支援活動（国際活動支援や本研究領域内で共用するための設備・装置の購入、開発、運用又は実験試料や資材の提供など）を効率的かつ効果的に行うための体制及び計画について記述すること。

**1) 領域代表者の研究領域の運営や推進に関するビジョン及び準備段階における状況**

**【領域推進のビジョン】**

**原理検証への明確な道筋の設定と共有：**ブラックホールの撮影による重力理論の検証のように、物理学では数理理論構築から実験的検証の流れが成功しているが、現在の脳科学ではそのような明確な理論検証は為されていない。検証という目標達成には、理論と実験の研究者が互いの研究を理解し連携した上で、基礎となる考え方や検証の具体的方針を共有し取り組む必要がある。そのためには歴史的な成功例がそうであるように、数理的な基礎に基づき理論体系を構築した上で検証実験を計画する必要がある。代表は博士課程までは実験をしていてその後理論に進んだため両方の研究を解る立場にある。同じく理論と実験の両方で実績があるA03銅谷・B03小松をはじめ連携して原理検証に取り組むことに高い意欲を持つベストメンバーが揃っている。理論が先導し、ベイズ推論・自由エネルギー原理という共通の定式化を使い共通言語・共通方針を設定することで検証のプロセスを明確化し共有していく。

**若手研究者の育成支援：**計算神経科学関連分野の若手研究者育成も本領域の重要なミッションである。本領域は計画班も過半数が若手研究者から構成されている。日本の現状として、計算神経科学を行う理論研究者・データサイエンティストが圧倒的に不足しており、実験と理論の連携ができる研究者も限られている。また良い研究を行うためにはキャリアの早い時期から良い理論に出会っておくことが大切である。本領域を中心に、理論と実験の研究者が集まり計算原理について学び互いにディープな議論ができる場を作ることで、その中から独創的なアイデアが数多く出るようにしたい。また斬新なアイデアを研究として実現できるように総括班を中心に支援していく。

**【準備段階における状況】**

計画班のメンバーにはこれまでのいくつかの新学術領域で研究を行ってきたメンバーも多く、合同シンポジウムで、相互の研究発表に接することは多かった（銅谷・高橋・岡本・小松・鈴木）。また、実験と理論研究の橋渡しと、若手研究者の育成を目的とする「脳と心のメカニズム」ワークショップは銅谷や高橋らが牽引してきた。

自由エネルギー原理は非常に学際的で幅広い分野を対象とする理論であり、かつ発展途上であるため、多くの・研究者・学生が習得に手間取り、その意義が十分に理解されていない。そこで本申請のメンバーはこれまでに自由エネルギー原理に関わる研究会・チュートリアル・学会誌の発行など学際的な議論を喚起する施策を継続して行ってきた。

- ・2018年9月「自由エネルギー原理」特集号(日本神経回路学会学会誌)（寄稿 磯村、乾）

【（b）領域マネジメント体制（つづき）】

- ・2019年7月 日本神経科学大会 大会長 岡本、プレナリーレクチャー Friston、教育講演 磯村
- ・2020年2月 京都大学国際シンポジウム「Computational Principles in Active Perception and Reinforcement Learning in the Brain」（登壇 磯村、小松）
- ・2021年12月 自由エネルギー原理の国際シンポジウム（共同主催者 磯村、登壇 岡本、Bogacz）

このように、計画班のメンバーはこれまで複数のグラント、シンポジウム、学会等で一緒に活動しており、相互の研究の中身やスタンスについて十分に理解している。今回の申請に先立ち、繰り返しWeb会議やメッセージツールを利用して、共通言語、共通テーマを設定し、最終目標の実現のための具体的な共同研究の構想を練ってきた。相互が信頼し、最終目標に向かう機運は高まり、準備は順調・円滑に進んでいる。

2) 領域代表者を中心とした領域推進に十分貢献できる研究者による有機的な連携体制

**A00：情報領域の脳知能研究（理論）、B00：物理領域の脳知能研究（実験）**という2つの計画研究グループを設定する。A00は計算神経科学・情報理論・機械学習、B00は神経科学・神経生理学・精神医学を軸としており、グループ内での研究フレームワーク共有に加えてA00理論とB00実験の連携を通じて脳の統一理論の開拓と検証という領域の目標に貢献する体制をとる（図3）。階層的にサカナ（B01）・マウス（A03）・マーモセット（B02, B03）・ヒト（健常者・精神疾患患者）（B04）の生体計測を配し、予測と行動を支える神経基盤の普遍的な理解を目指す。計測範囲や動物種の異なるデータの統合により、相補的に生成モデルの推定精度を向上できる。例えば、動物を使って広範囲・高精度のデータを得ることで網羅的な計測が難しいヒトの生成モデルの解析に活用できる。このように動物種間で保存される普遍的な原理を明らかにする。そしてそれに基づく理論の拡張（A01, A03）とその人工知能への実装（A02）を目指す。領域の目標達成にとりわけ重要な課題として、次の3つを設定し、複数の計画班が共通する実験・解析パラダイムを用いて取り組むことで普遍的な原理の解明を円滑に行える体制を整えた。カッコ内は主にその課題に取り組む計画班である。

**課題①高次な視聴覚入力の予測（B03小松, B04高橋, A02鈴木, A01磯村）** マーモセットとヒトを用いて高次な視覚・聴覚入力の予測を担う深層予測符号化モデルの脳内機序解明に取り組む。

**課題②離散状態空間内の行動計画・意思決定（B01岡本, B03小松, A03銅谷, A01磯村）** 動物は報酬を得るために数手先まで計画し行動する必要がある。サカナ・マウス・マーモセットが迷路課題を解くときの脳の計算原理の解明に取り組む。

**課題③連続状態空間内の行動計画・制御（B02蝦名, A03銅谷, A01磯村, A02鈴木）** 現実世界のエージェントは適切に行動計画し身体を制御する必要があるが、その機序は多くが未解明である。マウス・マーモセットのレバー押し課題を用いてその計算機序の解明に取り組む。

A00班は、脳がベイズ推論を実行しているとする理論から予想される回路構造と実験デー

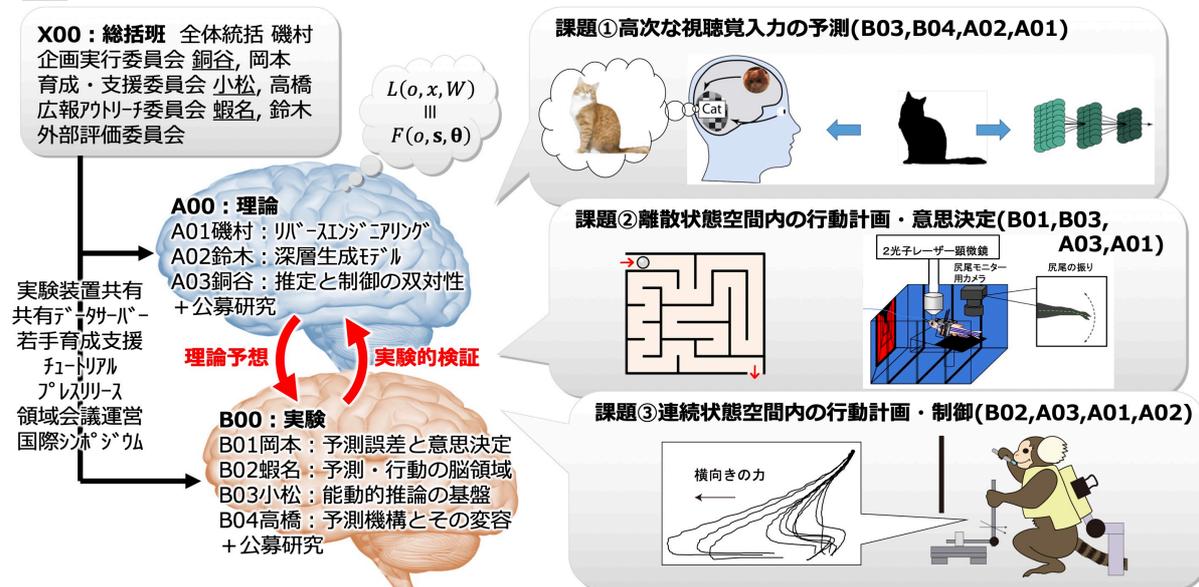


図3：領域の連携体制

【（b）領域マネジメント体制（つづき）】

タから推定される回路構造の比較（A03）や、学習前のデータのみから生成モデルを構成して動物の学習後の神経活動や行動を予測できるかをテストする（A01）ことで、理論の検証を行う。さらに得られた知見を深層生成モデルに基づく人工知能に実装し、脳の計算機序を取り入れることで既存の機械学習手法よりも優れた新たなアルゴリズムを開発する（A02）。このように、領域代表者を中心とした領域推進に十分貢献できる研究者による有機的な連携体制が構築され、脳の計算原理の解明という最終ゴールに向かう体制を整備した。

**3）総括班、研究組織の役割及び活動内容**

総括班の役割は、メンバー間の知識と情報の共有を図り実のある共同研究を実現すること、新領域の研究を担う若手の育成、女性研究者の支援を図ること、そして生まれた研究成果を広く社会に広め応用につなげていくことである。以下の組織構成で活動を展開する。

**<組織構成>**

研究代表者 磯村拓哉(脳情報理論) 全体統括

**企画実行委員会**

研究分担者 銅谷賢治(計算神経科学) 企画実行委員長、連携研究者 岡本仁(神経科学)

**育成・支援委員会**

研究分担者 小松三佐子(神経科学) 育成支援委員長、連携研究者 高橋英彦(精神医学)

**広報アウトリーチ委員会**

研究分担者 蝦名鉄平(神経科学) 広報連携委員長、連携研究者 鈴木雅大(機械学習)

**評価委員会**

国内外評価委員会 連携研究者：川人光男(ATR)、乾敏郎(追手門学院大学)、酒井宏(筑波大学)  
国際アドバイザーボード：Karl Friston(University College London)、Anil Seth(University of Sussex)、Rafal Bogacz(University of Oxford)

**<活動内容>**

**企画実行委員会**：年に2度の領域会議を開催し、計画研究、公募研究のメンバー間の知識共有をはかり、共同研究をうながす。2025年度と2027年度には国際ワークショップを企画し、研究成果を世界に発信するとともに、海外の先進的な研究者とのネットワークの強化を図る。

**育成・支援委員会**：若手と女性で構成され、メンバーからのニーズ収集・意見交換の場を設定する。Webも活用し、企画会議にも若手に参加してもらう。毎月、計算理論とアルゴリズム、生体情報の可視化や計測等のテーマでチュートリアルを開催し、夏にはサマースクールやハッカソンを開催し、その動画配信もする。こうして、専門外の知識と技術の向上と若手の育成を図る。またチュートリアルや研究交流会で積極的に若手を登壇させることで、若手が認知される機会を増やしキャリア構築を支援する。メンバー間のデータ解析とモデリングの共同研究の支援、さらに実験で得られた貴重なデータやアルゴリズムをより良く活用するため、データベースを構築・運営する。海外先進ラボとの間の若手・女性研究者を中心に派遣する。公募研究では若手を積極的に採用し、希望者にはメンター制度を導入し、自身の分野や他分野のシニアとアドバイスや共同研究の相談を気軽にできるようにする。

**広報アウトリーチ委員会**：当該領域の研究活動と成果を幅広く広報するため、Webサイトの立ち上げと更新、ニュースレターの発行、プレスリリース等を行う。また、多くの国民、特に学生の関心を喚起するためのアウトリーチ活動を企画実行する。さらに研究成果を社会応用につなげるため、産業、医学、教育界との連携をめざしたワークショップなどを企画する。

**評価委員会**：国内外評価委員に領域会議への出席を依頼し、活動の進行・達成状況に対して評価を受ける。国際アドバイザーボードは、関連分野を牽引する3名に依頼し、活動評価、国際共同研究の推進やシンポジウム開催について助言、協力を得る。

以上の活動を統括するため、研究内容の報告を主目的とする領域会議に加え、総括班会議を随時開催し各企画の構想の練り上げと進捗の把握を図る。

（c）領域推進の計画概要

当該研究領域の推進に当たっての、

- 1) 基本的な研究戦略（研究項目を設ける場合にはその考え方を含む）
- 2) 研究領域における具体的な研究内容（研究項目を設ける場合には項目ごとの研究内容を含む）
- 3) 各研究項目、各計画研究の必要性及び研究項目間、計画研究間での有機的連携を図るための具体的方法について、必要に応じて概念図等を用いて示すこと（2頁以内）。

1) 基本的な研究戦略

最先端の計測技術を用いてB00班がゼブラフィッシュ・マウス・マーモセット・ヒトの脳から神経細胞の活動データを大規模・高精度に取得し、A00班がデータから脳が保有する生成モデルをリバースエンジニアリングし理論予想と比較することで、両班が連携し脳の計算原理と目される仮説を実験的に検証し、その神経基盤を解明する。実験データから構成した生成モデルを脳型人工知能として動作させ、実験動物の行動や脳活動、学習を予測できるかをテストすることにより理論の妥当性を検証し、改良・拡張する双方性の連携を中核とする。

2) 研究領域における具体的な研究内容

**計画研究A01**：A03と連携しリバースエンジニアリング手法を連続状態空間モデルに拡張する。B00班が取得する神経細胞の活動データから生成モデルをリバースエンジニアリングし、理論モデルとの整合性を検証する。B01, B02, B03が取得する学習前のデータのみから生成モデルを構成し、A02と連携してそれを脳型人工知能に実装、それを自由エネルギー原理に従って学習させ、動物の学習後の神経活動や行動を予測できるかをテストし理論の検証を行う。

**計画研究A02**：A01と連携してリバースエンジニアリング手法によって脳の生成モデルの仮説として深層生成モデルを設計する。このモデルはB00班と連携して脳から得られるデータから学習し、神経活動等が予測できるかによって妥当性を検証する。また、A03と連携して能動的推論に基づく行動制御を実現する深層生成モデルについての研究を進め、A03とB02によって得られる実験結果に基づき検証を行う。そして、これらの深層生成モデルを学習する方法として予測符号化に基づく生物学的に妥当な形でのパラメータ更新方法を研究する。

**計画研究A03**：知覚のためのベイズ推定と行動のための強化学習が共通の計算手法で実現可能であるという「推定と制御の双対性」を手がかりに、大脳皮質の感覚野と運動野の作動原理を相補的に解明することをめざす。A01, A02と連携して脳回路で実現可能なアルゴリズムの候補を探索するとともに、マウスの感覚野と運動野の多層の神経細胞の活動を記録し、それぞれの神経回路での情報表現と処理のダイナミクスを解析する。

**計画研究B01**：体が透明な変異体のゼブラフィッシュ成魚を閉ループ仮想現実空間で様々なルールのもとで遊泳させ、その学習過程での神経細胞の活動をイメージングすることに成功し(Nature Commun, 2021)、予測と予測誤差に基づく行動制御を可能とする仮説的神経回路モデルを提唱するに至っている。本研究では、神経系の様々な階層での予測と予測誤差の計算の実体と、それが脳の階層的相互作用によって行動の制御にどのように使われているのかを明らかにする。

**計画研究B02**：変動する環境の中で運動制御課題を実行中の小型霊長類コモンマーモセットの大脳皮質で、領野・神経細胞体・神経軸索という様々なスケールにおける神経活動を計測・制御することで環境の予測と行動選択、運動制御のための脳領域ネットワークの実態を明らかにする。また、A00班と共同で、どのようなアルゴリズムが実際の脳活動に対応し、このような機能を実現し得るのかを検証する。

**計画研究B03**：予測情報処理および予測に基づく生成モデルの更新がサル大脳皮質でどのように行われているのかを検証する。広域皮質脳波電極を用いて感覚野だけでなく前頭葉を含む広域の神経活動を学習前後にわたり計測する。それらのデータを領域内で広く共有し、理論検証および種間比較を通して脳内情報処理の統一的理解を目指す。

## 【（c）領域推進の計画概要（つづき）】

**計画研究B04**：予測情報処理がヒト脳でどのように行われているのか、精神疾患では、どのような変容が認められるのかを検証する。fMRI信号と深層学習モデルを組み合わせて、間接的に予測成分を検討することに加え、定位的頭蓋内脳波を活用し、直接的に予測成分の抽出を試みる。

**3）各計画研究の必要性及び研究項目間計画研究間で有機的連携を図るための具体的方法**

本領域の目標達成に重要な共通課題を3つ設定する。カッコ内の計画班が主に担当する。

**課題① 高次な視聴覚入力の予測（A01, A02, B03, B04）**：B03小松、B04高橋がそれぞれマーモセットとヒトを用いて高次な視覚・聴覚入力の予測を担う深層予測符号化モデルの脳内機序の解明に取り組む。影絵から動物の姿を予測・補完することは高次の概念の知覚や想像力の典型例である。現状の人工知能が未だ生物に及ばないこれらの能力を担う神経基盤を解明する。得られた知見からA02鈴木が深層生成モデルの新たな構造／学習則を導き人工知能に実装することでより脳の理解を深める。

**課題② 離散状態空間内の行動計画・意思決定（A01, A03, B01, B03）**：迷路やパズル、ボードゲームは離散状態空間内の行動計画最適化の典型例である。動物は報酬（エサ）をもらうために、あるいは罰から逃れるために、1手先だけではなく数手先まで計画して行動する必要がある。そこで、動物に迷路課題を解かせ、そのときの神経細胞の活動を計測し解析することで脳の計算原理を解明する。B01岡本がゼブラフィッシュで、A03銅谷がマウスで、B03小松がマーモセットで、共通性のある十字迷路課題を用いて計算機序の解明に取り組み、そこから普遍的な原理を導き出す。近年、深層学習を用いてこの研究分野は大きく進展しているが、データ効率、常識に基づく推論、想像力など生物に及ばない点も多々あるため、脳の仕組みを解明することでさらなる発展が期待できる。これらは一般に部分観測マルコフ決定過程でモデル化でき、標準的な神経回路のダイナミクスとの精密な対応関係が確立されている。実験的に取得する神経細胞の活動データからA01磯村が生成モデルをリバースエンジニアリングし、A, B班が連携して学習を予測させることで妥当性を検証する。例えば正確な（2層）生成モデルを持つ個体と簡易（1層）生成モデルを持つ個体では、学習・行動戦略に違いが生じるはずであるためそれを検証する。得られた知見から新たなモデル構造／学習則を導く。

**課題③ 連続状態空間内の行動計画・制御（A01, A02, A03, B02）**：現実世界のエージェントは身体を有しそれを適切に制御する必要がある。その状態数は離散よりもはるかに大きく、そのため長期の予測や計画はより困難である。脳は一体どのようにこれを実現しているのだろうか？これはロボット制御への応用にも重要である。A03銅谷がマウス、B02蝦名がマーモセットにおいてコントローラー押し課題を用いてその計算機序の解明に取り組む。連続状態空間の脳内表現としては予測符号化モデルや推定と制御の双対性に基づく銅谷モデルが有力である。リバースエンジニアリングを連続状態空間モデルに拡張することで、データから推定した回路構造と理論モデルの整合性を確認し、さらにその予測性能を検証する。得られた知見から深層生成モデルの新たな構造／学習則を導き人工知能に実装する。上記の離散状態空間の問題も本来は連続状態空間内の最適化を粗視化しているに過ぎないため、最終段階としては迷路課題のような複雑な課題も連続状態空間で扱えるように拡張し統一理論を構築するとともに、それを脳型人工知能に実装することでより脳の理解を深める。

とりわけ、ある実験データに基づき構成した生成モデルが、それ以外の実験データの予測にも利用可能かは、理論の妥当性・普遍性の検証に重要である。B01, B02, B03が学習前から学習後まで継続的に神経細胞の活動を計測し、学習前のデータのみから構成した生成モデルが動物の学習後の神経活動や行動を予測できるかテストすることで理論の検証を行う。

（c）領域推進の計画・方法①

領域推進の計画・方法について記述すること。特に次の点について具体的かつ明確に記述すること（2頁以内）。

- 1) 研究領域及び各計画研究の具体的な達成目標
- 2) 1) を実現する具体的な計画・方法（共同研究、設備の共用化、若手研究者の育成等の取組を含む）  
研究目的を達成するための具体的な研究計画・方法について記述すること。
- 3) 国内外への情報発信などの取組内容

1) 研究領域及び各計画研究の具体的な達成目標

本領域の目的は、動物やヒトの詳細な脳活動記録を通じて、予測と行動を担う脳の計算原理を検証し、生物学的基盤を明らかにするとともに、人工知能への実装を通してさらなる脳の理解を目指すものである。研究領域の具体的な達成目標として、代表的な理論仮説とそれに対して計画班が連携して取り組む検証の達成目標を次に示す：

理論仮説	検証の達成目標
<b>ベイズ脳仮説</b> ：脳がベイズ推論を行っているという主張。神経細胞の活動が隠れ状態の事後期待値を符号化していると考えられている。	課題①②③において計測した神経細胞の活動と、ベイズ推論を実行するエージェントが持つ隠れ状態の事後期待値との整合性を確認する。
<b>自由エネルギー原理</b> ：脳が自由エネルギーを最小化することでベイズ推論を行い、知覚・学習・行動を最適化しているという主張（図1）。	課題①②③の実験データから生成モデルと自由エネルギーを計算し、その生成モデルに基づきベイズ推論を実行するエージェントが動物の行動・脳活動・学習過程を予測できることと、学習により自由エネルギーが減少することを確認する。
<b>予測符号化</b> ：脳は予測誤差を最小化することで知覚を行っているという主張。大脳皮質には期待値と予測誤差を符号化する別々の神経細胞集団がいると考えられている。	課題①における神経細胞の活動を計測し、期待値と予測誤差を符号化する神経細胞集団を見つけ出し、高次な知覚予測を行う神経回路構造と理論の対応を明らかにする。
<b>能動的推論</b> ：脳は将来の自由エネルギーを最小化するような一連の行動をベイズ推論することで行動計画や行動制御を最適化しているという主張。	課題②により、迷路課題実行時の行動計画・意思決定を担う神経細胞集団の活動が自由エネルギー最小化により説明・予測可能であることを確認する。
<b>推定と制御の双対性</b> ：感覚野において予測符号化を担う局所回路と双対となるような回路構造が運動野にも存在し、それにより行動の最適化が行われているという主張。	課題③の連続状態空間のタスクにおいて、感覚野と運動野の各層の神経細胞の活動を計測し、仮説により対応づけられた情報表現が行われていることを確認する。

これらの目標達成のため各計画班は以下を実行する：

**計画研究A01**：神経細胞の活動データから生成モデルをリバースエンジニアリングして脳型人工知能として実装することで、動物と人工知能に同じ感覚入力を与えたときに動物の行動・脳活動・学習過程を予測することが可能な、神経生物学的にもっともらしい脳型人工知能を創出する。それにより脳の計算原理を数理的に明らかにする。

**計画研究A02**：深層生成モデルの構造として脳に基づく仮説を実装し、脳から得られたデータに基づき学習し予測することで、脳の構造に関する仮説の妥当性を評価する。また、深層生成モデルによる能動的推論に基づく行動の抽象化と制御について研究を進め、脳における推定と制御の双対性を立証する。そして、これらの深層生成モデルを予測符号化に基づく生物学的妥当な形で学習する方法を研究し、脳型人工知能の実現と脳の理解を目指す。

## 【（c）領域推進の計画・方法①（つづき）】

**計画研究A03**：マウス用のジョイスティックでの到達課題において、横方向の力場の方向と強度を確率的に変化させながら、体性感覚野および一次運動野の各層の神経細胞の活動をプリズム型内視鏡と多色カルシウムイメージングにより計測し、身体や外界の状態の尤度、事前確率、事後確率や、報酬、行動価値、状態価値などの情報が、推定と制御の双対性にしながら形で表現されているかを検証し、必要に応じて仮説の修正改善を行う。

**計画研究B01**：神経系の様々な階層での予測と予測誤差の計算の実体と、その階層的相互作用によって行動の制御にどのように使われているのかを、脳全体が小さいが、その基本的構造が哺乳類と保存されているゼブラフィッシュを利用して、広範な脳部位の神経活動の同時計測を行うことによって、明らかにする。

**計画研究B02**：環境の予測から行動制御に至る一連の脳機能に必要な大脳皮質広域での神経活動がどのような神経細胞個々の活動の結果生じるのか、他領域からのどのような入力の結果生じるのかを明らかにする。また、A01、A02、A03と共同で、実際の生体脳における階層的な領域間ネットワークの変動がどのような神経回路モデル・アルゴリズムによって説明され得るのかを検証し、環境の予測と行動選択・制御のための脳機能の実体を明らかにする。

**計画研究B03**：①視覚・聴覚の予測課題、②離散状態における選択課題をそれぞれ実行中のマーモセットから、広域皮質脳波電極を用いて感覚野だけでなく前頭葉を含む広域の神経活動を計測する。それらのデータを領域内で広く共有し、理論検証および種間比較を通して脳計算原理の統一的理解を目指す。

**計画研究B04**：視覚物体認知の際のfMRI信号と深層学習モデルの内部信号とを対比し、予測信号を分離し、予測情報処理に伴う信号を同定する。統合失調症や発達障害患者に対し同様の方法を適用し、精神疾患の予測機構の異常の神経基盤を明らかにする。定位的頭蓋内脳波を用いて、視覚・聴覚の予測課題・選択課題での脳情報を計測し、前頭葉からの予測成分の抽出をする。

**2）1）を実現する具体的な計画・方法（共同研究、設備の共用化、若手研究者の育成等）**

代表的な共同研究の具体例は、「3）各計画研究の必要性及び研究項目間計画研究間で有機的連携を図るための具体的方法」で示した。動物実験・人間を対象にコンセプトの似た実験課題を導入し実施する。総括班では、領域内共用の設備として2光子顕微鏡を購入する。共用データベースも構築し、実験動物のデータや人間から得た生体情報は二次利用が可能な同意を得て領域内で共同利用する。データ解析手法や数理アルゴリズムの共有も図っていく。

代表のA01磯村はかつて脳科学若手の会の世話人を務め、B04高橋は日本若手精神科医の会の創設メンバーで、若手とのネットワークを有しニーズの拾い上げが可能である。育成支援委員長のB03小松は、女性からの目で、女性研究者のニーズにあった細やかな企画や支援をしていく。若手研究者の海外派遣、海外拠点からのインターンシップ受け入れ、拠点間での定期的な情報共有、異分野のためのチュートリアル、海外拠点とのワークショップ等、企画・運営をバックアップし、世界規模での活躍が期待される若手研究者の育成に注力する。

**3）国内外への情報発信などの取組内容**

関連分野で国際的に発信力のある研究者で構成される国際アドバイザーボードを設置し、国際的なコミュニティに情報発信するとともに、情報のインプットも強化する。国際アドバイザーボードとすでに構築されている海外連携拠点間で、学術発展・人材育成のための交流を推進する。国際ワークショップは、領域メンバーに限定せず、国内外の関連諸学会からも広く希望を募り、積極的に機会を提供する。国内外の研究者の招聘にあたっては、国内の関連施設等で講演、ワークショップなども積極的に企画し、学術交流を強く推進する。

**（c）領域推進の計画・方法③**

領域推進の計画・方法について記述すること。特に次の点について具体的かつ明確に記述すること（1頁以内）。

## 5) 公募研究の役割

公募研究は、領域設定期間の1年目に令和6(2024)～令和7(2025)年度分、3年目に令和8(2026)～令和9(2027)年度分の公募が行われる。本研究領域の推進に当たって、公募研究がどのような位置付けにあり、研究領域全体として、どのような研究を期待するのかについて記述すること。本研究領域として何件程度の公募研究課題を必要とし、1研究課題当たり、どの程度の研究経費が必要かを示すこと。公募研究の金額を設定する際は、研究遂行が十分可能な研究経費を計上すること。なお、公募研究の設定に当たっては、次の最低基準のいずれかを上回るように設定すること。

- ・1年目と3年目それぞれの採択目安件数が15件を上回ること
- ・公募研究にかかる経費の総額（令和6(2024)～令和9(2027)年度の合計）が研究領域全体の研究経費（5年総額）の15%を上回ること

採択された研究領域の公募研究への配分総額は、領域計画書に記載の「公募研究」の「研究経費」の範囲内で設定される。

計画研究の代表者は**A00**：計算神経科学・情報理論・機械学習、**B00**：神経科学、神経生理学、精神医学という各専門分で既に国際的に高い評価を得ている研究者であるが、予測と行動の統一理論の開拓と検証を行うためには、計画班とは異なる視点に立つ発想や、オリジナリティーの高い技術や解析手法・理論も相補的に組みあわせる必要があり、専門分野も幅広い分野から募集する。とりわけ理論と実験を繋ぐために必要なデータサイエンスを行う提案や、公募班と計画班の共同研究による相乗効果の大きさも重視したい。例えば、領域全体の実験データ取得からデータ解析・理論構築という流れを意識したうえで、計画研究では扱わない機能や計測スケールの実験データを有する公募研究と計画研究の理論研究者との共同研究、計画研究で絞られた生物学的ターゲットを計測・制御する独自の技術をもつ公募研究との共同研究、公募研究者の独自の理論を計画研究のデータで検証する共同研究などが考えられる。

本領域の発展には多才な人材による公募研究が極めて重要な役割を担うと考えている。特に柔軟で新しい視点を有した若手研究者や女性研究者に積極的な応募を呼び掛ける。そこで公募研究には、総予算の目安年間約3億円のうち約9千万円をあて、2024～2025年度と2026～2027年度にそれぞれ20名程度を採択する。各研究者には年間500万円の範囲で研究経費を配分する。具体的な例は以下に示す。

**A00：情報領域の脳知能研究（理論）**

前半では、普遍理論構築につながる理論研究や、予測や行動に関連する具体的な脳機能を対象にした独自の視点を持った理論の提案、計画班の計測データや既存のデータベースを活用したデータ解析により理論の検証を行うような提案を募集したい。後半では、計画班や公募班の計測データを積極的に用いて理論検証する提案を募集したい。また後半では、人工知能応用も重視し、スパイキングニューラルネットワークによるエネルギー効率の高い計算方法の実装などを含め、次世代人工知能開発につながるようなアイデアを含む提案も募集する。

**B00：物理領域の脳知能研究（実験）**

前半では、動物またはヒトを対象としたオリジナリティーの高い生体情報の計測技術や解析手法を有した提案を募集したい。したがって必ずしも、生物学が専門のバックグラウンドである必要はない。また、幅広い動物種からの提案を募集する。後半は、より生体情報を制御・操作にフォーカスし、それにより因果関係にまで踏み込んで理論予想の検証をするような提案を募集したい。

**（d）過去の採択研究領域等からの発展性（該当する研究領域のみ）**

過去の採択研究領域等からの発展性について記述すること。特に次の点について具体的かつ明確に記述すること（1頁以内）。

- 1) 過去に「学術変革領域研究」「新学術領域研究（研究領域提案型）」又は他の研究費制度において採択された領域型研究を更に発展させる提案については、当該研究費で期待された成果が十分得られたか（中間評価・事後評価（当該研究費の配分機関が行うものに限る）がある場合はその結果についても記述）、更に強い先端的な部分の発展・飛躍的な展開を図る内容となっているか（当該研究領域における研究計画との違いについても記述）
- 2) 応募情報（Web入力項目）において「該当しない」を選択した場合は、その旨を記述

該当しない

**（e）研究経費①**

本研究領域の研究内容及び研究体制等を踏まえ、応募する研究経費の必要性・妥当性について研究期間との関連性を含めて記述すること（1頁以内）。

なお、研究期間内の特定の年度に重点的に研究費を配分する場合、又は年間の応募研究経費の総額（総括班、公募研究を含む。）が、応募上限である3億円を超える場合は、年度ごとに3億円では対応ができない理由、その必要性について記述すること。

1年度目は計画研究のみが研究等を開始することになるため、注意すること。

**総括班：**設備備品として、初年度に2光子レーザー顕微鏡（Thor Lab）を計上する。顕微鏡は理化学研究所和光キャンパスに設置し、計画班と公募班が自由に使える共有設備とする。顕微鏡を含む実験系を沖縄科学技術大学院大学にあるA03班の設備と同様のものとする。都内近郊の研究者がA03班と共通の実験パラダイムを用いることができるようにする。初年度には備品としてデータベース用サーバーも計上する。また、領域会議や班員間の共同研究のための旅費、若手研究者の支援や共同研究のための渡航費を毎年、計上する。2025年と2027年には国際シンポジウムの招聘旅費を計上する。領域の運営を担う補助者の人件費、領域会議・シンポジウム等の会議費、ニュースレター発行やホームページ管理などの広報費などを毎年、計上する。

**計画研究班：**初年度～次年度にかけて、各計画班で必要となる大型備品を購入する。

- ・ A01では、A02, A03と協力し構成する神経回路モデルのシミュレーションや、B00班の提供する神経細胞の活動データに基づいた生成モデルの予測能力の検証を行うために、データ解析用PC、大容量メモリ搭載GPU計算機を導入する。

- ・ A02では、A01, A03と協力し構成する深層生成モデルを計算するためのデータ解析用PC、GPU計算機を導入する。

- ・ A03では、2波長内視顕微鏡システムを導入し、脳皮質予測コーディング仮説の実験的検証のためにマウス大脳皮質局所回路の入力と出力信号、興奮性と抑制性の細胞の活動を区別して同時記録をする。

- ・ B01では、ゼブラフィッシュの神経細胞の活動を計測するための多重電極計測用アンプを導入する。

- ・ B02では、本研究専用のマーモセットの飼育・繁殖のためにマーモセット飼育架台を導入する。

- ・ B03では、マーモセット行動課題用タッチパネルシステムを導入する。

- ・ B04では、定位的頭蓋内脳波のための多チャンネル脳波計を導入する。

このような大型備品は本提案の目的達成には重要度の高いものである。

人件費に関しては、本領域では、様々な階層で包括的なデータ収集、データ解析をしていくことが想定されるため、効果的に研究を推進するために、博士研究員や研究補助員等の人件費や研究協力者への謝金が必要である。旅費は領域会議・国内外の共同研究の打ち合わせ・成果発表の旅費や招聘旅費が必要である。その他、会議費、オープンアクセス料、機器使用料等が必要である。