テーマ別調査・分析

CRDS-FY2024-RR-04

調査報告書

# プレプリントサーバーarXivを利用したAI分野の研究動向俯瞰調査

# Panoramic Survey of Trends in AI Research Field using arXiv

CRDS 研究開発戦略センター
Center for Research and Development Strategy

国立研究開発法人科学技術振興機構

# エグゼクティブサマリー

本報告書は、プレプリントサーバーarXivを用いて、AI分野の最新研究動向を調査した結果をまとめたものである。調査では、arXivのプレプリントと主要国際会議のプロシーディングスを対象とし、技術的類似性に基づく俯瞰マップを使って、直近1～2年の研究傾向を分析した。

調査対象は、2010年以降にarXivに収録された全25カテゴリの文献情報427,846件と、AIに関する30の主要国際会議のデータ200,663件である。

調査結果から、以下の領域が主要なトレンドとして確認された：

- 画像処理用トランスフォーマー
- 自然言語処理のニューラルネットワーク
- 生成AI（深層生成モデル）
- 大規模言語モデル
- ロボットの動作生成における深層学習
- AIチップ（深層ニューラルネットワーク向け専用プロセッサ）

また、以下の新興領域も注目を集めていることが分かった：

- マルチモーダル認識
- 自己教師あり学習
- 音声データの分析・生成
- 知識グラフ
- ロボットの行動・動作認識

この分析により、arXivがAI分野の動向把握やトレンド分析、新しいテーマの発見に有効であることが示された。arXivには、国際会議プロシーディングスにない新興研究トピックや未発表の研究成果が多数含まれており、迅速かつ幅広い研究動向の把握に役立つ。一方、国際会議プロシーディングスにのみ含まれる情報もあるため、両者を組み合わせて活用することで、AI分野の最新研究開発動向をより包括的に理解し、効果的な研究開発戦略の立案に役立てることができる。

# 目次


**1 ｜ 背景と研究動向** …… 1

1.1 AIの進展 …… 1

1.2 情報科学系分野でのプレプリント利用の拡大 …… 2

1.3 AI分野の主な研究動向 …… 4

**2 ｜ 調査の概要** …… 9

2.1 調査目的 …… 9

2.2 文献データ母集団 …… 9

2.3 トピックモデルによる類似度評価 …… 12

2.4 被引用数 …… 13

2.5 時系列予測 …… 13

**3 ｜ ＡＩ論文俯瞰マップ** …… 17

3.1 俯瞰マップ全体概要 …… 17

**4 ｜ arXivと国際会議の比較** …… 29

4.1 全体比較 …… 29

4.2 経年変化の比較 …… 30

4.3 マクロ動向把握 …… 33

4.4 標準化被引用回数 …… 36

**5 ｜ エマージング領域の同定** …… 38

5.1 国際会議への移行状況 …… 38

5.2 引用回数の変化 …… 41

5.3 キーフレーズレベル速報性 …… 49

5.4 時系列予測に基づく新興テーマ探索 …… 56

**参考文献** …… 63

# 1 ｜ 背景と研究動向

## 1.1 AIの進展

人工知能（AI）は、人間の知的活動をコンピュータで再現する技術であり、認識、判断、計画、学習などの機能を持つ。現在のAIは、特定の機能に特化した「特化型AI」と、様々な状況に対応できる「汎用AI」に大別される。現在の第3次AIブームでは特化型AIが主流だが、基盤モデルの登場により、AIの汎用性と多機能性が急速に向上している[1]。

AI技術はビッグデータと密接に関連しながら発展している。ビッグデータの存在がAIの精度を高め、AIの進化がより詳細なデータ分析を可能にするという好循環が生まれている。ビッグデータとは、膨大なデータの収集と分析技術を指す。センサーやIoTデバイスの進歩により、実世界のデータが増加しており、これらのデータは社会問題の解決や現実世界の現象をリアルタイムで把握するのに役立つと期待されている[2]。

第3世代AIは特定の用途で高い性能を示す一方で、大量のデータと計算資源を必要とし、学習範囲外の状況への対応が苦手という課題がある。これを解決するため、知覚・運動系と言語・知識系のAI技術を統合した第4世代AIの研究が進んでいる[3]。図1-1に基盤モデル・生成AIに関わる研究開発課題の全体像を示した。次世代AIモデルの開発では、計算脳科学や認知発達ロボティクスなどの知見も重要な役割を果たしている。

AI技術の社会への浸透に伴い、AIの安全性、公平性、透明性に関する社会原則や倫理指針の策定が進められている。これらを基に、AIソフトウェア工学、人とAIの協働、AI・データ駆動型の問題解決への取り組みが重視されている。

このようなAI技術の社会実装を目指す中で、科学政策や産業政策の方向性を決定するには、最新の研究動向の把握が不可欠である。本調査では、国際会議のプロシーディングスやプレプリントを分析し、AI技術の最新動向を明らかにすることで、研究開発戦略立案に役立つ基礎データの提供を目指している。

**図1-1　基盤モデル・生成AIに関わる課題の全体像[3]**

## 1.2 情報科学系分野でのプレプリント利用の拡大

### 1.2.1 プレプリントとは

原著論文は査読や出版の過程に時間がかかるため、新しい研究分野の動向を迅速に把握するのは難しい。また、異分野融合研究や新概念の研究は、従来の査読基準で不利になる可能性がある[4]。

これらの課題に対処するため、本調査では査読前の「プレプリント」の分析に注目し、新興研究分野や評価が定まっていないトピックスを捉える方法を提案する。特に、AI分野の最新動向を把握するため、プレプリントサーバーのarXiv[5]に着目した。

プレプリントとは、査読付きジャーナルに投稿前の草稿であり、査読も出版もされていない（図1-2）。1990年代のインターネット普及により、プレプリントを公開して広くフィードバックを求める活動が一般化した。arXivは1991年にコーネル大学のPaul Ginsparg氏によって設立され、物理、数学、情報科学分野で広く利用されている先駆的なプレプリントサーバーである。他にも、生物学のbioRXiv[6]、化学のChemRxiv[7]、医学のmedRxiv[8]、心理学のPsyArXiv[9]など、分野別のプレプリントサーバーが運用されている。2022年3月には、JSTによる国産プレプリントサーバーJXiv[10]の運用も始まった。

プレプリントには主に2つの特徴がある。第一に、論文投稿前にサーバーに登録・公開されるため、査読や出版までの期間に関する情報を得やすい。第二に、掲載前のプレプリントと掲載後の論文を比較できる。これらの特徴により、既存の論文データベースでは得られない情報を補完する役割が期待できる。

しかし、arXivの具体的な実態調査結果は少ない。本報告では、arXivに投稿された論文の傾向分析を試み、研究分野の動向把握や科学・産業政策への寄与について考察する。これにより、arXiv利用の意義や影響を明らかにし、科学研究全体の発展に貢献することを目指す。

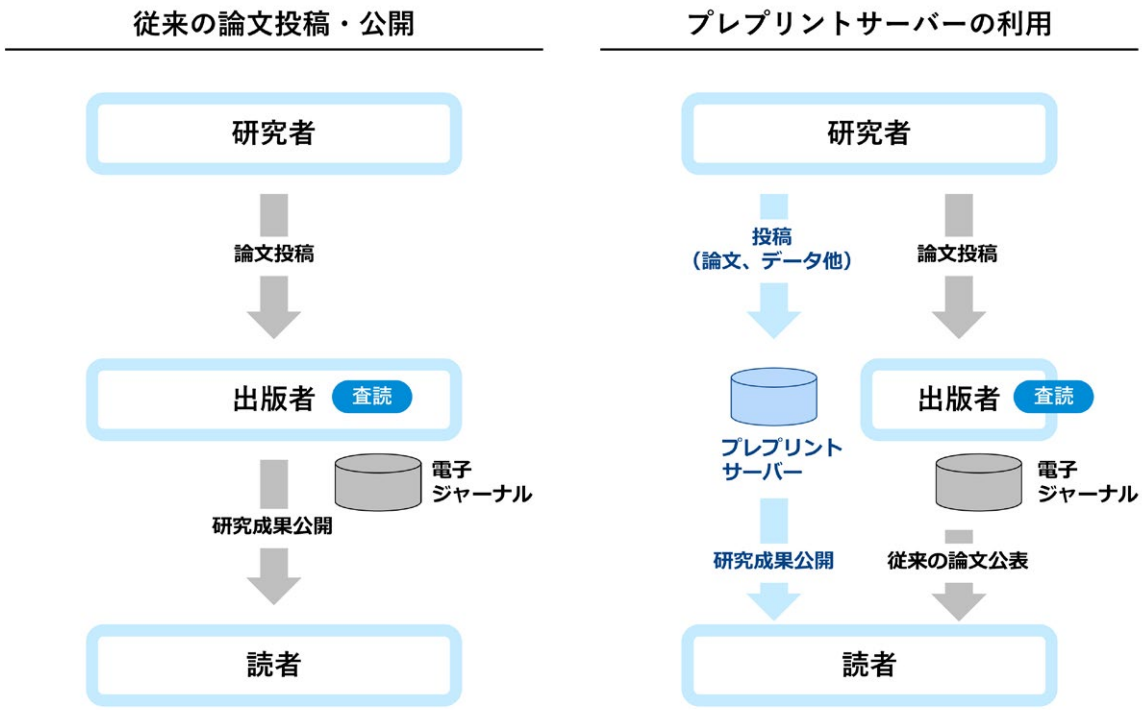


**図1-2　原著論文とプレプリントの学術情報流通の違い（[11]を元にCRDS作成）**

### 1.2.2 情報科学分野でのarXiv利用拡大の背景

情報科学分野では、過去10年でプレプリントの利用が著しく増加している（図1-3）。この利用拡大には複数の要因があると考えられる[12]。

まず、この分野には、NatureやScienceのような絶大な影響力を持つジャーナルがない。そのため、研究の評価は主に国際会議を通じて行われる。ジャーナル論文の有無よりも、国際会議での引用と評価が研究の認知度と影響力を決める主な要因となっている。しかし、国際会議は年に数回しか開催されず、採択率も低いため、発表機会はジャーナル論文に比べて限られているといえる。そのためarXivへの投稿は、国際会議発表の妨げにならず、研究成果を素早く共有できる手段として機能している。また、学術界で競争する必要のない企業の研究は、国際会議での発表の動機が乏しく、プレプリントのみで公開されることも多い。

機械学習研究の活発化がさらにarXivの利用を後押しした。研究者数の増加に対し、年に数回しか開催されない国際会議だけでは発表機会が不足し、研究成果の共有が困難になっている。そこで、新しい着想や結果をまずarXivなどのプレプリントサーバーに公開する方法が定着したと考えられる。

かつてインターネットが普及する前は、「テクニカルレポート」がプレプリントの役割を果たしていた。これは、評価が確立していない研究を公開する独自の文化であり、研究集会や会議での発表がテクニカルレポートとしてまとめられることが多かった。査読付きジャーナルや会議での採択率が低い場合、価値ある研究でも広く認知される機会を失いがちである。しかし、プレプリントサーバーでは、完璧でなくても興味深い論文が評価されることがあり、多くの研究者に利用されている。また、arXivへの先行投稿が会議発表の妨げにならないため、研究者間の迅速な情報共有とフィードバックを促進している。

情報分野は研究開発のサイクルが速く、新しいアイディアを受け入れやすい特徴がある。研究者はarXivに投稿された論文を参照し、公開されたコードを実装して迅速に独自評価を行える。これにより、正式な査読では得られない貴重なフィードバックを得られる。

機械学習分野では、プレプリントが特に重要な役割を果たしている。最新の研究がプレプリントとして公開されると、多くの研究者や技術者が即座にそのコードを試し、研究の価値を実際に検証できる。これは、他分野の伝統的な査読プロセスとは大きく異なる。自己評価としての実装とテストが、公式の査読を必要としない新たな形態として機能している。

加えて、情報科学分野は新しいアイディアに対して比較的寛容で、受け入れやすい環境があることも影響していると考えられる。例えば、実用性の低いアルゴリズムが報告されても、社会に大きな悪影響を与えるわけではなく、研究の試行錯誤として受け入れられる。そのため、多くの研究者がプレプリントを通じて積極的に成果を共有している。

これらの要因が組み合わさり、情報分野でのarXiv利用が広がっている。その結果、研究成果の早期共有と評価が促進され、分野全体の発展につながっている。情報科学の研究開発は他分野と比べて速いサイクルで進行しており、独自の発展を遂げているといえる。次節のarXivのデータ分析でも、これらの傾向が数値的に裏付けられている。

本報告書では、このような分野の特性を考慮し、動向分析にはarXivの利用が適切だと判断してこの分析を行った。ただし、調査対象とする分野によってはarXiv以外のプレプリントサーバーを分析すべき場合もある。また、情報分野に限っても、国際会議プロシーディングスの草稿と思われる論文がarXivに投稿されている割合は約3割にとどまる。そのため、今後もarXivと国際会議議事録を相互補完的に使用して分析する必要がある点に注意したい。

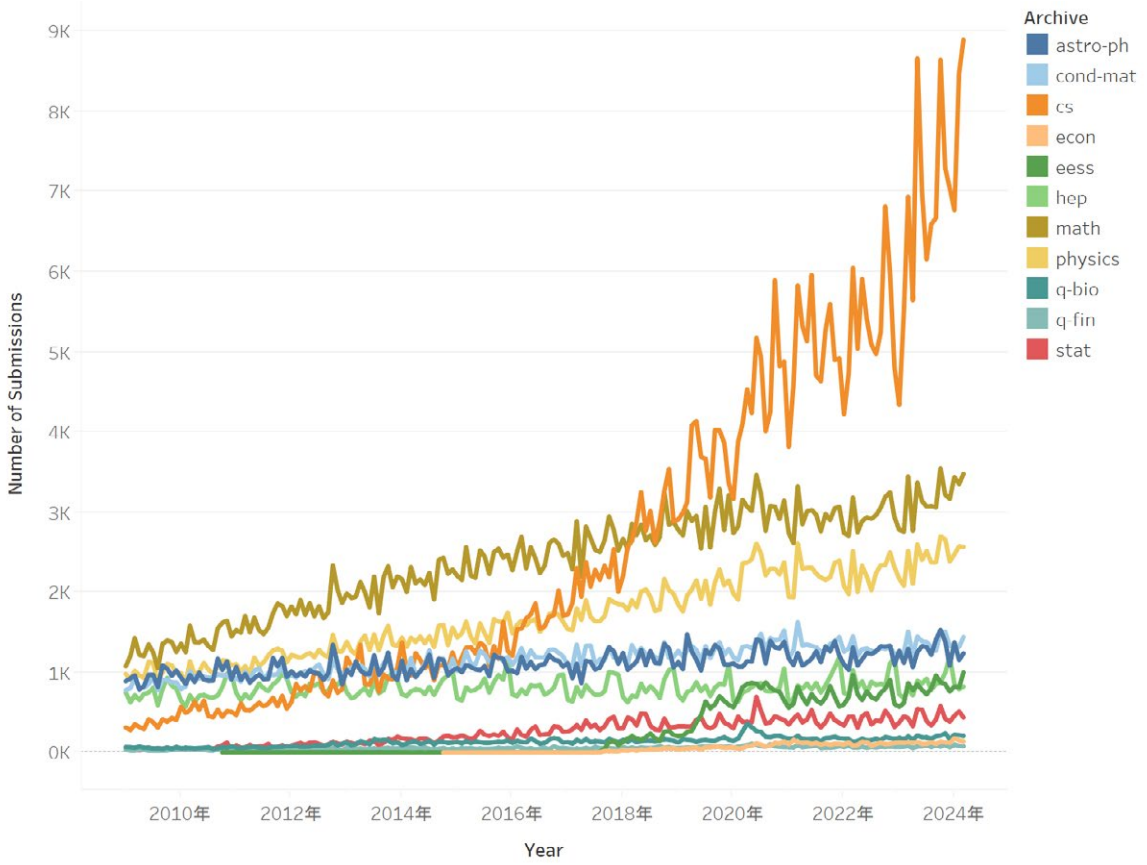


**図1-3　　情報科学分野（cs）の伸び［13］**

## 1.3 AI分野の主な研究動向

### 1.3.1 機械学習・パターン認識

機械学習は、訓練と判定の二つの主要フェーズから成る。訓練フェーズでは、学習データから統計的な規則性を抽出してモデルを構築する。判定フェーズでは、構築したモデルを使って新しいデータの予測や分類を行う。訓練データにラベルがある場合を教師あり学習、ない場合を教師なし学習と呼ぶ。また、強化学習は特定環境下での行動を通じて最適な戦略を学ぶ方法である。

過去には、人間の脳の働きを模した研究としてニューラルネットワークとその学習アルゴリズムが注目された。1958年に提案されたパーセプトロンは単純な線形関数の学習に有効だったが、排他的論理和を学べないという欠点があった。これを克服するため、1986年に誤差逆伝播法が提案され、ニューラルネットワーク研究が再興した。

1992年には、機械学習研究に大きな影響を与えたカーネル学習器SVM（Support Vector Machine）が登場した。2006年以降、Geoffrey Hintonらによって提案された深層学習は、多層ニューラルネットワークの効果的な学習法として注目を集めた。特に2012年、深層学習が画像認識コンペティションで圧倒的な成果を上げ、その有効性が広く認識されるようになった。

その後もさまざまな改良が加えられた。ResNetでは残差学習が導入され、深層ネットワークの学習効率が向上した。RNNやLSTMは時系列データや自然言語処理で有効性を示し、最近ではアテンション機構を中心としたトランスフォーマーが新たな標準として確立され、幅広く応用されている。

### 1.3.2 画像系のトランスフォーマーと自己教師あり学習

近年、画像認識を含むパターン認識分野の深層学習アーキテクチャは、CNN型からトランスフォーマー型へ移行している。その代表例が2020年に発表されたビジョントランスフォーマー（Vision Transformer：ViT）である[14]。ViTは画像を重ならないパッチに分割し、各パッチをベクトル化して位置情報を加えることで、自然言語処理で使用されるトランスフォーマーと同様の処理を適用する。このアプローチにより、ViTは従来のCNNモデルより高性能で計算コストも大幅に削減された。これを受け、ViTの様々な派生モデルが開発されている[15]。

深層学習モデルの高精度化には大量の教師データが必要だが、その準備は容易ではない。この問題に対し、半教師あり学習、能動学習、転移学習、ドメイン適応、データ合成やデータ拡張などの手法が試みられている。特に自己教師あり学習は大きな効果を示している。自然言語処理分野では、BERT以降、テキストの一部にマスクをかけた穴埋めタスクを設定し、ラベルなしでも教師あり学習を可能にする手法が用いられ、モデルの大規模化を促進している。

画像処理分野の自己教師あり学習は、主に対照学習[16]とマスクドオートエンコーダー（MAE）[17]の二つの方法で行われる。対照学習では、画像データに各種変換を施し、同一画像の変換画像間で特徴量を最大化し、異なる画像の変換画像間で特徴量を最小化する訓練を行う。これにより、教師あり学習に匹敵する認識精度が得られると報告されている。一方、MAEは画像の一部をマスクし、その部分を推測する学習で同様の効果を実現している。

これらの技術進化により、深層学習はさらに多くの応用分野での利用が期待され、新たな派生技術の開発が進んでいる。各技術の効果的な組み合わせによって、画像認識や自然言語処理分野での革新が今後も続くと予想される。

### 1.3.3 深層生成モデルと画像生成AI

機械学習のクラス分類手法は、識別モデルと生成モデルに大別される。識別モデルはデータの所属クラスを特定し、生成モデルはデータの生成過程をモデル化する。識別モデルの代表例は畳み込みニューラルネットワーク（CNN）である。一方、生成モデルでは深層生成モデルの発展が著しい[18]。主な深層生成モデルには、GAN（敵対的生成ネットワーク）[19]、VAE（変分自己符号化器）[20]、フローベースモデル[21]、拡散モデル[22]がある。

GANは2014年にIan J. Goodfellowが開発し、生成器と識別器の二つのネットワークで高精細な画像生成を実現したが、学習の安定性に課題がある。VAEは入力データを潜在変数に変換し再生成する方式で、技術進歩によりGANとの性能差は縮まっている。

フローベースモデルは正規化フロー技術で確率分布をモデル化し、新しいサンプルを生成する。拡散モデルはノイズ付加プロセスを逆に辿ってデータを生成する。これらは計算コストが高いが、安定した学習と高品質な生成結果が得られ、広く研究されている。深層生成モデルは架空の人物の顔や特定画風の絵画生成に応用されている。2022年にはテキストから画像を生成する技術が注目を集め、DALL-E[23, 24]、Imagen[25]、Parti[26]、Muse[27]などのモデルが簡単なテキストから高品質な画像を生成し、多くの利用者を魅了している。

これらの技術は機械学習の可能性を広げ、新しい形のアートやデザイン創出に貢献している。大規模データ学習により、画像生成の精度と多様性が向上し、モデルのパラメータ数も数十億から数百億個に増加している。

生成モデルの応用は芸術的創作だけでなく商業利用にも広がっている。広告業界では特定キャンペーン用の画像やビデオの自動生成が可能になり、教育研究分野では視覚教材やデータセット生成に活用されている。

将来的には、一般の人々も高品質な生成コンテンツを容易に作成できるようになると期待される。そのためには、技術的知識がなくても操作できる直感的なインターフェースの開発が重要である。技術の発展により、様々な分野での創造的活動がさらに豊かになることが期待される。



### 1.3.4 ニューラルネット自然言語処理

深層学習の発展により、画像認識では主に畳み込みニューラルネットワーク（CNN）が、自然言語処理では初期に回帰型ニューラルネットワーク（RNN）やLSTM（Long Short-Term Memory）ネットワークが採用された[28, 29]。これらの技術に注目点を動的に決定するアテンション機構が導入され、特に機械翻訳での系列変換の質が向上した。アテンション機構の成功を受け、自然言語処理だけでなく画像処理にも応用が広がった[30,31]。

2017年、Googleがアテンション機構のみで構成されるトランスフォーマーを発表した[32]。このモデルはRNNやCNNを使わずに計算効率と学習のしやすさを大幅に改善し、自然言語処理分野で顕著な成果を上げ、機械翻訳や意味理解タスクで高性能を示した。トランスフォーマーの登場により、自然言語処理のベンチマークスコアが次々と更新された。ニューラルネット自然言語処理には大量の訓練データが必要だが、全タスクで大量データを用意するのは困難である。そこで、大量のラベルなしデータでモデルを事前学習し、少量のラベル付きデータで追加学習を行う手法が採用された。2018年にGoogleが発表したBERT[33]は、この事前学習アプローチを用い、MLM（Masked Language Model）を通じて言語理解の新標準を確立した。

BERTの成功後、多くの改良モデルが提案され、自然言語処理ベンチマークで最高スコアを更新している。RoBERTa[34]、ALBERT[35]、ERNIE[36]、UniLM[37]などは特定の言語タスクをより効率的に処理するよう設計されている。また、ViLBERT[38]やVL-BERT[39]のようなマルチモーダルモデルも登場し、テキストに加え画像や音声データも統合的に扱えるようになった。

近年は大規模言語モデルの開発が進み、OpenAIのGPT-3[40]やMicrosoftとNVIDIAのMT-NLG[41]などのモデルが大量データで事前学習され、自然言語処理の可能性を大きく広げている。これらのモデルは事前学習だけでなく、ゼロショット学習を通じて多様なタスクに適応する能力を持ち、自然言語処理技術の新たな地平を開いている。

### 1.3.5 大規模言語モデル・基盤モデル

2018年にGoogleが発表したBERT以降、トランスフォーマー型の大規模言語モデル（Large LanguageModel：LLM）は主に自己教師あり学習で構築されるようになった。2020年にOpenAIが発表したGPT-3は、BERTの約3000倍のデータ規模と500倍以上のパラメーター数を持つ[40]。このモデルは少数の例示（Few Shot Learning）で様々な自然言語処理タスクに対応し、ブログや小説の生成、プログラムコードやHTMLコードの作成、多様な質問への回答など、幅広く活用された。

2022年11月、GPT-3の改良版であるGPT-3.5が発表され、人間のフィードバックを基にした強化学習（Reinforcement Learning from Human Feedback）を加えたChatGPTがWeb公開された[42]。このモデルはチャット形式で多様な自然言語タスクに応答し、大きな注目を集めた。

さらに、テキストだけでなく画像や動画も扱える「基盤モデル」（Foundation Model）[43]と呼ばれる新カテゴリのモデルが登場した。例えば、Google DeepMindのFlamingoやGatoは、テキスト、画像、動画などの様々なデータ形式を扱い、Few Shot Learningで新しいタスクに適応できる。

テキストから画像を生成する技術も進化し、OpenAIのDALL-E[23, 24]やGoogleのImagen[25]が主要な進展として挙げられる。これらの技術は、大量で多様なデータを用いた訓練により、汎用性や多様なタスクへの適応能力を高めている。

### 1.3.6 深層ニューラルネット（DNN）アクセラレーター

2011年、深層ニューラルネットワーク（DNN）が画像分類精度で従来手法を大きく上回ることが報告され、注目を集めた。この技術は1980年代に提唱されたバックプロパゲーション（BP）を基盤とし、大規模学習データ、高性能計算機、BP改善手法の組み合わせにより、画像・音声認識、自動翻訳、自動運転など多様な分野で応用が広がっている。

学習と推論処理の高速化・低電力化を目指し、Google社のTPUのようなDNN処理エンジンが提案され、新しい情報処理アーキテクチャーとして注目されている[44–46]。また、GPUコンピューティングが学習処理の中心的アーキテクチャーとなり、DSPベースのアプローチが画像処理用組み込み機器向けアクセラレーターとして提案されている。さらに、データフローマシン[47]、FPGAコンピューティング[48]、リコンフィギュラブルコンピューティング[49]など、様々な新技術が登場し、競争が激化している。

国内では、東京工業大学[50]と北海道大学[51]がFPGAとリコンフィギュラブルコンピューティングの研究を進めている。ルネサスエレクトロニクス社[52]やプリファードネットワークス社も注目技術を発表しており、特にプリファードネットワークス社のDNNディープラーニングアクセラレーターチップは、スーパーコンピューターMN-3に搭載され、Green500で1位を獲得し大きな注目を集めた[53]。

DNNのスケーリング則に関する研究も進展し、より大きなDNNモデルが未学習データに対して高い汎化性能を示すことが明らかになった。これは従来の機械学習理論であるオッカムの剃刀に反する発見である。この新理論に基づく「宝くじ理論」では、DNN学習を通じて良い部分ネットワークを選び出せるとされる。東京工業大学はこれに基づいた推論チップを発表し、重みの乱数生成により推論時のメモリーアクセスを削減し、電力効率を向上させる技術を開発した。

### 1.3.7 世界モデルと深層予測学習

人間は限られた知覚情報から脳内で外界のモデルを作り、それを用いて意思決定や行動を行う。このモデルは「内部モデル」「力学モデル」とも呼ばれ、AI分野では「世界モデル」[54]として知られている。このモデルは無意識的な反射行動にも影響を与え、例えばバットでボールを打つ際、視覚情報からボールの位置を予測し筋肉を動かしている。

世界モデル構築の研究例として、Google DeepMindのGQN（Generative Query Network）[55]がある。 GQNは複数の視点からの画像を基に内部の世界モデルを作り、未知の視点からの画像を予測できる。この技術はVAEベースの深層生成モデルを使用し、深層学習研究の新たな注目トピックとなっている。

知覚情報に基づくモデルを用いたシミュレーションや予測は、知覚系と運動系を即応的ループとして結びつける上で重要である。この考え方は、産業技術総合研究所と早稲田大学が開発したタオル畳みロボットに応用されている[56]。このロボットは人間のタオル畳み動作を模倣しながら、動作を予測し、実際の動作と予測との差異を学習する。

認知発達ロボティクスでは、予測誤差最小化原理に基づくモデルの更新や環境への働きかけが、人間の認知発達に深く関わるとされる。この分野は人間の知能の認知発達メカニズムを構成論的に解明しようとするもので、深層学習の進展と重なる部分が多く、今後の発展が期待されている。

### 1.3.8 深層学習によるロボットの動作生成

産業用ロボットの動作生成技術は、近年大きな変革を遂げている。従来の演繹的アプローチから、深層学習を活用した帰納的アプローチへの移行が進んでいる[57]。

演繹的アプローチは、環境や対象のモデルを先に作り、それに基づいて最適な動きを求める。しかし、こ

の方法は環境の変化に弱く、使える範囲が限られていた。一方、帰納的アプローチは、センサーデータを基に統計的な修正を加えながらモデルを作り、動きを生み出す。特に注目されているのが、End-to-End学習だ。これは、環境の認識から動作の生成まで一貫して学習し、人間の介入を最小限に抑える方法である。

深層強化学習も重要な技術である。これは、行動の結果得られる報酬を基に、最適な行動の選び方を学習する。Google DeepMindのAlphaGoは、この技術を用いて世界トップクラスのプロ棋士に勝利した有名な例である。

ロボットの動作生成では、実機での多数の試行が難しいため、コンピューターシミュレーションを活用するSim-to-Realアプローチが採用されている[58]。ただし、シミュレーションと実機の差を埋めるための調整が必要となる。最近では、自然言語処理や画像処理で成果を上げているトランスフォーマー型の大規模モデルも、ロボットの動作生成や制御に応用され始めている。PaLM-SayCanやRT-1など、言葉による指示に基づく動作生成や、実機での学習を進める研究が行われており、今後の発展が期待される分野である。

これらの新しい技術により、ロボットはより複雑で多様な環境に適応できるようになると考えられる。その結果、人間の作業者と協力してタスクを効率的に遂行し、産業全体の生産性向上に貢献することが可能になるだろう。

ロボットは単なるプログラム実行機から、状況を理解し最適な判断を下す知的な機械へと進化しつつある。深層学習や強化学習の進展により、ロボットの学習プロセスは人間の学習方法に近づき、未知の状況や複雑な問題への対応能力が向上している。ただし、実世界での動作精度の向上、長期的な学習と記憶の安定性、異なるタスク間での知識の転移など、まだ多くの課題が残されている。これらの課題を解決することで、ロボットはより高度な技術支援を人間に提供できるようになるだろう。

### 1.3.9 AIロボット駆動科学

AIは科学的発見の分野に新たな可能性をもたらしている。囲碁界でAlphaGoがトップ棋士に勝利したように、AIは人間の想像を超える手法を生み出し、科学の世界でも同様の影響を及ぼす可能性がある。

AIの強みは、人間の認知能力を超えた仮説の生成・探索と、その評価・検証の効率化にある。深層学習を用いることで、複雑な現象から規則性を見出し、異分野の知識を組み合わせた推論が可能になる。さらに、ロボット技術の導入により、科学的発見プロセスの自動化や効率化が進んでいる。

国際的には、2050年までにノーベル賞級の科学的発見を目指すAIシステムの開発が進められている。英国のアラン・チューリング研究所による「The Turing AI Scientist Grand Challenge」や、「Nobel TuringChallenge」などのプロジェクトが立ち上がっており、日本でも関連研究が推進されている。

実際に、材料科学や生命科学、創薬の分野ではAIを活用した科学的発見が行われている。物理学でも深層学習モデルが現象理解に貢献し、新法則の発見や新物質の開拓が期待されている。ただし、AIが人間には理解困難な結果を導き出す場合、その科学的価値の評価方法について議論が必要である。

AIとシミュレーションの融合も進んでいる。物理シミュレーションの結果を学習データとして利用することで、数値シミュレーションの高速化や調整の容易化が実現している。また、複雑系や社会科学の課題に対しても、機械学習とシミュレーションの組み合わせが有効とされている。

これらの取り組みは科学的探求の新たな地平を開くものであり、処理フローの効率化や絞り込みアルゴリズムの更なる発展が求められている。AIと科学の融合は、今後の科学技術の発展に大きな影響を与えると考えられ、その進展が注目されている。

# 2 ｜ 調査の概要

## 2.1 調査目的

本調査の目的は、AI分野における最新の研究開発動向を把握し、直近1～2年の研究潮流を明確に分析することである。この目的を達成するため、技術的類似性に基づいた俯瞰マップを用いて文献情報を分析した。

調査では、主要国際会議のプロシーディングスに掲載された発表済み論文の書誌情報に加え、プレプリントサーバーarXivに投稿された未発表論文の書誌情報も収集した。これらの情報を同一の俯瞰マップ上に統合し、可視化および分析を行った。

分析手法として、俯瞰マップ上における研究テーマの空間的な関係性と時間的な変化を詳細に解析した。これにより、AI分野における研究開発の最新動向を包括的に把握し、分析することが可能となった。

この手法を用いることで、発表済みの論文だけでなく、最新の未発表研究も含めた幅広い視点からAI分野の研究潮流を捉えることができる。また、時系列での変化を追跡することで、研究テーマの発展や新たな領域の出現などを迅速に識別することが可能となる。本調査の結果は、AI分野の最新動向を理解する上で有用な情報を提供し、今後の研究開発の方向性を検討する際の重要な指針となることが期待される。



## 2.2 文献データ母集団

arXivに収録された文献について、表2-1に挙げた全25個のカテゴリについて、2010年以降の全テキストを収集した。カテゴリ付与に重複があり重複を除いて429,669件、さらにそこから記述の短いものを除外して、最終的に427,846件を分析対象とした。

国際会議プロシーディングスについては、AIに関係する30の主要な国際会議で発表された文献について2010年以降のデータをエルゼビア社のScopusより取得した（表2-2）。会議の選定には、Research.comのImpact Scoreを参考とした。合計206,730件収集され、類似した文献や、記述が短くて詳細な分析に適さない文献を除外して最終的に200,663件を分析対象とした。

**表2-1　分析対象としたarXivのカテゴリ**

| No. | カテゴリ | 内容 | 件数 |
|---|---|---|---|
| 1 | cs.AI | Artificial Intelligence | 72,512 |
| 2 | cs.AR | Hardware Architecture | 4,022 |
| 3 | cs.CE | Computational Engineering, Finance, and Science | 5,786 |
| 4 | cs.CL | Computation and Language | 50,136 |
| 5 | cs.CR | Cryptography and Security | 28,324 |
| 6 | cs.CV | Computer Vision and Pattern Recognition | 107,288 |
| 7 | cs.CY | Computers and Society | 16,020 |

| | | | |
|---|---|---|---|
| 8 | cs.DB | Databases | 7,399 |
| 9 | cs.DC | Distributed, Parallel, and Cluster Computing | 18,758 |
| 10 | cs.ET | Emerging Technologies | 3,686 |
| 11 | cs.GR | Graphics | 5,100 |
| 12 | cs.HC | Human-Computer Interaction | 14,242 |
| 13 | cs.IR | Information Retrieval | 14,539 |
| 14 | cs.LG | Machine Learning | 156,193 |
| 15 | cs.MA | Multiagent Systems | 6,588 |
| 16 | cs.MM | Multimedia | 5,536 |
| 17 | cs.NE | Neural and Evolutionary Computing | 12,741 |
| 18 | cs.PF | Performance | 3,366 |
| 19 | cs.PL | Programming Languages | 6,592 |
| 20 | cs.RO | Robotics | 27,448 |
| 21 | cs.SD | Sound | 12,396 |
| 22 | cs.SE | Software Engineering | 13,976 |
| 23 | cs.SI | Social and Information Networks | 18,526 |
| 24 | stat.AP | Applications | 16,467 |
| 25 | stat.ML | Machine Learning | 61,869 |
| | | のべ合計 | 689,510 |
| | | 正味合計 | 429,669 |



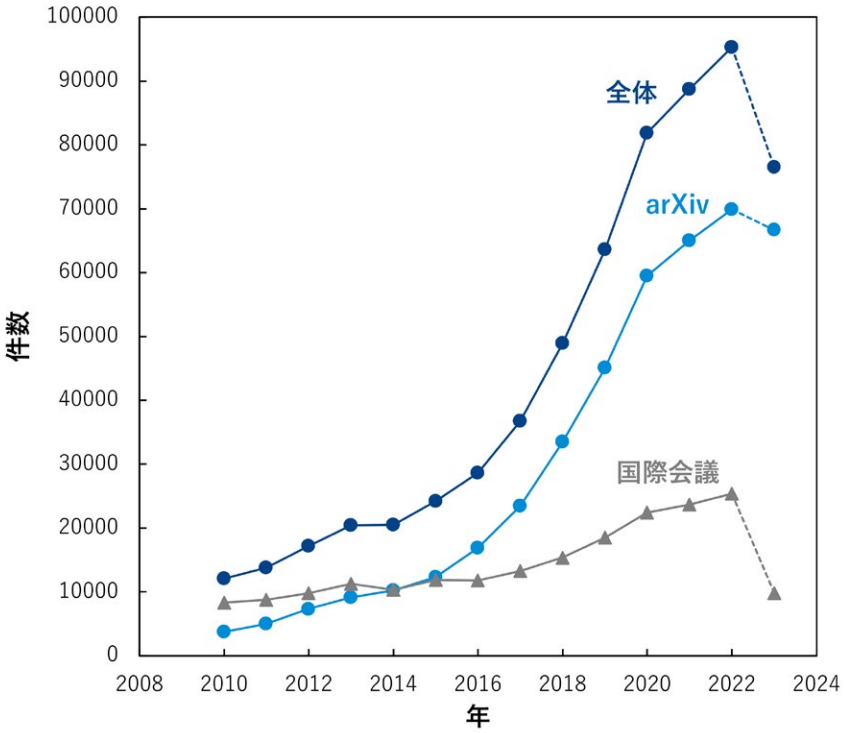


**図2-1　　分析対象の論文数の経年変化**

**表 2-2　　分析対象とした国際会議**

| No. | 略称 | 国際会議名 | 件数 |
|---|---|---|---|
| 1 | AAAI | AAAI Conference on Artificial Intelligence | 15,604 |
| 2 | ACL | Meeting of the Association for Computational Linguistics | 7,768 |
| 3 | CCS | Computer and Communications Security | 2,965 |
| 4 | CVPR | IEEE/CVF Conference on Computer Vision and Pattern Recognition | 11,430 |
| 5 | CVPRW | Conference on Computer Vision and Pattern Recognition Workshops | 4,542 |
| 6 | ECCV | European Conference on Computer Vision | 6,926 |
| 7 | EMNLP | Empirical Methods in Natural Language Processing | 6,188 |
| 8 | ICASSP | International Conference on Acoustics, Speech, and Signal Processing | 20,288 |
| 9 | ICCV | International Conference on Computer Vision | 6,927 |
| 10 | ICCVW | International Conference on Computer Vision Workshops | 1,957 |
| 11 | ICLR | International Conference on Learning Representations | 4,392 |
| 12 | ICML | International Conference on Machine Learning | 7,136 |
| 13 | ICRA | International Conference on Robotics and Automation | 12,661 |
| 14 | ICSE | International Conference on Software Engineering | 5,170 |
| 15 | IHC | Human Factors in Computing Systems | 14,417 |
| 16 | IJCAI | International Joint Conference on Artificial Intelligence | 8,362 |
| 17 | Ijcnlp | International Joint Conference on Natural Language Processing | 2,341 |
| 18 | INTERSPEECH | Conference of the International Speech Communication Association | 9,447 |
| 19 | IROS | International Conference on Intelligent Robots and System | 12,262 |
| 20 | MICCAI | Medical Image Computing and Computer-Assisted Intervention | 4,387 |
| 21 | MM | ACM Multimedia | 4,356 |
| 22 | NAACL-HLT | North American Chapter of the Association for Computational Linguistics - Human Language Technologies | 4,340 |
| 23 | NIPS | Neural Information Processing Systems | 13,026 |
| 24 | SIGIR | Research and Development in Information Retrieval | 3,843 |
| 25 | SIGKDD | Knowledge Discovery and Data Mining | 4,542 |
| 26 | SIGMOD | International Conference on Management of Data | 2,282 |
| 27 | SP | IEEE Symposium on Security and Privacy | 1,220 |
| 28 | USENIX | Security USENIX Security Symposium | 1,364 |
| 29 | WADS | Workshop on Algorithms and Data Structures | 346 |
| 30 | WWW | World Wide Web | 6,241 |
| | | 合計 | 206,730 |

## 2.3 トピックモデルによる類似度評価

本調査では、論文・特許の分類にトピックモデル（LDA: Latent Dirichlet Allocation）を用いた（図2-2）。トピックモデルは、文書中の単語の出現確率を推定し、類似した単語を含む文章を似たものとして評価する自然言語処理の手法である。

トピックモデルでは、各文章が複数のトピックで構成され、各トピックの単語分布の組み合わせで単語が生成されると仮定する。この手法は、トピックと文書の単語が完全に一致しなくてもよいという柔軟性があり、文章の特徴を複数のトピック構成として捉えることで、文章間の類似度を定量的に評価できる。文書の可視化は以下の3ステップで行った：

(1) トピックモデル（LDA）を用いて、各文献の内容を自動抽出されたトピックの組み合わせで表現する。
(2) 文献間でのトピックの共有関係から、トピック空間における高次元類似度を求める。
(3) 多様体学習により2次元に次元削減を行い、図示化する。

二次元座標にマッピングした図では、原点が母集団で最も頻出する概念を示す。原点からの距離は、その中心的概念からの差異の度合いを表す。文書間の距離も同様に差異を示し、近接する文書は類似した概念を含むと解釈できる。ただし、クラスターや文書の絶対的な位置ではなく、相対的な位置関係が重要である。

各クラスターへのラベル付けは、文書ごとに推定されたトピックに基づいてグループ化し、トピック内のキーワードの組み合わせなどを元に手作業で行った。

LDAの特徴として、キーワードの組み合わせによるトピックをさらに組み合わせて文献の内容を表現するため、単一キーワードのみによる類似判定が生じにくい。例えば、「cell」という単語だけでなく、「stemcell（幹細胞）」や「solar cell（太陽電池）」のように、より具体的な概念での類似判定が可能となる。

一般的に、筆頭トピックが同じ文献は近くに集まるが、2番目以降のトピックの違いにより、その集積領域内での相対位置が変わる。ノイズ除去の際には、筆頭トピックが関心外でも、2番目以降のトピックに関心のある技術が含まれている場合は採用した。この手法により、文献の内容をより正確に把握し、関連性の高い研究や技術を効率的に特定することが可能となる。

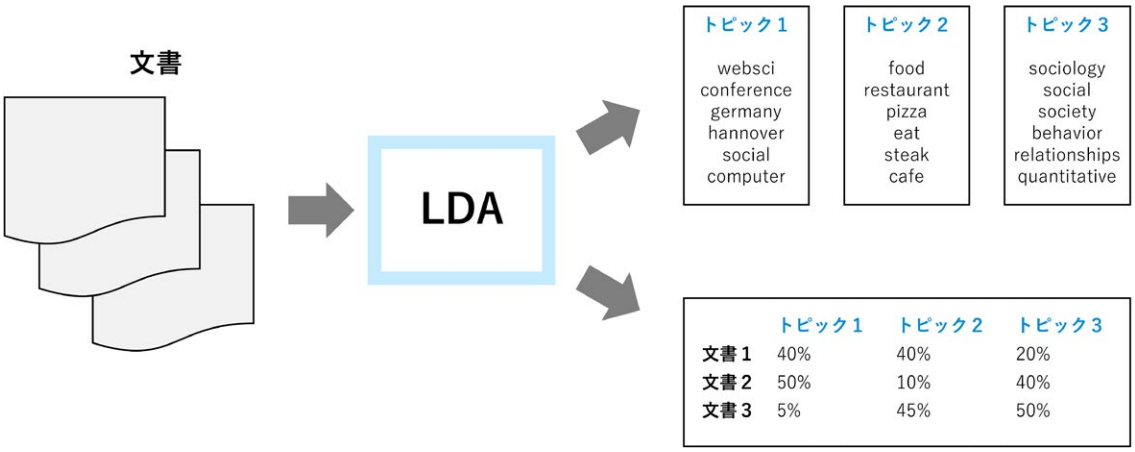


**図2-2　　トピックモデル（LDA）**

## 2.4 被引用数

本調査では、arXivとScopusデータベースに収録されている国際会議議事録の被引用情報を分析した。arXivには引用情報がないため、Semantic Scholarから取得した。図2-3初回取得した引用情報に基づく平均被引用数の経年変化を示した。データ取得は2023年10月（初回）と2024年1月（二回目）に実施し、約80日の間隔があった。この期間における被引用回数の変動を調査することで、研究発表後の影響の推移を明らかにしようと試みた。具体的な取得日程は、arXivは2023年10月20日～25日（初回）、2024年1月4日～5日（二回目）、国際会議（Scopus）は2023年10月18日（初回）、2024年1月12日（二回目）である。

被引用数の変化の分布を図5-4に示した。収集したデータには被引用回数が増加するケースと減少するケースの両方が含まれるが、変化なしの割合が50%以上を占め、大半のデータで顕著な変動は見られなかった。

一方で、増加および減少の両方において数万件の変動が記録される事例も確認された。これらは初回の集計時の誤りが修正された結果、または二回目での集計ミスによるものと推測される。適切に処理されたデータと考えられるのは数件の増加に留まるケースだが、数件の減少などの誤り修正とを系統的に区別することが困難であった。

このため、本調査では収集したデータをそのままの形で分析を行った。このapproachは、データの完全性や正確性に一定の制限があることを認識しつつ、全体的な傾向を把握する上では有効であると判断した。今後の調査では、データの信頼性をさらに高めるための方法を検討し、より精緻な分析を行うことが課題となる。また、被引用数の変動パターンと研究内容や分野との関連性についても、さらなる調査が必要である。

## 2.5 時系列予測

本調査では、arXiv登録文献と国際会議議事録における新技術キーワードの出現タイミングを分析した。2014年以降、arXivに掲載された論文に新技術のキーワードが先行して現れる傾向が確認された。この事実に基づき、arXivを利用した新興技術テーマの早期検知の可能性を探った。

arXivは日単位でデータが収録されているため、Scopusと比較してより精密に技術トレンドを捉えられると期待された。しかし、日単位では統計的精度が不十分であったため、本調査では月単位でのキーワード出現の推移を分析した。例として、arXiv文献中に出現した「large language model」の数を図2-5に示した。

技術出現の予測には、MetaのProphetを使用した。Prophetは、予測値に加えて、トレンドや周期性パターンも出力する。予測値自体は周期性予測により不安定だが、トレンドの動きは安定しており、キーフレーズ出現の予測に有効と判断した。また、Prophetは学習データが少ない場合、警告を出し自動的にMCMC法で不足分を生成して予測を行うため、時系列分析に不慣れな利用者でも適切な予測が可能という利点がある。

予測値の算出方法は以下の通りである（図2-6）：

- 2018年1月の予測値：2010年1月～2017年12月のデータを用いて2018年1月から12ヶ月後までを予測し、その平均値を使用。
- 2018年2月の予測値：2010年1月～2018年1月のデータを用いて2018年2月から12ヶ月後までを予測し、その平均値を使用。
- 以降も同様に、予測対象月の前月までのデータを用いて12ヶ月分の予測を行い、その平均値を使用。

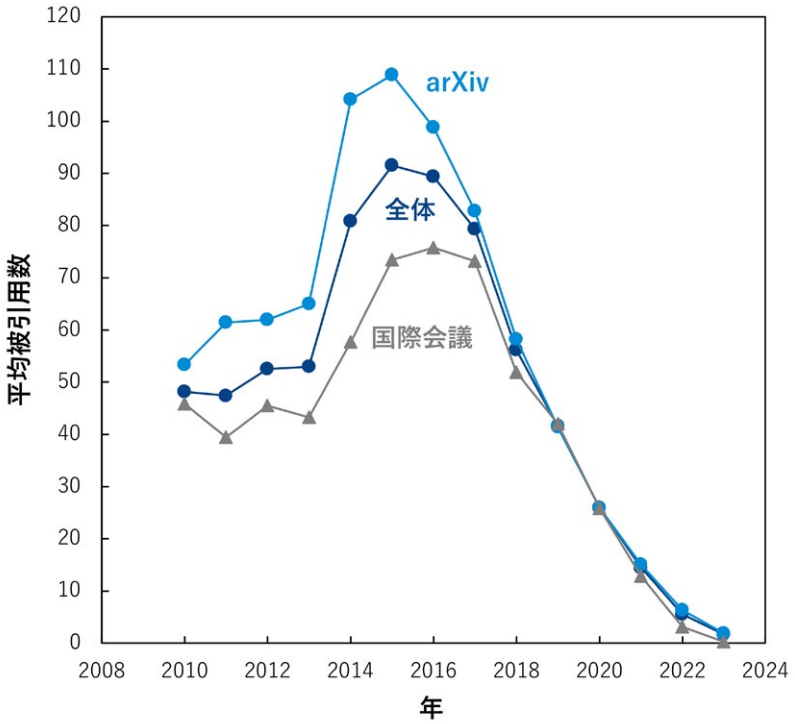


**図2-3　　分析対象の論文の平均被引用数の経年変化**

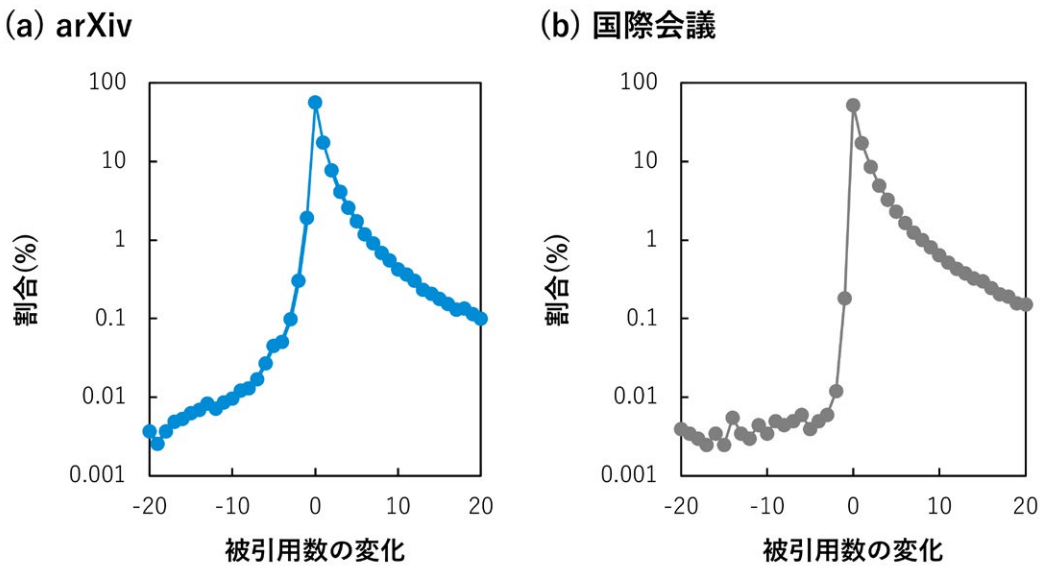


**図2-4　　初回と二回目での被引用数の変化の分布**

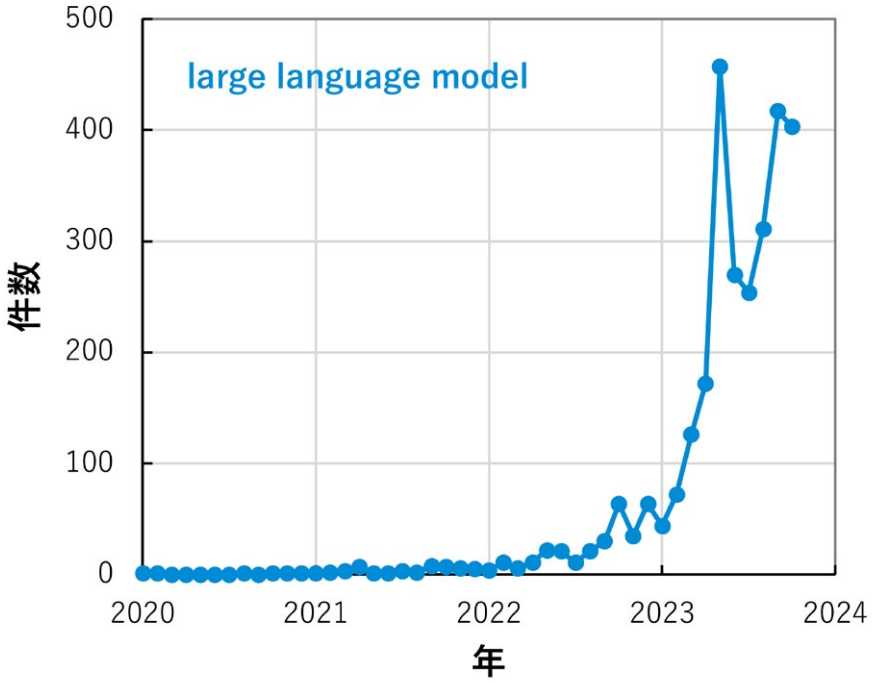


**図2-5　arXiv文献中の「large language model」の出現推移（月単位）**



arXivにおける実測値との比較（バックテスト）を5つのキーフレーズに対して実施した結果、トレンド値が「1」を超えた時点で新興技術の出現と判断することが適切であることが示された（図2-7）。

この手法により、新興技術の早期検知が可能となり、研究開発戦略の立案や技術動向の把握に有用なツールとなることが期待される。今後は、より多様なキーフレーズでの検証や、他のデータソースとの比較分析を通じて、予測精度の向上と適用範囲の拡大を図ることが課題となる。

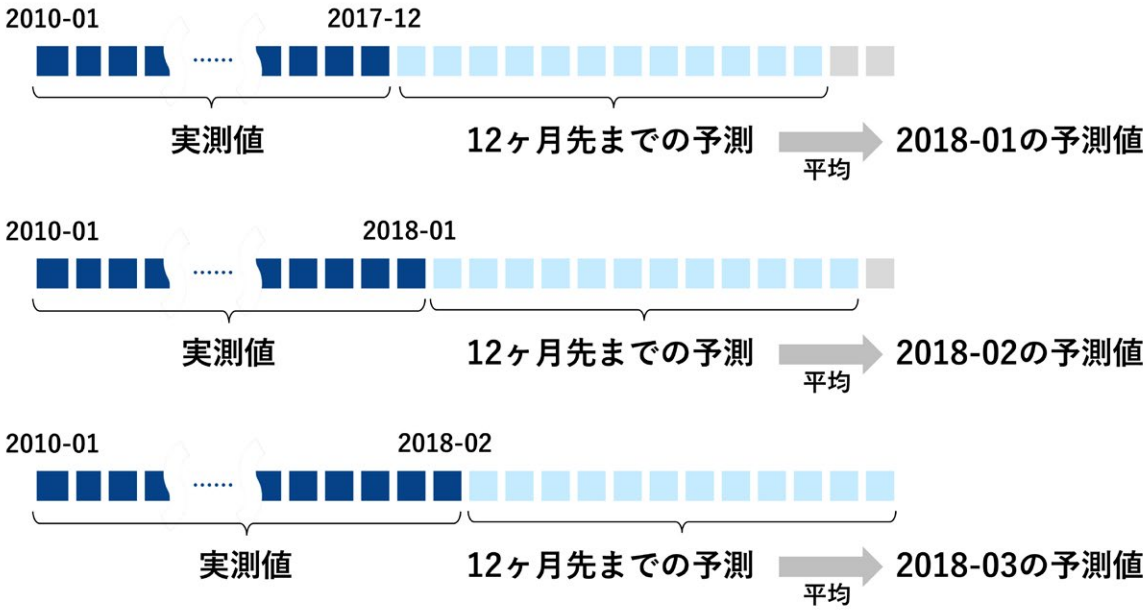


**図2-6　予測値の算出方法**

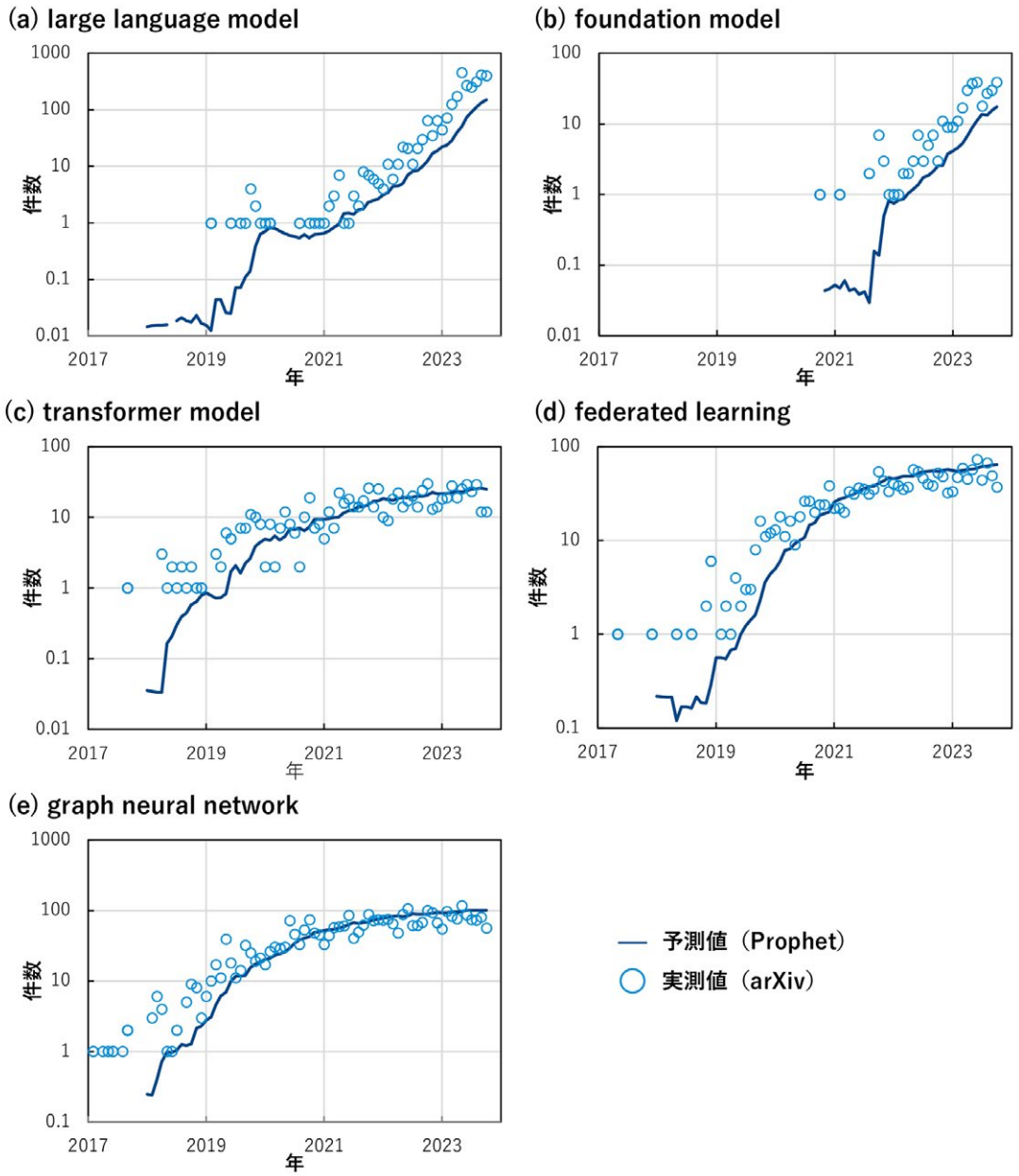


**図2-7　　時系列予測のバックテスト**

# 3 ｜ AI論文俯瞰マップ

## 3.1 俯瞰マップ全体概要

AI関連研究に関するarXiv収録データ427,846件とScopus収録の国際会議議事録データ200,663件、合計628,509件を対象に、トピックモデルと多様体学習による俯瞰解析を行った（図3-1）。分析対象は各文献のタイトルと要約である。

俯瞰図上では、トピックモデルの主要トピックが共通する領域を赤の破線楕円で示し、これを「技術領域」と呼ぶ。さらに、類似した赤楕円群を目視で判断し、青の実線楕円で囲んだ。この群を「技術区分」と呼ぶ。同定した技術区分と技術領域は表3-1のとおりである。

各領域の名称は、以下のデータを順に参考にして決定した：

(1) トピックのキーワード
(2) RAKE（Rapid Automatic Keyword Extraction）によるキーフレーズ
(3) タイトル、要約のテキスト

図3-2に俯瞰マップ中心の2領域「機械学習手法・モデル」、「ニューラルネットワーク」の各密集領域（赤い楕円）の詳細を示す。 KWは、集積の源となった筆頭トピックに含まれる上位5キーワードを示す。この俯瞰マップにより、特にニューラルネットワークが機械学習手法の範疇にどのように組み込まれているか、また、どの技術が外に拡張しているかが明らかになる。

ニューラルネットワークは、深層学習の活性化関数やファインチューニング、転移学習、自己教師あり学習など、現在の自然言語処理や画像処理で広く用いられている技術を包含しており、これらは機械学習手法の領域に位置している。このことから、ニューラルネット技術がいかにして進化し、現代の様々な応用へと適用されているかが解析できる。「機械学習手法・モデル」領域は、実質、ニューラルネット発展技術と言えるものの、俯瞰マップ上では異なる場所に位置している。これらの観察から、ニューラルネットワーク技術と機械学習手法との間には密接な関連がありながらも、それぞれの発展と応用において独自の進化を遂げていることが示唆される。

図3-3に俯瞰マップ上側の2領域「画像処理」、「顔分析」の各密集領域（赤い楕円）の詳細を示した。俯瞰マップ中央の「機械学習手法・モデル」の応用先として「画像処理」の領域が接して形成されたとみなせる。「画像処理」から、更に枝分かれして「顔分析」と「動画・動産分析」の領域が接して形成されている。「画像処理」の「自然言語処理」にやや近い領域にマルチモーダル認識、キャプション生成などの技術がある。

図3-4に俯瞰マップ左側の3領域「自然言語処理」、「音声処理」、「動画・動作分析」の各密集領域（赤い楕円）の詳細を示す。「機械学習手法・モデル」のもう一つの主要な応用先として「自然言語処理」が接している。この「自然言語処理」の周囲に「画像処理」、「動画・動作分析」、「音声処理」が配置された。大規模言語モデル、プロンプトエンジニアリングなどの密集領域も形成されている。

図3-5に俯瞰マップ中央右側の6領域「最適化問題」、「意思・判断」、「情報・予測」、「クラスタリング」、「予測」、「ドメイン適応」の各密集領域（赤い楕円）の詳細を示した。「機械学習手法・モデル」に接して「最適化問題」と「意思・判断」が配置された。「最適化問題」には、「ニューラルネットワーク」と同様に、深層学習における基礎技術である勾配降下法などが含まれている。これら2領域は、深層学習の基礎技術として「機械学習手法・モデル」に接していると解釈できる。「意思・判断」領域には反実仮想説明の領域が形成された。「意思・決定」に接して「情報・予測」、「最適化問題」の付近に「クラスタリング」、（時系列）「予測」、

「ドメイン適応」の領域がある。

**表3-1　区分と領域**

| 技術区分（領域数） | 技術領域 |
|---|---|
| 機械学習手法・モデル（30） | 「Disentangledな表現学習」「データ拡張」「特徴選択」「教師あり学習・ラベル」「脳波解析」「自己教師あり学習・対照学習など」「ゼロショット学習」「クラス（未見、少数派）問題」「損失関数」「ファインチューニング」「表現学習」「マルチタスク学習」「転移学習」「分類器」「継続学習」「評価指標（関数）」「能動学習」「不確実性推定」「深層学習」「予測」「カーネル学習法」「大規模データセット」「サポートベクターマシン」「学習アルゴリズム」「スパースモデリング」「自動機械学習　（AutoML）など」「スパース表現」「構造化」「公平性評価」「ランキング学習」 |
| ニューラルネットワーク（7） | 「CNN」「予測精度」「量子化」「ニューラルネットワーク全般」「ニューロン　（モデル）」「SNN」「RNN」 |
| 画像処理（32） | 「形状識別・分析」「輝度・照明識別」「人物認識」「姿勢推定」「深度推定」「カメラ位置推定・キャリブレーション」「多視点解析」「線分認識」「シーン検出・解析」「物体認識・COCOデータセット」「点群データ処理」「画像マッチング・検索」「マルチモーダル認識」「顕著性検出」「キャプション生成」「微細画像認識」「解像度処理」「ハイパースペクトル画像分析」「画像再構築」「細胞検出・識別」「医療画像処理（がん）」「カラー識別・ハイライト」「画像位置合わせ」「心臓マッピング」「医療画像処理（腫瘍）」「敵対的生成ネットワーク」「画像処理技術全般」「セグメンテーション」「医療画像セグメンテーション」「画像説明」「距離測定・分析」「ローカル/グローバル特徴認識」 |
| 顔分析（4） | 「（顔）なりすまし攻撃対策」「表情認識」「顔認識」「属性識別」 |
| 自然言語処理（20） | 「アテンション機構」「視覚認識・特徴」「材料分析」「Transformerモデル」「語彙・感情分析」「コンテキスト認識」「テキスト認識・分類」「文表現・文意」「単語（埋込み）処理」「プロンプトエンジニアリング」「文書識別」「大規模言語モデル」「エンティティ認識」「関係抽出・分類」「長・短期記憶」「知識グラフ」「QA」「対話分析・生成」「言語モデル・自然言語処理」「機械翻訳」 |
| 音声処理（8） | 「音声データ分析・生成」「話者認識」「音声言語認識」「音声認識全般」「音声イベント検出」「音声強調」「音源分離」「ノイズ処理」 |
| 動画・動作分析（8） | 「動作認識」「オプティカルフロー推定」「アクション認識」「動画解析」「ビデオオブジェクトセグメンテーション」「物体追跡」「ビデオコーディング」「ターゲット追跡」 |
| 最適化問題（16） | 「（非負）行列因数分解」「関数近似・モデリング」「高次交互作用モデリング」「最適化手法・アルゴリズム」「最適化問題・解法」「多様体学習」「次元削減」「ガウス過程」「確率分布」「マルコフ連鎖モンテカルロ法」「変分推論」「サンプル選択・サイズ」「プロセスマイニング」「変数・変数選択」「パラメータ推定」「勾配降下法」 |
| 意思・判断（7） | 「決定木分析」「反実仮想説明」「ブラックボックス」「オンラインアルゴリズム」「多腕バンディット問題」「データセット」「意思決定」 |
| 情報・予測（4） | 「通信量推定・安全性」「相互情報量最大化」「リアルタイム予測」「市場予測」 |
| クラスタリング（2） | 「クラスタリング」「グラフニューラルネットワーク」 |
| 予測（2） | 「予測手法」「時系列予測」 |
| ドメイン適応（2） | 「ドメイン適応」「分布シフト分析」 |
| ロボット・自動運転（28） | 「コンピュータビジョン」「行動・動作認識」「SLAM」「位置推定」「マップ構築」「自律探査ロボット」「センサー」「ロボット制御・システム」「触覚センシング」「外科ロボット」「把握制御」「力制御」「手指動作分析・推定」「ソフトロボット制御」「（タスク指向）ダイアログ」「人間参加型AI」「模倣学習」「強化学習・ポリシー」「状態空間モデル」「プランニング」「自動制御・安全制御」「（マルチ）エージェント」「ロボット学習」「経路計画」「航路予測」「自動運転制御」「交通制御」「UAV制御」 |
| クラウド・セキュリティ（17） | 「検索手法」「クエリ分析・最適化」「データ分析・管理全般」「システム負荷分散」「エネルギー分析・予測」「連合学習」「プライバシー保護」「クラウドコンピューティング」「サービス品質分析」「IoTデータ分析」「信頼性評価」「サイバーセキュリティ対策」「Webデータセキュリティ」「暗号化」「（無線）通信セキュリティ」「セキュリティ攻撃分析」「セキュリティ分析」 |
| ハードウェアアーキテクチャ（4） | 「GPU処理技術」「アーキテクチャ検索」「メモリ効率」「FPGAボード」 |

| | |
|---|---|
| プログラミング（4） | 「論理プログラミング」「プログラミング言語分析」「プログラム解析」「コード分析・生成」 |
| ネットワーク（1） | 「ネットワーク・システム全般」 |
| 医療（5） | 「遺伝子分析・特定」「ウイルス感染分析・調査」「医療データ分析・診断」「因果推論」「効果推定」 |
| Web・システム（8） | 「通信システム制御」「対話システム」「システム全般」「レコメンダーシステム」「SNS」「アクセス設計・制御」「コミュニティ」「ソーシャルメディア」 |
| システム設計・開発（6） | 「VR・AR開発」「設計・デザイン全般」「法律解釈・判断」「教育システム・分析」「ソフトウェア評価」「ソフトウェアエンジニアリング」 |
| AI・ゲーム（2） | 「ゲーム理論」「AI倫理・システム」 |
| ブロックチェーン（4） | 「コンセンサスプロトコル」「スマートコントラクト」「ブロックチェーンシステム」「トランザクション制御」 |
| テスト（2） | 「バグ検出・分析」「テスト支援・自動化」 |
| 量子計算・暗号（3） | 「マルチパーティ計算・プロトコル」「鍵暗号・量子鍵配送」「量子コンピューティング」 |
| その他（32） | 「タンパク質構造分析」「ニュース分析」「ハッシュ関数推定」「感情識別」「将来予測（研究・仕事）」「スタイル変換」「グループ分析・比較」「知識蒸留」「視線分析・推定」「カルマンフィルター」「要約生成」「リスク評価・管理」「イベント検出」「投票プロトコル・システム」「バイアス分析」「CNN（音楽情報・画像検索）」「周波数分析」「光学文字認識」「敵対的攻撃分析」「脳マッピング・機能分析」「コンセプト推定」「分布外検出」「オートエンコーダ」「テンソル分解」「パターン分析」「異常検知」「正規化フロー」「方程式解法」「忠実度」「依存関係分析」「マルウェア検出」「トピックモデル」 |

図3-6に俯瞰マップ左下の領域「ロボット・自動運転」内の各密集領域（赤い楕円）の詳細を示した。中央の「機械学習手法・モデル」からやや離れて「ロボット・自動運転」の領域が形成された。俯瞰マップ中心寄りに強化学習の技術、外側にハード的なロボットの技術と自動運転の技術が配置された。ロボットとその周辺には国際会議議事録が多い。

図3-7に俯瞰マップ右下の5領域「クラウド・セキュリティ」、「ハードウェア・アーキテクチャ」、「プログラミング」、「ネットワーク」、「医療」の各密集領域（赤い楕円）の詳細を示した。「クラウド・セキュリティ」領域は、データ分析を中心に、その実施体系であるクラウドコンピューティングとそのエネルギー予測やシステム負荷分散などの実用例、更にその外側にセキュリティ関連の技術が集積した。この「クラウド・セキュリティ」を実現するための技術として「ハードウェア・アーキテクチャ」と「プログラミング」、「ネットワーク」などの技術が接している。またこれらの技術のメジャーに適用先としての「医療」の領域が形成された。

図3-8に俯瞰マップ下側の6領域「Web・システム、「システム設計・開発」、「AI・ゲーム」、「ブロックチェーン」、「テスト」、「量子計算・暗号」の各密集領域（赤い楕円）の詳細を示した。「クラウド・セキュリティ」領域の技術の応用先として、「Web・システム」、「AI・ゲーム」、「ブロックチェーン」の領域が形成された。その他、「Web・システム」の実現技術としての「システム設計・開発」や、「量子計算・暗号」、「テスト」などの領域も俯瞰マップ最外部に集積した。

図3-9にその他俯瞰マップ周辺の孤立した各密集領域（赤い楕円）の詳細を示す。俯瞰マップ外周部には、タンパク質構造分析、ニュース分析、投票プロトコル・システムなど、他の関連技術が少ない孤立した技術が配置された。これらもAI関連研究の重要な適用先、あるいは要素技術である。

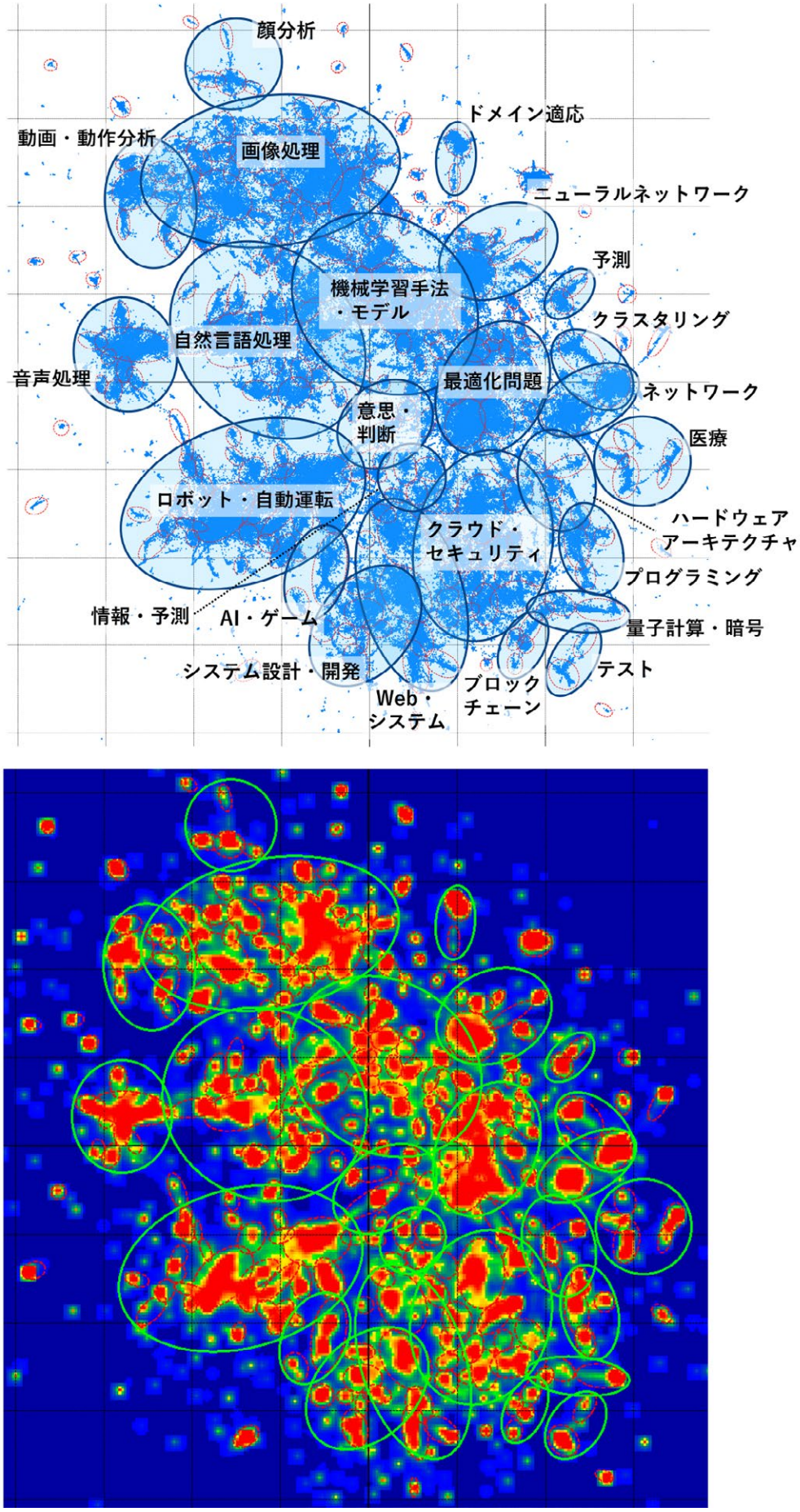


図3-1　　AI論文俯瞰マップ

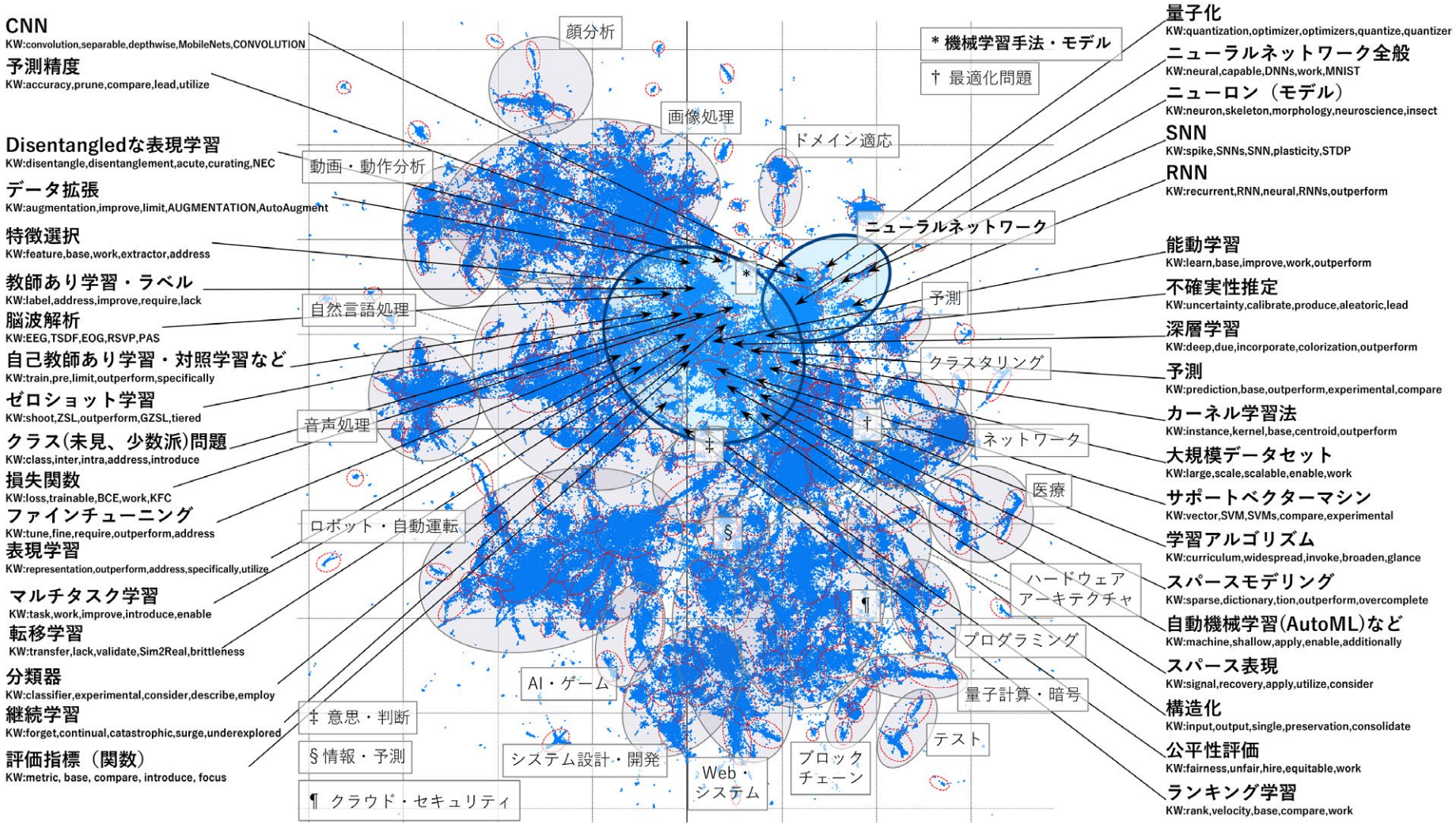


図 3-2　「機械学習手法・モデル」「ニューラルネットワーク」

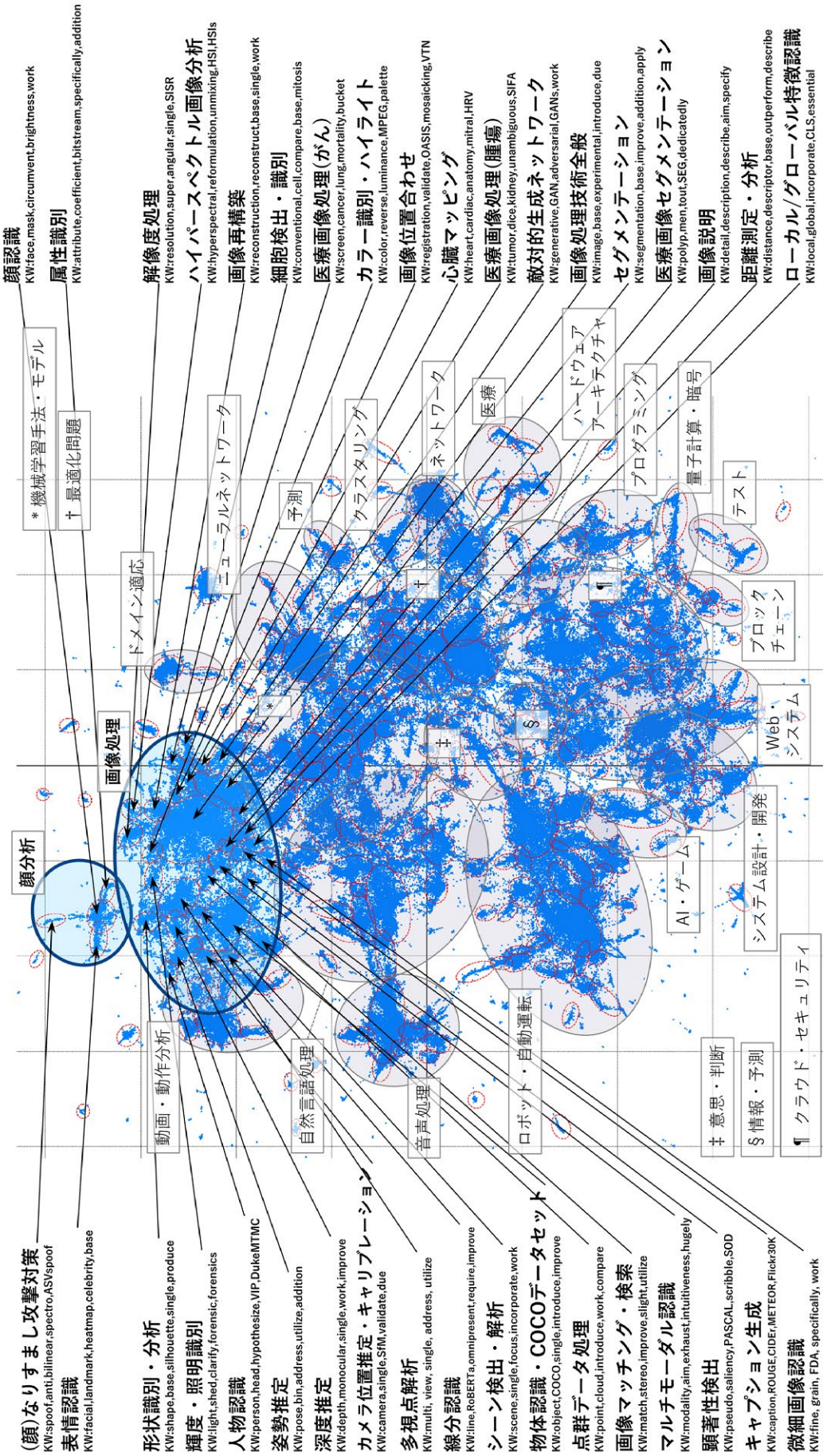
顔認識
KW:face,mask,circumvent,brightness,work
属性識別
KW:attribute,coefficient,bitstream,specifically,addition
解像度処理
KW:resolution,super,angular,single,SISR
ハイパースペクトル画像分析
KW:hyperspectral,reformulation,unmixing,HSI,HSIs
画像再構築
KW:reconstruction,reconstruct,base,single,work
細胞検出・識別
KW:conventional,cell,compare,base,mitosis
医療画像処理(がん)
KW:screen,cancer,lung,mortality,bucket
カラー識別・ハイライト
KW:color,reverse,luminance,MPEG,palette
画像位置合わせ
KW:registration,validate,OASIS,mosaicking,VTN
心臓マッピング
KW:heart,cardiac,anatomy,mitral,HRV
医療画像処理(腫瘍)
KW:tumor,dice,kidney,unambiguous,SIFA
敵対的生成ネットワーク
KW:generative,GAN,adversarial,GANs,work
画像処理技術全般
KW:image,base,experimental,introduce,due
セグメンテーション
KW:segmentation,base,improve,addition,apply
医療画像セグメンテーション
KW:polyp,men,tout,SEG,dedicatedly
画像説明
KW:detail,description,describe,aim,specify
距離測定・分析
KW:distance,descriptor,base,outperform,describe
ローカル/グローバル特徴認識
KW:local,global,incorporate,CLS,essential
(顔)なりすまし攻撃対策
KW:spoof,anti,bilinear,spectro,ASVspoof
表情認識
KW:facial,landmark,heatmap,celebrity,base
形状識別・分析
KW:shape,base,silhouette,single,produce
輝度・照明識別
KW:light,shed,clarify,forensic,forensics
人物認識
KW:person,head,hypothesize,VIP,DukeMTMC
姿勢推定
KW:pose,bin,address,utilize,addition
深度推定
KW:depth,monocular,single,work,improve
カメラ位置推定・キャリブレーション
KW:camera,single,SfM,validate,due
多視点解析
KW:multi, view, single, address, utilize
線分認識
KW:line,RoBERTa,omnipresent,require,improve
シーン検出・解析
KW:scene,single,focus,incorporate,work
物体認識・COCOデータセット
KW:object,COCO,single,introduce,improve
点群データ処理
KW:point,cloud,introduce,work,compare
画像マッチング・検索
KW:match,stereo,improve,slight,utilize
マルチモーダル認識
KW:modality,aim,exhaust,intuitiveness,hugely
顕著性検出
KW:pseudo,saliency,PASCAL,scribble,SOD
キャプション生成
KW:caption,ROUGE,CIDEr,METEOR,Flickr30K
微細画像認識
KW:fine, grain, FDA, specifically, work
顔分析
画像処理
ドメイン適応
* 機械学習手法・モデル
† 最適化問題
ニューラルネットワーク
予測
クラスタリング
ネットワーク
医療
ハードウェア アーキテクチャ
プログラミング
量子計算・暗号
テスト
ブロック チェーン
Web・システム
AI・ゲーム
システム設計・開発
動画・動作分析
自然言語処理
音声処理
ロボット・自動運転
‡ 意思・判断
§ 情報・予測
¶ クラウド・セキュリティ

**図3-3　「画像処理」「顔分析」**

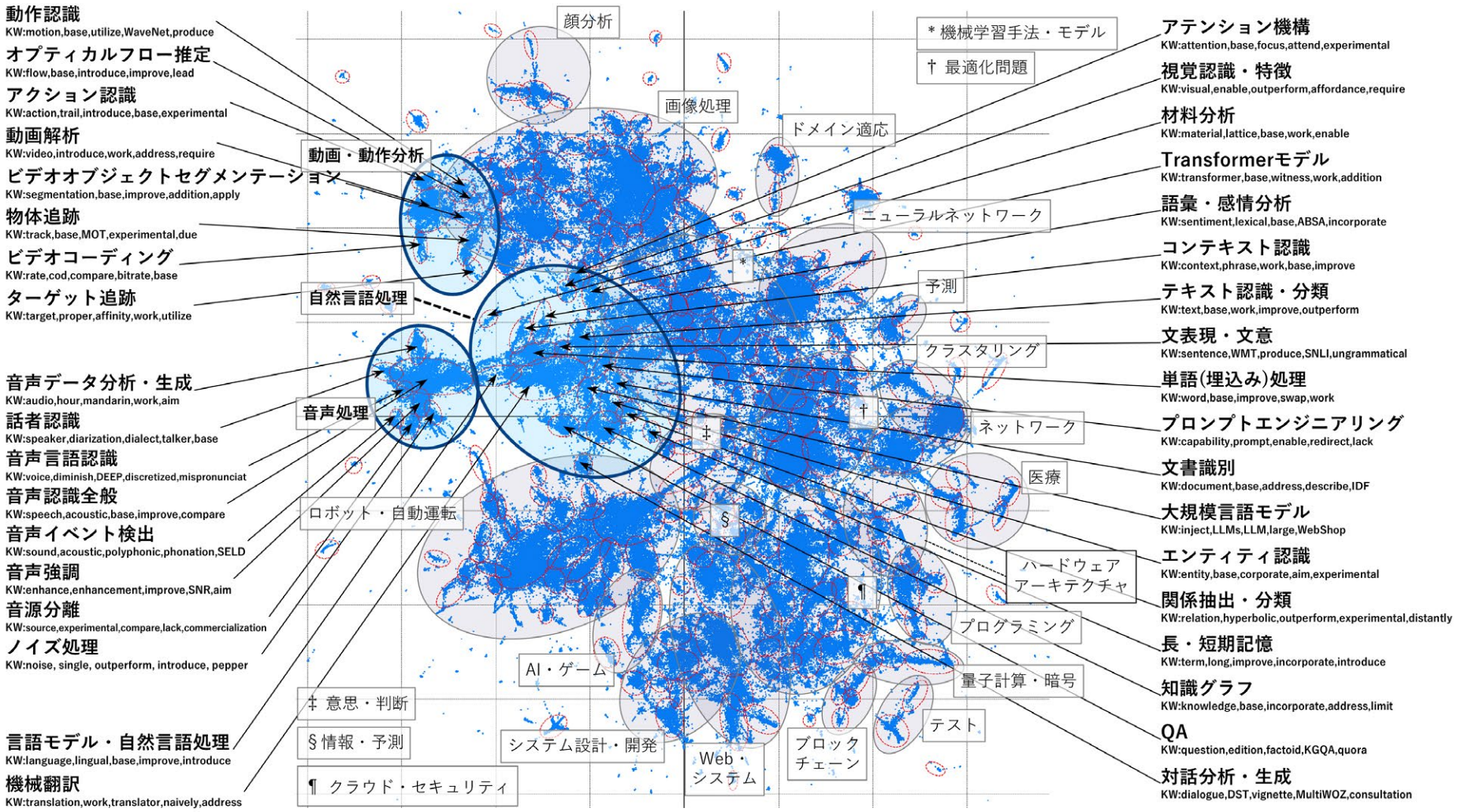


**図 3-4　「自然言語処理」「音声処理」「動画・動作分析」**

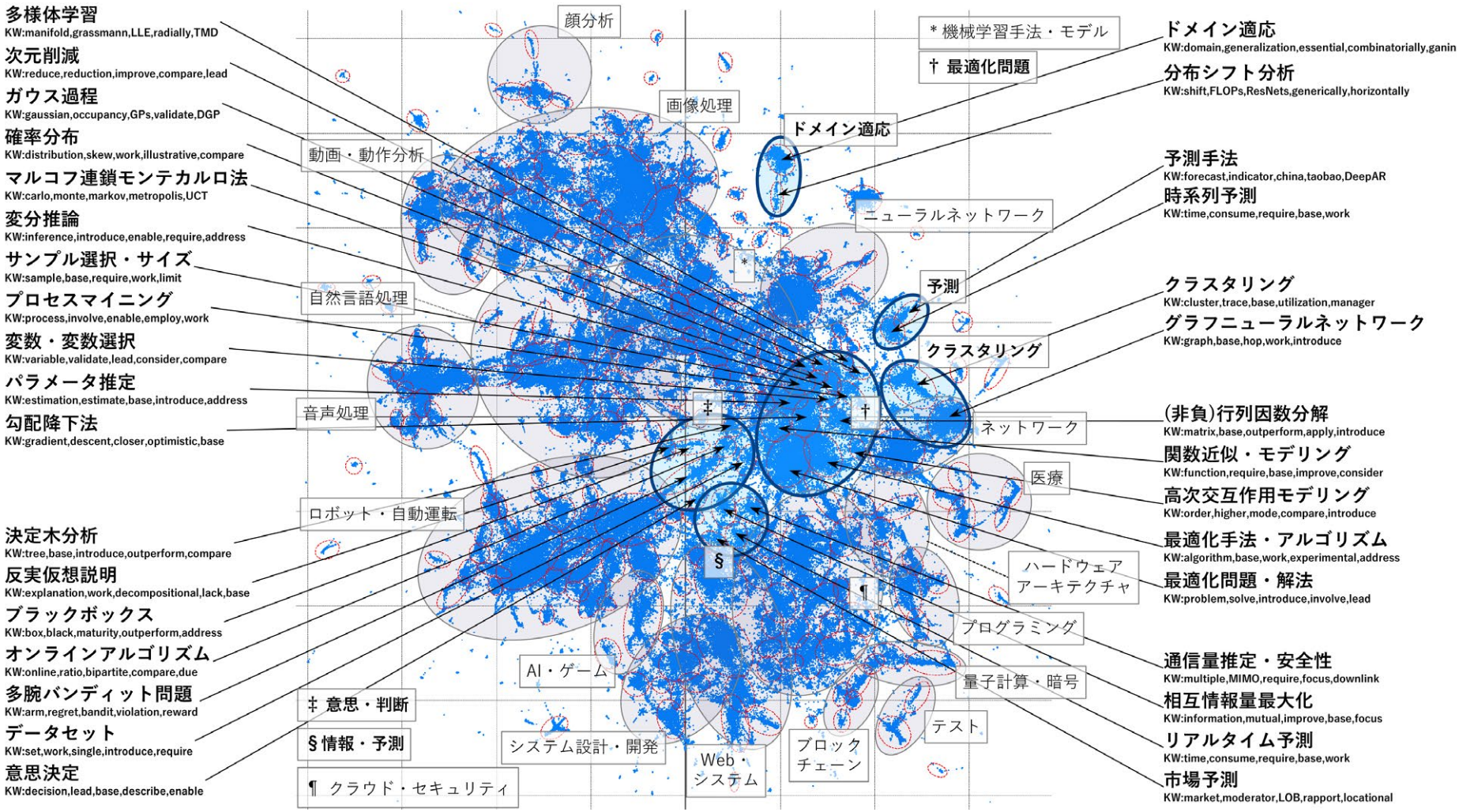


**図 3-5　「最適化問題」「意思・判断」「情報・予測」「クラスタリング」「予測」「ドメイン適応」**

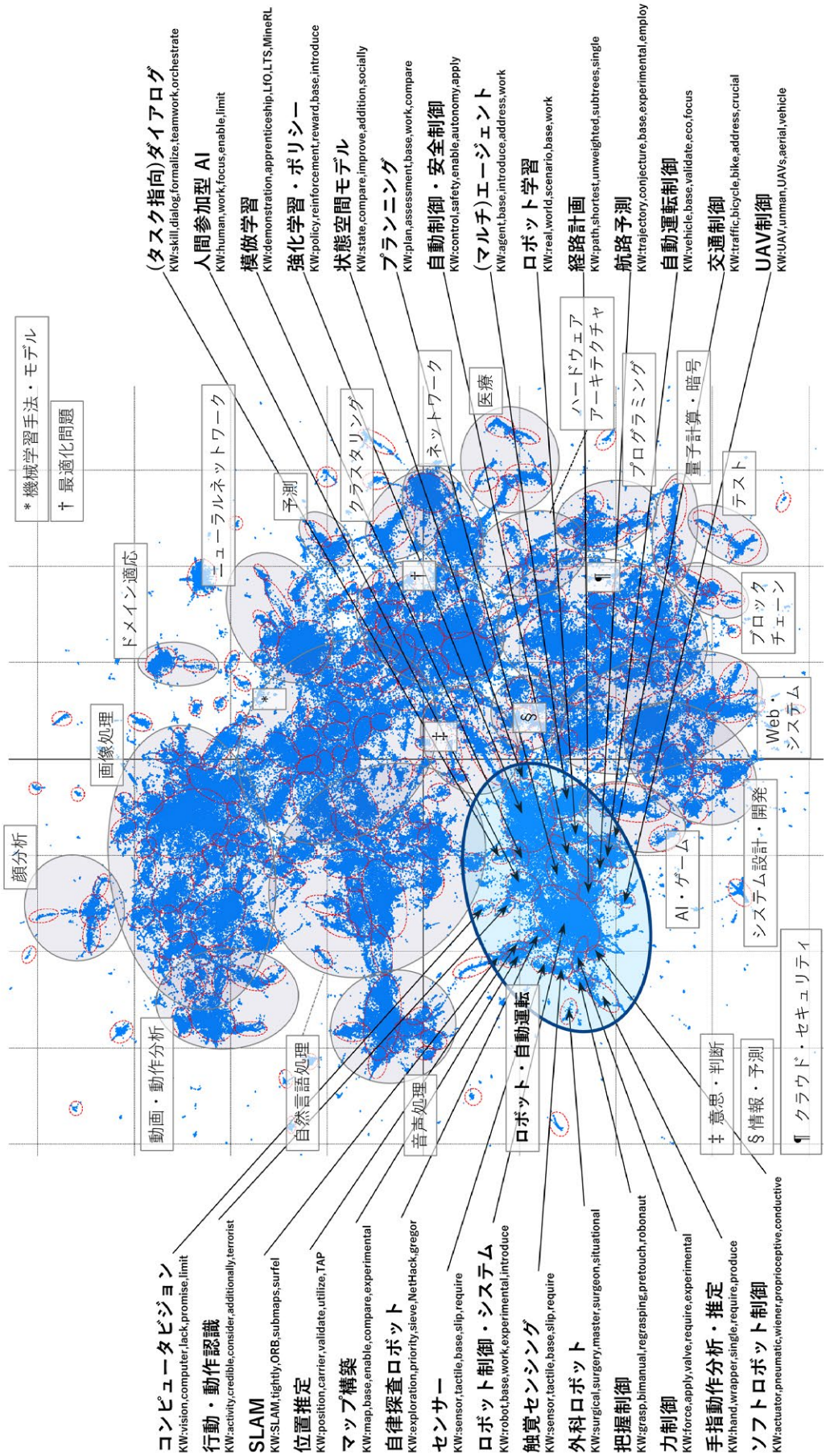


**図3-6　「ロボット・自動運転」**

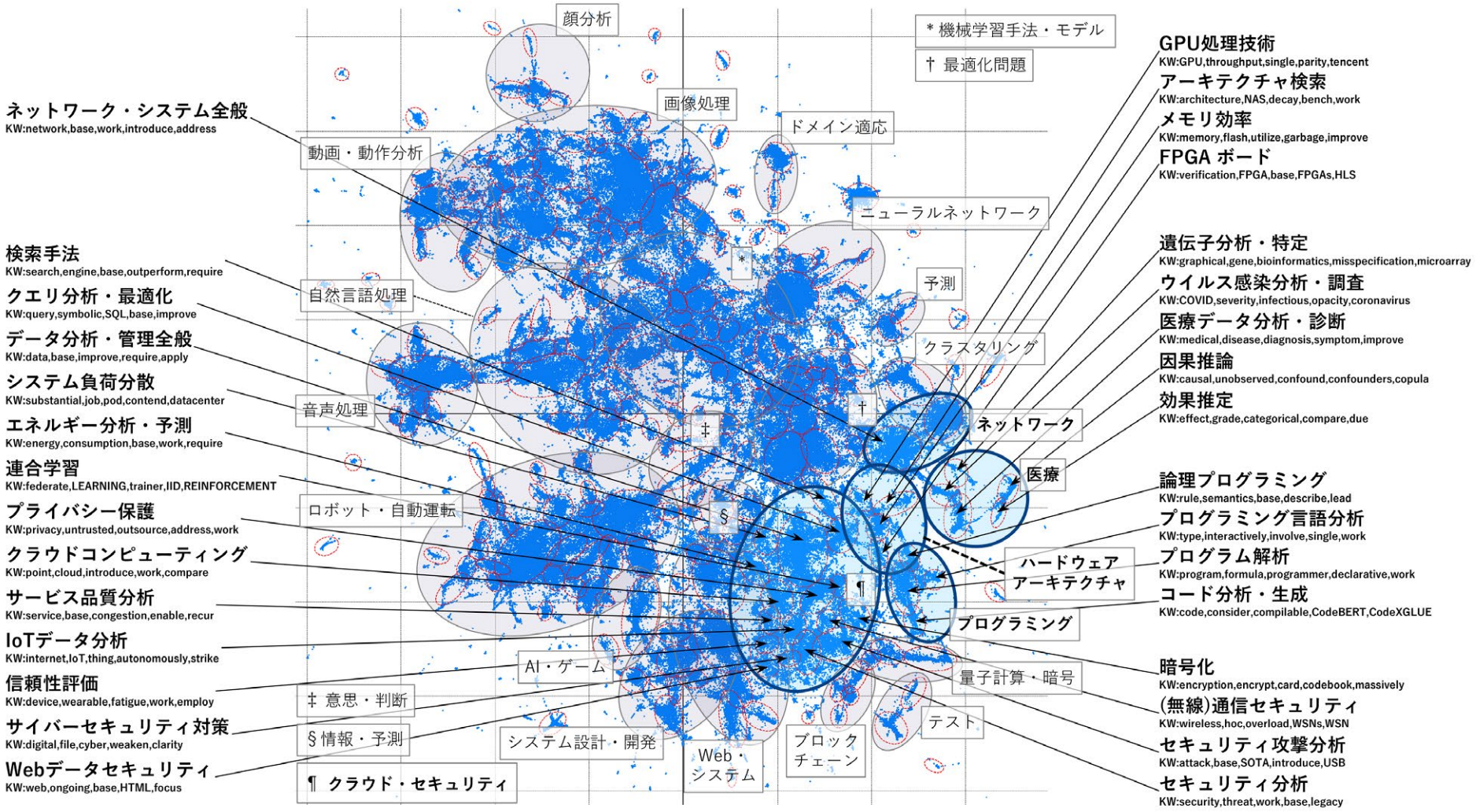


**図 3-7　「クラウド・セキュリティ」「ハードウェア・アーキテクチャ」「プログラミング」「ネットワーク」「医療」**

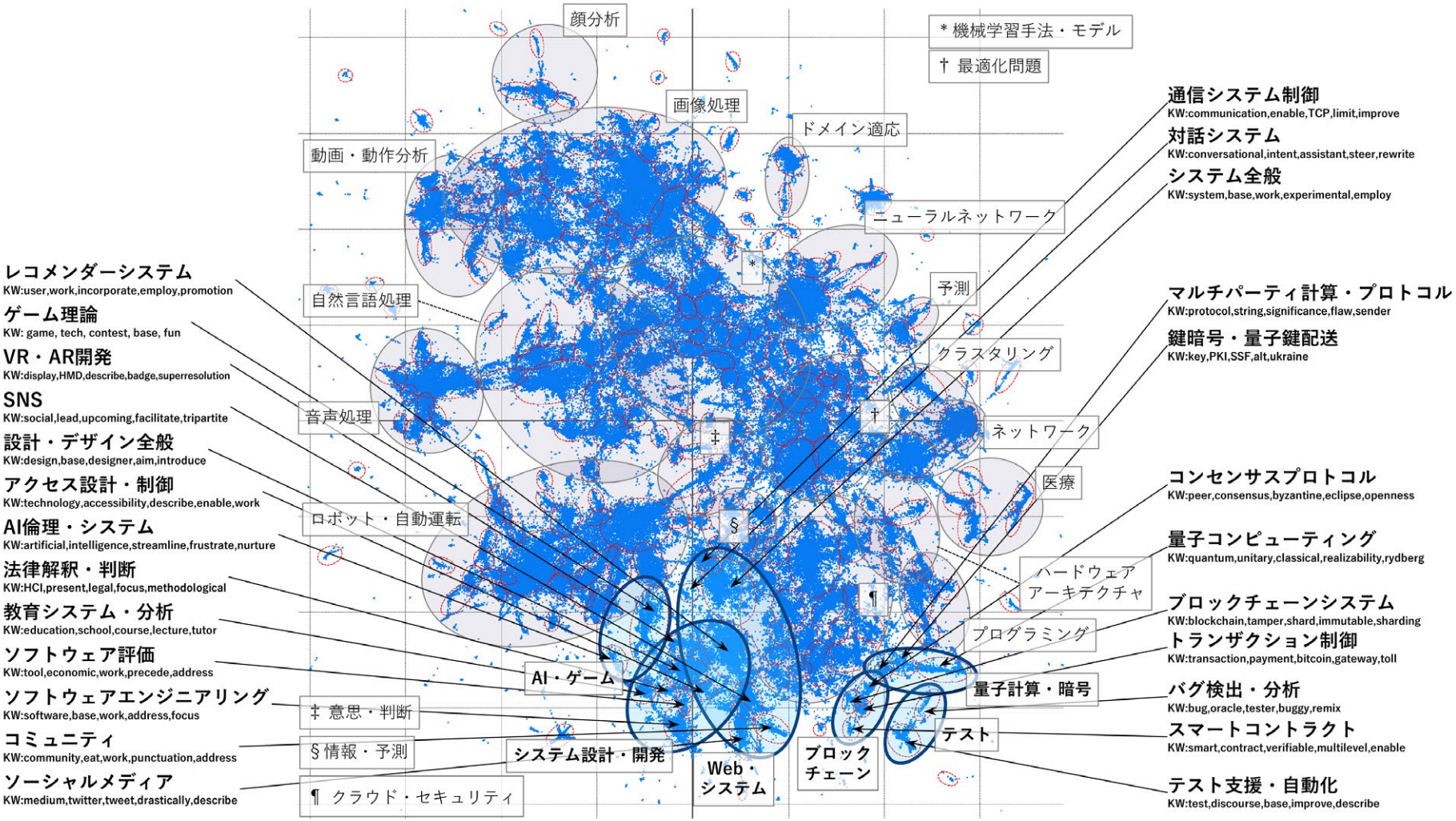


**図 3-8　「Web・システム」「システム設計・開発」「AI・ゲーム」「ブロックチェーン」「テスト」「量子計算・暗号」**

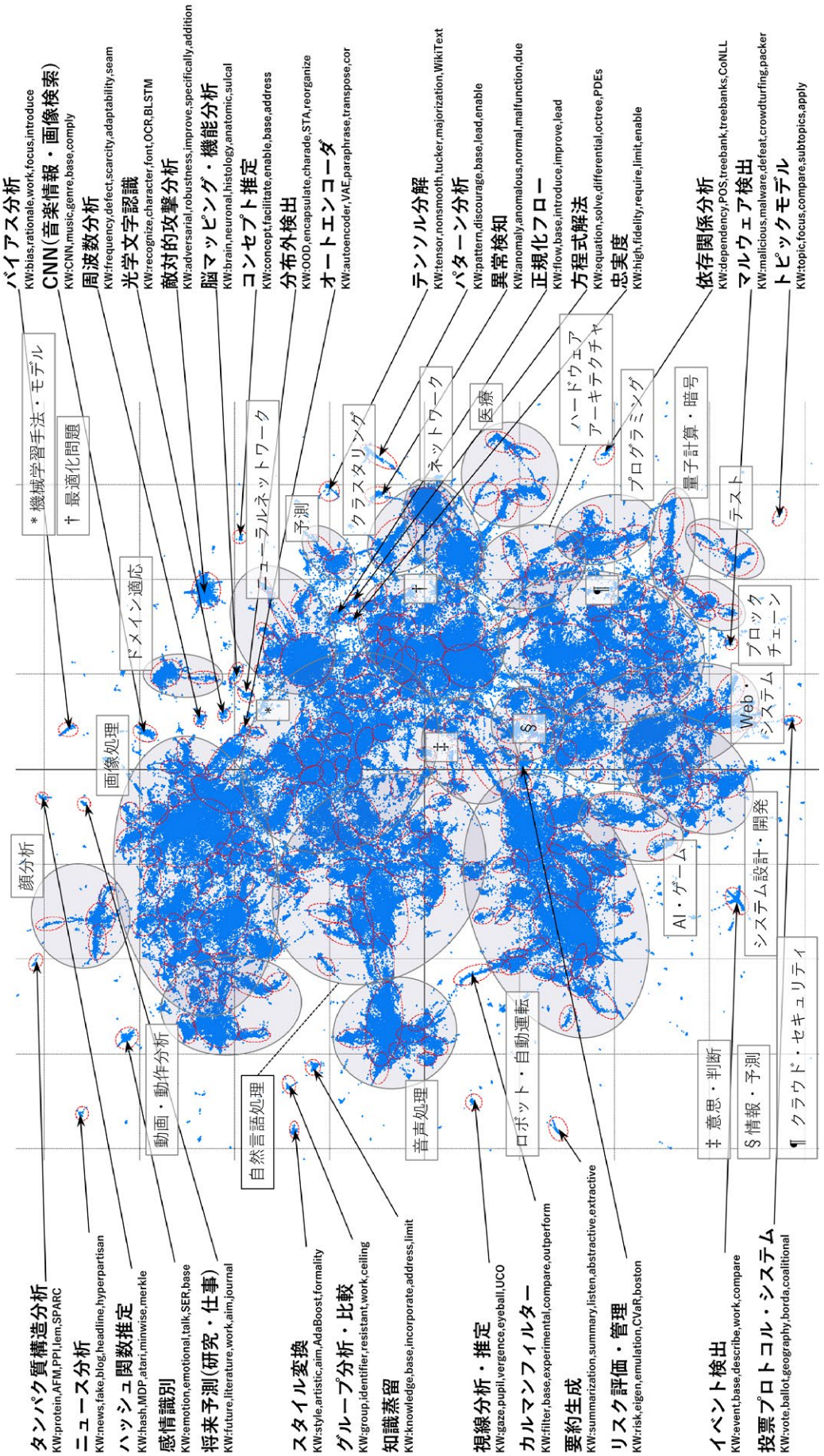
バイアス分析
KW:bias,rationale,work,focus,introduce
CNN(音楽情報・画像検索)
KW:CNN,music,genre,base,comply
周波数分析
KW:frequency,defect,scarcity,adaptability,seam
光学文字認識
KW:recognize,character,font,OCR,BLSTM
敵対的攻撃分析
KW:adversarial,robustness,improve,specifically,addition
脳マッピング・機能分析
KW:brain,neuronal,histology,anatomic,sulcal
コンセプト推定
KW:concept,facilitate,enable,base,address
分布外検出
KW:OOD,encapsulate,charade,STA,reorganize
オートエンコーダ
KW:autoencoder,VAE,paraphrase,transpose,cor
テンソル分解
KW:tensor,nonsmooth,tucker,majorization,WikiText
パターン分析
KW:pattern,discourage,base,lead,enable
異常検知
KW:anomaly,anomalous,normal,malfunction,due
正規化フロー
KW:flow,base,introduce,improve,lead
方程式解法
KW:equation,solve,differential,octree,PDEs
忠実度
KW:high,fidelity,require,limit,enable
依存関係分析
KW:dependency,POS,treebank,treebanks,CoNLL
マルウェア検出
KW:malicious,malware,defeat,crowdturfing,packer
トピックモデル
KW:topic,focus,compare,subtopics,apply
タンパク質構造分析
KW:protein,AFM,PPI,lem,SPARC
ニュース分析
KW:news,fake,blog,headline,hyperpartisan
ハッシュ関数推定
KW:hash,MDP,atari,minwise,merkle
感情識別
KW:emotion,emotional,talk,SER,base
将来予測(研究・仕事)
KW:future,literature,work,aim,journal
スタイル変換
KW:style,artistic,aim,AdaBoost,formality
グループ分析・比較
KW:group,identifier,resistant,work,ceiling
知識蒸留
KW:knowledge,base,incorporate,address,limit
視線分析・推定
KW:gaze,pupil,vergence,eyeball,UCO
カルマンフィルター
KW:filter,base,experimental,compare,outperform
要約生成
KW:summarization,summary,listen,abstractive,extractive
リスク評価・管理
KW:risk,eigen,emulation,CVaR,boston
イベント検出
KW:event,base,describe,work,compare
投票プロトコル・システム
KW:vote,ballot,geography,borda,coalitional
* 機械学習手法・モデル
† 最適化問題
顔分析
画像処理
ドメイン適応
動画・動作分析
自然言語処理
音声処理
ニューラルネットワーク
予測
クラスタリング
ネットワーク
医療
ハードウェア アーキテクチャ
プログラミング
量子計算・暗号
テスト
ブロック チェーン
Web システム
ロボット・自動運転
AI・ゲーム
システム設計・開発
意思・判断
情報・予測
クラウド・セキュリティ

図3-9 その他の領域

# 4 ｜ arXivと国際会議の比較

## 4.1 全体比較

arXiv収録データ427,846件とScopus収録の国際会議議事録データ200,663件、合計628,509件について詳細な分析を行った。両データソースを用いてヒートマップを作成し、データの分布と量的差異を視覚的に評価した（図4-1）。

分析の結果、arXivと国際会議データの間で、データの存在自体にはほとんど差が見られなかった。しかし、具体的なデータ量では顕著な違いが認められた。特に、国際会議のみのデータでは俯瞰マップの中心部のデータ量が相対的に少なく、俯瞰マップの全体的な構造はarXivのデータが主に反映されていると考えられる。今回分析対象とした国際会議議事録では、ハードウェア関連の技術記録が比較的少ないことが確認された。「ニューラルネットワーク」領域におけるニューラルネットワーク全般や、「画像処理」領域における画像処理技術全般など、データ件数の多い領域で、明らかにarXivに偏在している傾向が見られた。

興味深いことに、これらの2つの領域にある文献の被引用数は他の領域と比べて少ない傾向にある。この事実は、arXivの「玉石混交」的な性質を示唆している。つまり、arXivには質の高い研究と並んで、影響力の少ない研究も多く含まれていることが推測される。

また、国際会議議事録では、大規模言語モデルやプロンプトエンジニアリングのような新しい技術に関する文献の件数が、arXivと比較して少ないことも明らかになった。この傾向は、arXivが最新の研究トレンドをより迅速に反映する傾向があることを示している。これらの分析結果は、研究動向を把握する上で、arXivと国際会議議事録の両方を考慮することの重要性を示唆している。arXivは最新のトレンドや幅広い研究を捉えるのに適している一方で、国際会議議事録はより厳選された研究成果を反映している可能性がある。

今後の研究では、これらのデータソース間の差異がどのように時間とともに変化するか、また、各データソースの特性が研究の質や影響力とどのように関連しているかを詳細に調査することが有益だろう。このような分析は、AI研究の動向をより正確に把握し、将来の研究方向を予測する上で重要な洞察を提供する可能性がある。

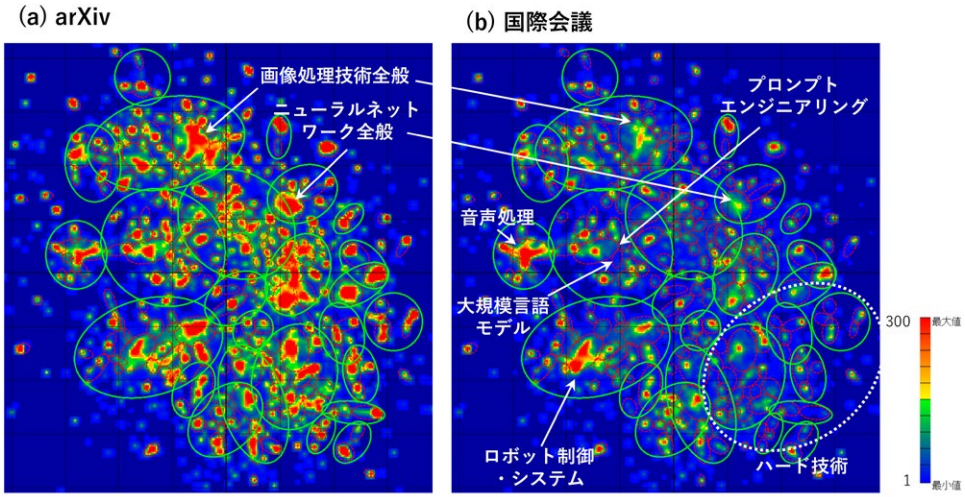


図4-1　論文俯瞰マップ（arXiv・国際会議議事録別プロット）

## 4.2 経年変化の比較

2010年から2023年までのarXivと国際会議議事録の件数推移およびその俯瞰マップ上での分布変化を分析した（図4-2）。深層学習が登場する以前の2010年当初は、国際会議議事録の件数が多かった。しかし、深層学習の出現後、arXivでの発表が急増し、現在では大きな差がついている。

国際会議での発表件数の伸びは比較的緩やかだが、これは発表枠に制限がある可能性がある。査読付き文献ではより顕著な増加が見られる可能性がある。なお、2023年のデータは収録途中であり、特に国際会議議事録の方がデータの欠損が多い点に注意が必要である。

俯瞰マッププロットでは、年ごとの密度が同じになるように感度を調整している。2010年当初は、マップの周辺部に多くのデータが分布し、中心部の深層学習関連はほぼ見られなかった。ただし、特徴選択などの技術にやや集積が見られた。深層学習が発表された2014年以降、ニューラルネットワーク全般や教師あり学習、マルチタスク学習など関連技術が出現し始めたが、その種類はまだ限られていた。2018年以降になると、現在主流となっている技術がほぼ全て出現している。2022年以降に新たに確認された技術には、Transformerモデル、大規模言語モデル、自己教師あり学習、反実仮想説明、マルチモーダル認識、ファインチューニングなどがある。これらの最新技術は、ほぼarXivでの収録が中心となっている。

この分析結果から、以下のような洞察が得られる：

- arXivは最新の研究トレンドをより迅速に反映する傾向がある。
- 深層学習の登場が、AI研究の量的拡大と質的変化をもたらした。
- 国際会議と比較して、arXivはより広範な研究を包含している。
- 最新技術の多くがarXivで先行して発表され、arXivが研究動向の早期把握に有用である可能性が高い。

今後の研究では、これらのトレンドがどのように継続または変化するか、また新技術の出現がAI研究全体にどのような影響を与えるかを注視する必要がある。さらに、arXivと国際会議での発表の質的差異や、各プラットフォームの特性が研究の普及や影響力にどのように関わるかを詳細に分析することも有益だろう。

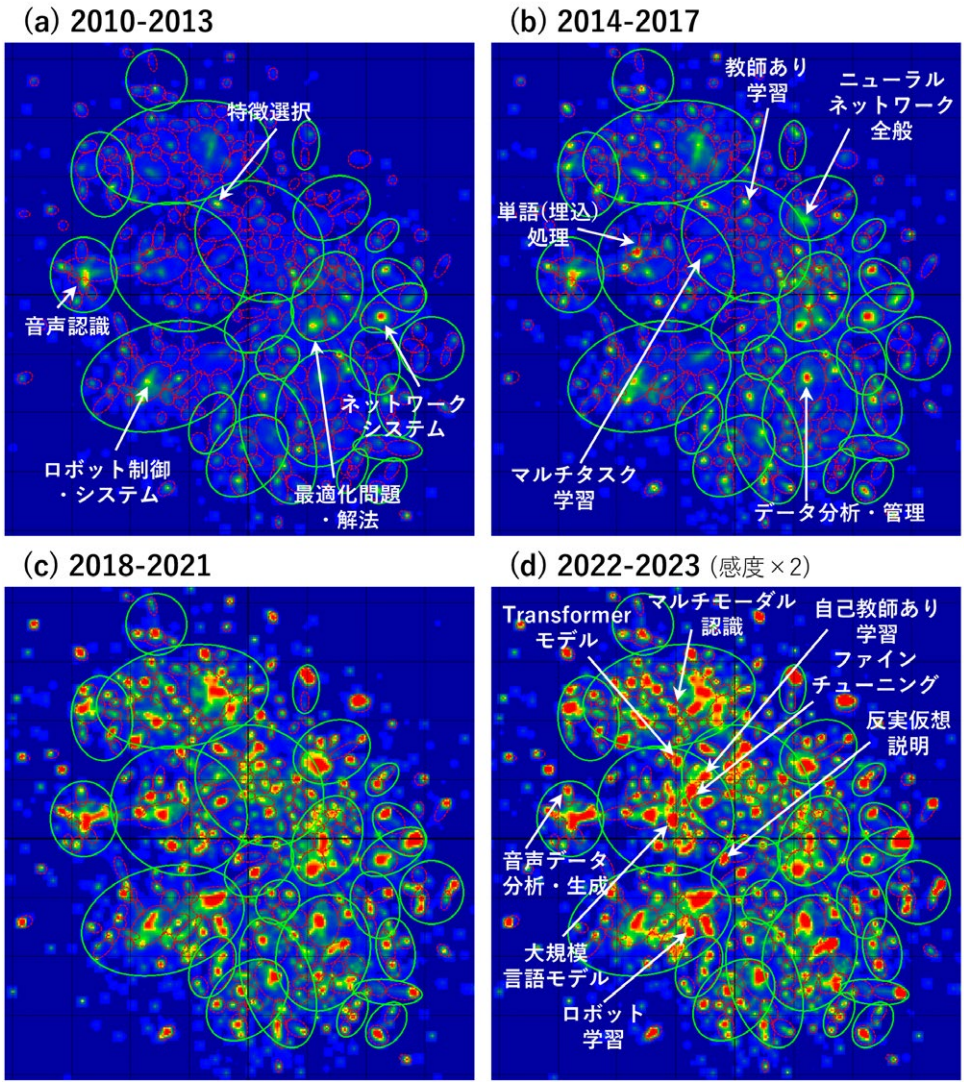


**図4-2　　論文俯瞰マップ（経年変化の比較）**

2010年から2023年までのarXivと国際会議議事録（Scopus収録）の俯瞰マップ上での分布変化を比較分析した（図4-3）。マッププロットでは、年ごとの密度が同じになるよう感度を調整している。

2010年当初、Scopus収録の国際会議議事録では特徴選択や教師あり学習など機械学習的技術が多く見られた一方、arXivでは俯瞰マップ下側の技術が中心であった。2014年以降になると、ニューラルネットワーク全般や画像処理技術全般でarXivの方が多くなったが、国際会議議事録側にも対応する文献は存在していた。

2018年以降は、arXivが圧倒的に優勢となり、より多くの技術が出現するようになった。2022年以降になると、大規模言語モデルやマルチモーダル認識などの新技術はほぼ全てarXiv側で明確に出現している。一方、国際会議議事録では2022年以降に出現した新技術は連合学習程度であった。ただし、arXivでは連合学習が2018年時点で既に確認できていた点は注目に値する。

この分析から、深層学習の黎明期ではarXivと国際会議議事録の間に大きな差はなかったものの、最近の大規模言語モデルなど最新技術の出現においてはarXivが国際会議議事録を大きく上回っていることが明らかになった。この傾向は、AI研究の動向把握におけるarXivの重要性を示唆しており、特に近年の急速な技術革新を反映している。

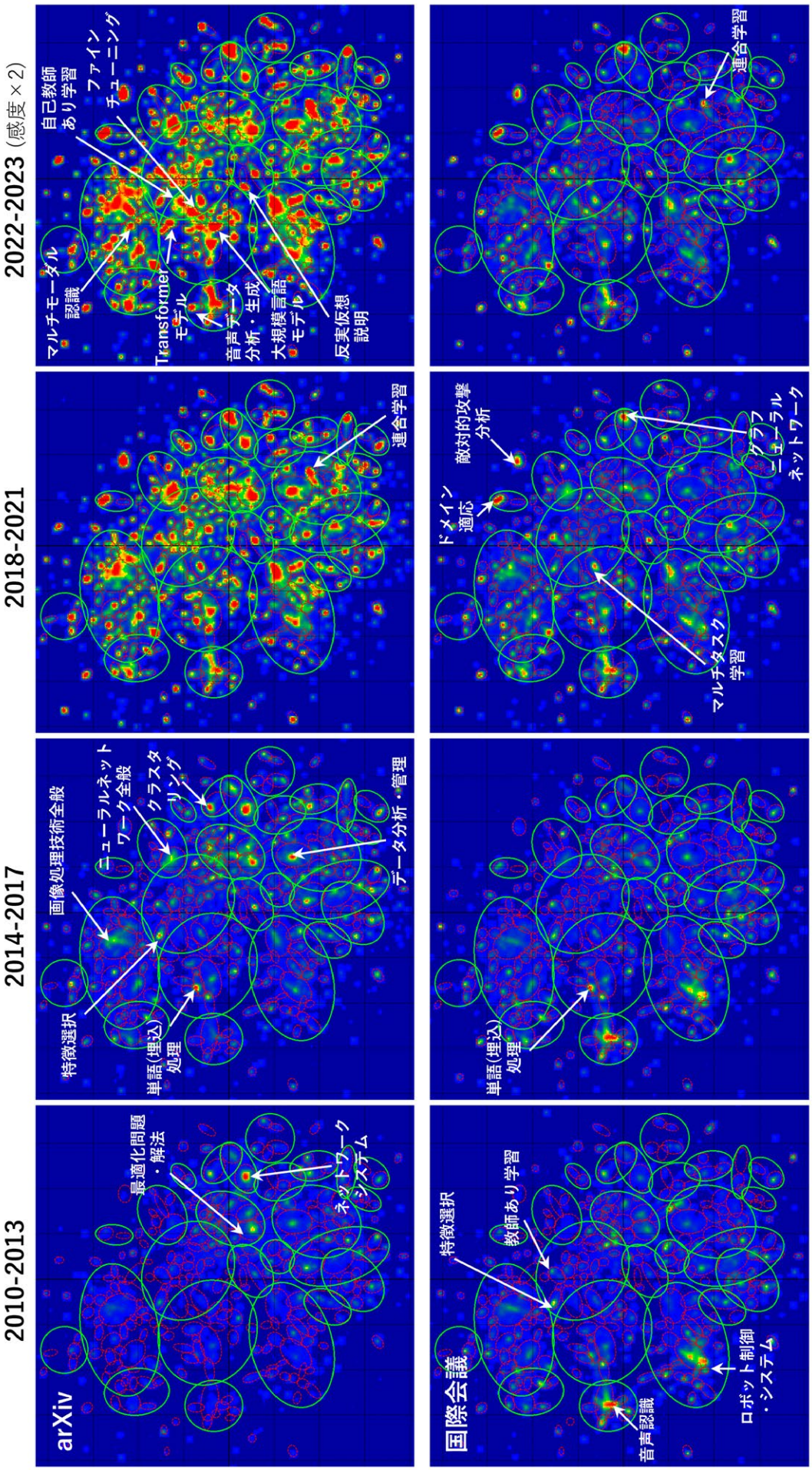


図4-3　論文俯瞰マップ（経年変化の比較）

## 4.3 マクロ動向把握

エマージングな領域の同定のため、俯瞰マップを縦横0.5×0.5のメッシュに分割して、各メッシュ別に成長曲線（ロジスティック関数）フィットを行い、立ち上り年を求めた（図4-4）。この値が新しい程、今後の成長が期待できる。データが欠損している2023年を除外した絶対評価（図4-4(a)）と、データを各年の平均値で相対化して2023年まで利用した相対評価（図4-4(b)）を行った。

絶対評価(a)では、全体的な増加傾向を反映して、5割以上の領域が成長判定となった。一方、相対評価(b)では、平均以上に増加している領域に着目し、直近で増加した領域を中心に今後も成長が続くと判定された。具体的には、大規模言語モデル、マルチモーダル認識、仮想現実説明などの領域が、直近の増加から今後も成長の可能性があると判断された。これらの注目領域は俯瞰マップ上で赤い破線の楕円で示されている。一方、Transformerモデルは直近では成長が見られたものの、文献データからは今後の継続的な成長を示す明確な兆候は見られなかった。

この分析手法により、AI分野における新興領域や成長が期待される技術を客観的に識別することが可能となった。ただし、技術の進展は予測困難な要素も多いため、この結果は一つの指標として捉え、他の要因も考慮しながら総合的に判断する必要がある。

図4-5に示したようにarXivと国際会議議事録それぞれに対して個別に成長分析を行ったところ、全体的な傾向に大きな差異はないことが確認された。 arXiv掲載から国際会議議事録への採録（およびScopusデータベースへの反映）までの時間差は約1年程度と見なせるため、立ち上がり年の判定には大きな影響を与えていないことが分かった。

詳細に見ると、俯瞰マップ上側の深層学習領域では、arXivと国際会議の個別判定結果が近似していた。一方、ハードウェアが中心的なトピックとして含まれる俯瞰マップ下側では、顕著な差異が見られた。

個別の技術に注目すると、大規模言語モデルについては興味深い結果が得られた。件数が少ない国際会議議事録では今後も成長すると判定されたが、既に多くの件数があるarXivでは成長判定とならなかった。これは、件数変化が急激すぎてフィッティングに失敗したためであることが詳細な確認で判明した。一方、プロンプトエンジニアリングについては、arXivで成長判定となった。この結果は、arXivが最新の研究トレンドをより迅速に反映する傾向があることを裏付けている。

これらの分析結果から、以下の洞察が得られる：

- arXivと国際会議議事録は、全体的な研究動向を把握する上で補完的な役割を果たしている。
- 深層学習のような主流な領域では、両データソース間で類似した傾向が見られる。
- ハードウェア関連など特定の領域では、データソース間で差異が生じる可能性がある。
- 急速に成長している新興分野（例：大規模言語モデル）では、従来の成長モデルが適切に機能しない場合がある。
- 最新のトレンド（例：プロンプトエンジニアリング）は、arXivでより早く捉えられる傾向がある。

今後の研究では、これらの差異の原因をより詳細に分析し、各データソースの特性を考慮した成長予測モデルの開発が求められる。また、急速に発展する分野に対応できる新たな分析手法の探索も重要な課題となるだろう。また、この手法によって同定されたエマージング領域に対して、より詳細な質的分析を行うことで、その成長要因や潜在的な応用可能性を探ることも重要である。

これらの分析結果は、研究動向の把握や将来予測において、複数のデータソースを統合的に分析することの重要性を示唆している。同時に、各データソースの特性や限界を理解した上で、慎重に解釈を行う必要性も強調している。効果的な研究開発戦略の立案のための重要な参考情報となり得るが、技術の急速な進歩や予期せぬブレークスルーの可能性を考慮し、常に最新の情報をモニタリングし、柔軟に対応することが肝要であろう。

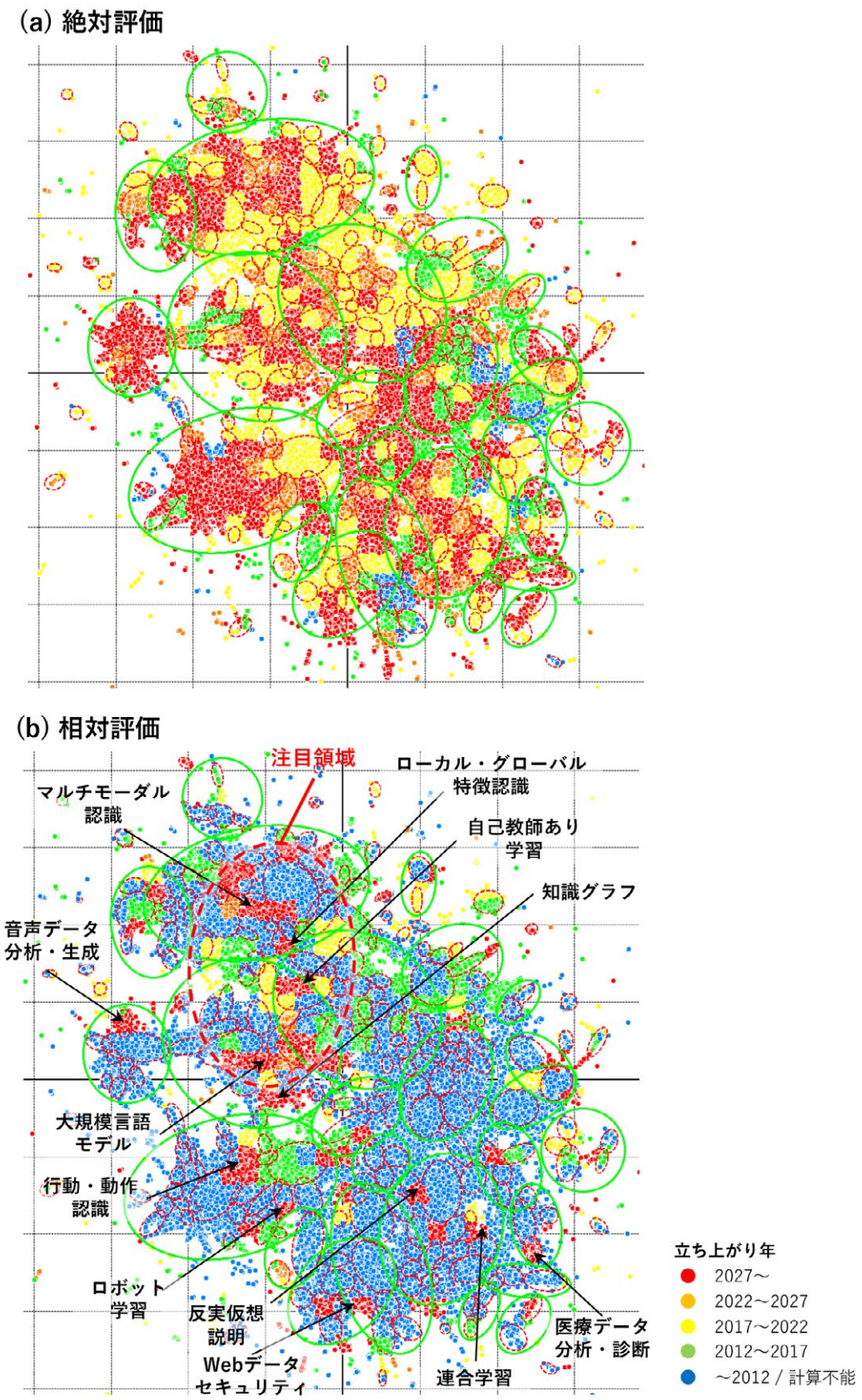


**図4-4　成長曲線フィットによる立ち上がり年（絶対評価と相対評価の比較）**

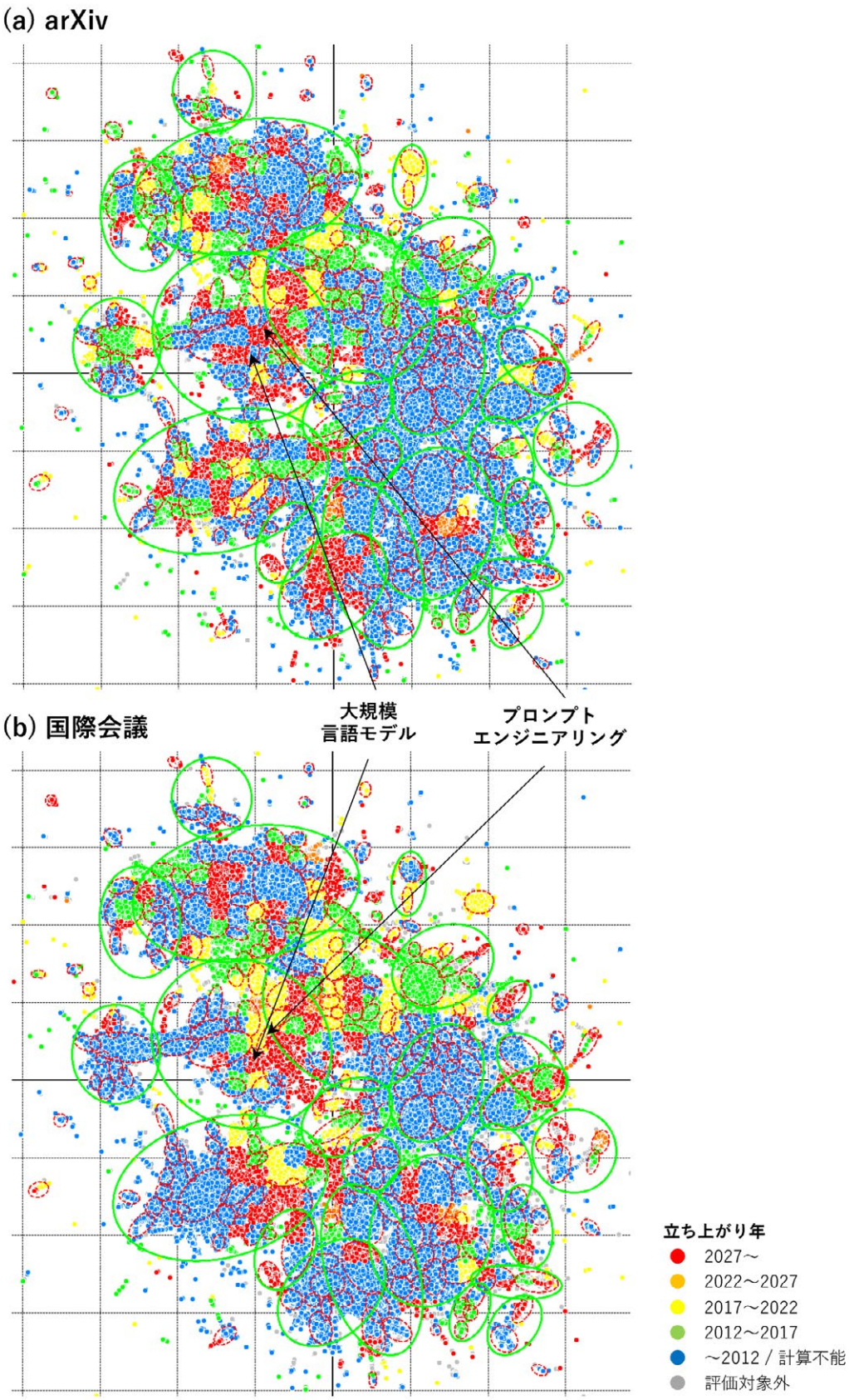


**図4-5　成長曲線フィットによる立ち上がり年（arXivと国際会議の比較）**

## 4.4 標準化被引用回数

技術領域の重要性を評価するため、俯瞰マップを0.5×0.5のメッシュに分割し、各メッシュの標準化被引用回数の平均を算出した（図4-6）。被引用回数は論文の掲載年が古いほど多くなる傾向があるため、掲載年ごとに平均値と標準偏差を求め、標準化した値で比較を行った。この値が0前後であれば平均的な注目度、正の値であれば平均以上の注目度を示している。なお、arXivの引用情報はSemantic Scholarから取得し、データ件数の差を考慮して全体を標準化した。

分析結果は、arXivと国際会議議事録のいずれを用いても全体的な傾向が一致しており、両者が同様の学術文献であることを示している。特に注目すべき領域は、俯瞰マップ上部の「機械学習手法・モデル」「画像処理」「自然言語処理」の境界領域にあり、具体的にはTransformerモデル、プロンプトエンジニアリング、大規模言語モデルなどが挙げられる。この結果は、成長曲線フィットによる立ち上がり年を用いた成長領域判定の結果とほぼ一致している。

詳細に分析すると、arXivと国際会議議事録の間にいくつかの差異が見られた。まず、大規模言語モデルやプロンプトエンジニアリングの評価が高いのはarXivのみであり、これはarXivの情報の先行性を示している。一方、ニューラルネットワークや画像処理技術などの広範な領域では、arXivの被引用数が国際会議議事録よりも低い傾向にある。これらの領域では、arXivの文献数が多く、国際会議議事録の文献数が少ないことから、この被引用数の差はarXivの内容の質にばらつきがあることを反映していると考えられる。

この分析結果は、新興技術の情報源としてarXivを評価する際の注意点を示唆している。 arXivは情報の先行性という利点を持つ一方で、内容の質にばらつきがあるという短所も認識すべきである。したがって、arXivを新興技術の唯一の有効な情報源とみなすのではなく、その特性を理解した上で適切に活用することが重要である。

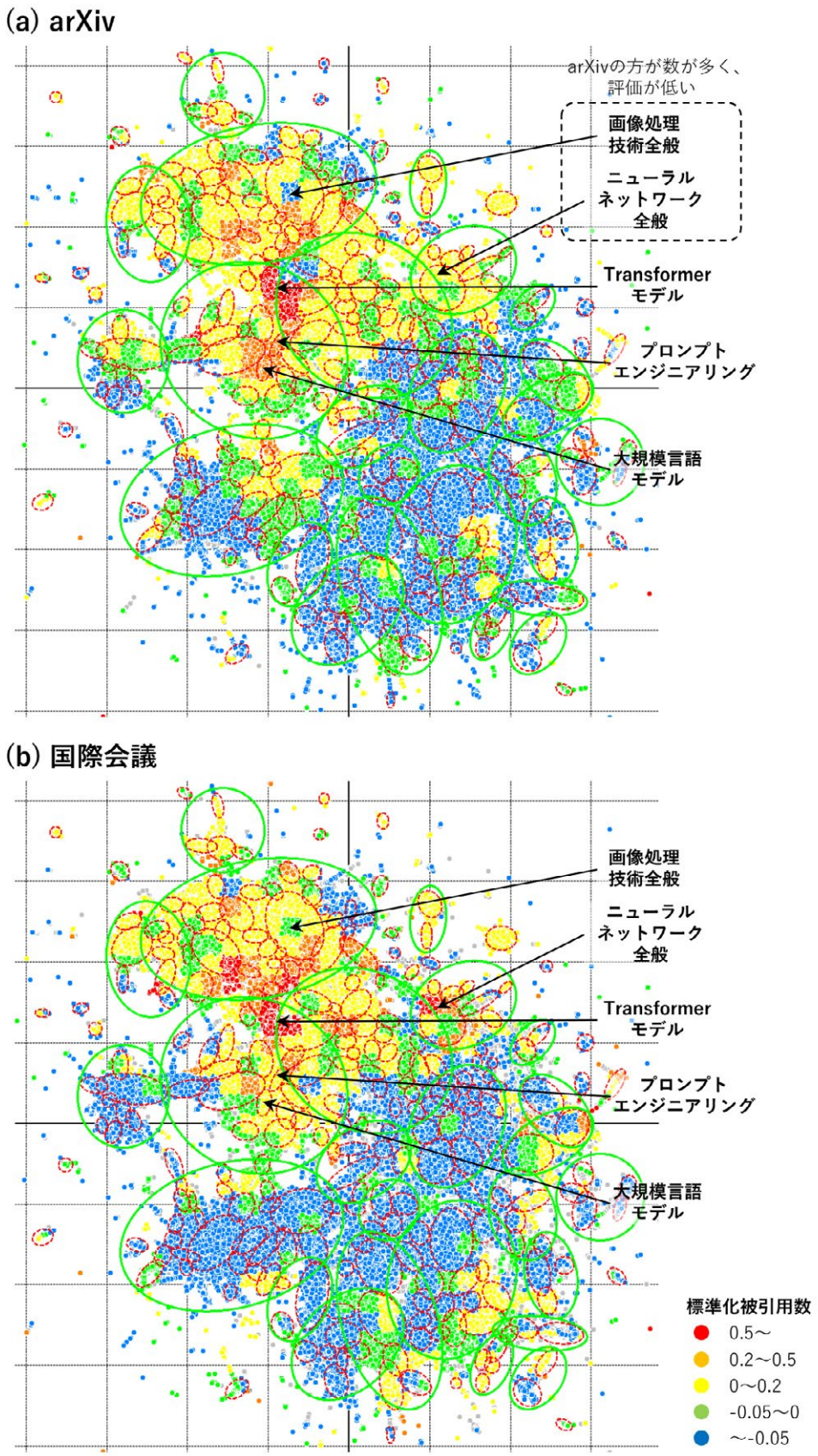


**図4-6　　標準化被引用数（arXivと国際会議の比較）**

# 5 ｜ エマージング領域の同定

## 5.1 国際会議への移行状況

arXiv掲載後に、国際会議議事録に採録されるまでの年のずれを調べた。arXivデータと国際会議議事録データ（Scopus）が同一と判定するための条件として、「arXivと国際会議議事録の両方のデータに記載があり、DOI番号が一致するもの」という条件（判定の結果、同一の論文だと判断された文献数は5,284件）に加えて、以下の2種類の条件を使用した。いずれにせよ、当該30国際会議に掲載された発表の20～30%はarXivに掲載されていることが分かる。

**条件A** arXivの「journal_reference」もしくは「arxiv_comment」に当該30学会のいずれかへの発表の記載があり、かつその記載学会名（1つ）データにタイトルが一致するデータがある文献（38,933件）。このとき、大文字・小文字、ハイフン・スペースの差異は無視した。先述のDOI番号一致とこの条件で同一と判定された論文は40,901件であった。これはarXiv全データ427,846件の約9.6%、国際会議議事録全データ200,663件の約20.4%に相当する。

**条件B** タイトルが一致する文献（61,988件）。このとき、大文字・小文字、ハイフン・スペースの差異は無視した。先述のDOI番号一致とこの条件で同一と判定された論文は62,398件であった。これはarXiv全データ427,846件の約14.6%、国際会議議事録全データ200,663件の約31.1%に相当する。

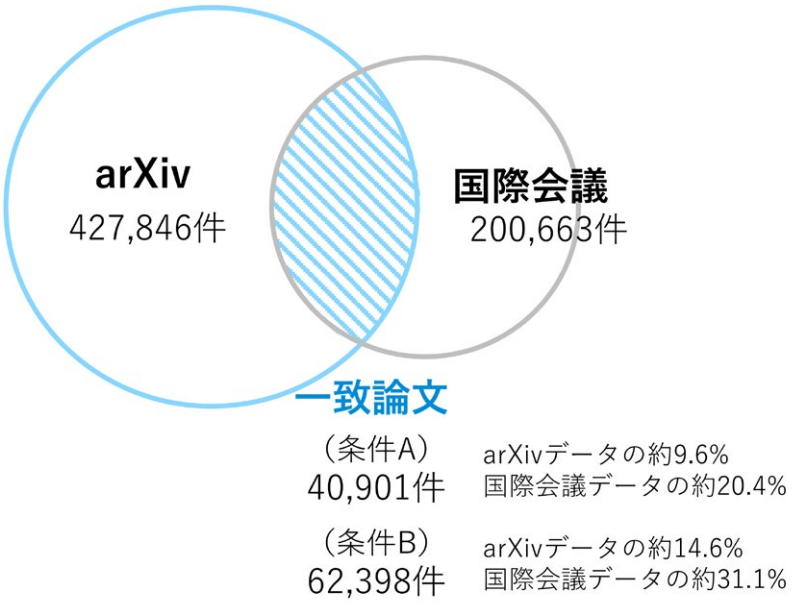


**図5-1　arXivと国際会議で同一と判定された論文**

arXiv収録年と国際会議議事録掲載年の差異の分布を、条件AとBの2通りについて図5-2に示した。一般的にはarXiv収録が先行すると予想されるが、データ上では国際会議議事録掲載後にarXiv収録されるケー

スも存在した（図5-2で差が負値のもの）。

このような国際会議発表後のarXiv登録は、セルフアーカイビングによるものと考えられる。セルフアーカイビングとは、研究者が自身の論文をオープンアクセスリポジトリに公開し、広くアクセス可能にする行為である。これにより研究成果の迅速な普及と共有が促進され、学術コミュニティの発展に寄与する。

図5-2には差の平均値として、arXivと国際会議議事録の前後を問わない全体の平均値と、arXivが先行するケースのみについての平均値（つまり、非負値の平均値）の2通りで求めた値を示した。

条件AとBのいずれの場合でも、arXivは国際会議に対して平均0.2～0.3年の先行性があり、ほとんどの場合1年以内であることが分かった。俯瞰図上で数年の差異として現れるのは、ある程度の件数が蓄積して初めて顕在化するためである。つまり、個別の論文では1年未満の差異であっても、俯瞰図上では数年の差として現れる点に注意が必要である。

文献情報データベースのユーザーの視点からは、プレプリント（arXiv）と国際会議プロシーディングス（Scopus）の掲載年の差は、前述の1年未満の差に加え、Scopusデータベースへの収録時間が影響する。Scopusは手作業による登録プロセスを含むため、即時収録ではない。具体的には、訴求登録や出版年月日の付け替え（早期公開日での初期登録後、刊行時に最終出版日へ変更）が行われる。

この登録プロセスにより、Scopusへの収録には遅れが生じる。一般的に、論文発表からScopusデータベースへの正式収録までは半年未満と推定される。一方、arXivに登録された論文はより早くアクセス可能となり、最新の研究成果を迅速に共有する点で有用である。

したがって、ユーザーは両データベースの特性を理解し、最新の研究動向を把握するために適切に使い分ける必要がある。arXivは速報性に優れ、Scopusは体系的な収録を行うという特徴を踏まえ、研究目的に応じて効果的に活用することが求められる。

arXivに収録されたデータのうち、論文や学会発表に関する記載情報を含むものは、そうでないものと比較して、より重要な内容を含んでいる可能性が高い。この記載情報の有無は、標準化被引用数と同様に、論文の重要性を示す指標となることが期待される。

この仮説に基づき、本調査では条件A、Bに加えて、新たに条件Cを設定して分析を行った。条件Cで抽出された論文は217,155件に達し、これはarXivに収録された文献の半数以上を占めている。結果は、arXivに掲載される文献の多くが、何らかの形で正式な論文掲載や学会発表がなされていることを示している。このことから、arXivが単なるプレプリントサーバーではなく、査読済みの研究成果を含む重要な学術情報源となっていることが分かる。

**条件C** arXivの「journal_reference」もしくは「arxiv_comment」に当該30学会に限らず、何らかの学会・雑誌への発表・掲載記述がある文献（217,155件）。このとき、大文字・小文字、ハイフン・スペースの差異は無視した。先述のDOI番号一致とこの条件で同一と判定された論文は234,104件で、これはarXiv全データ427,846件の約54.7%に相当する。

30の主要国際会議への掲載情報を持つarXiv文献が、優れた研究の指標となり、被引用数と同等の評価基準になるという仮説を検証した。図5-3(a)には、DOI情報から掲載が確認できる、または掲載情報とタイトルが一致するarXiv記事（条件A）の分布を標準化して示した。比較のため、右図に標準化された被引用数の分布を示している。

DOI一致または、掲載情報かつタイトルが一致したarXiv文献（40,901件）の位置を割合でプロット（左図）すると、被引用の高い領域（右図）とやや近い分布となった。特に、顕著性検出やTransformerモデルの分野では良く一致している。一方、敵対生成ネットワークの分野では一致度が低く、大規模言語モデルに関しては、まだ国際会議議事録に掲載されていない可能性があり、著しい不一致が見られた。

この結果から、発表からある程度時間が経過した著名な30国際会議に掲載情報のあるarXiv論文は、有

用な技術情報を含む可能性が高いと考えられる。ただし、国際会議の掲載情報には時間差があるため、最新の情報では不一致が生じる可能性があることに注意が必要である。

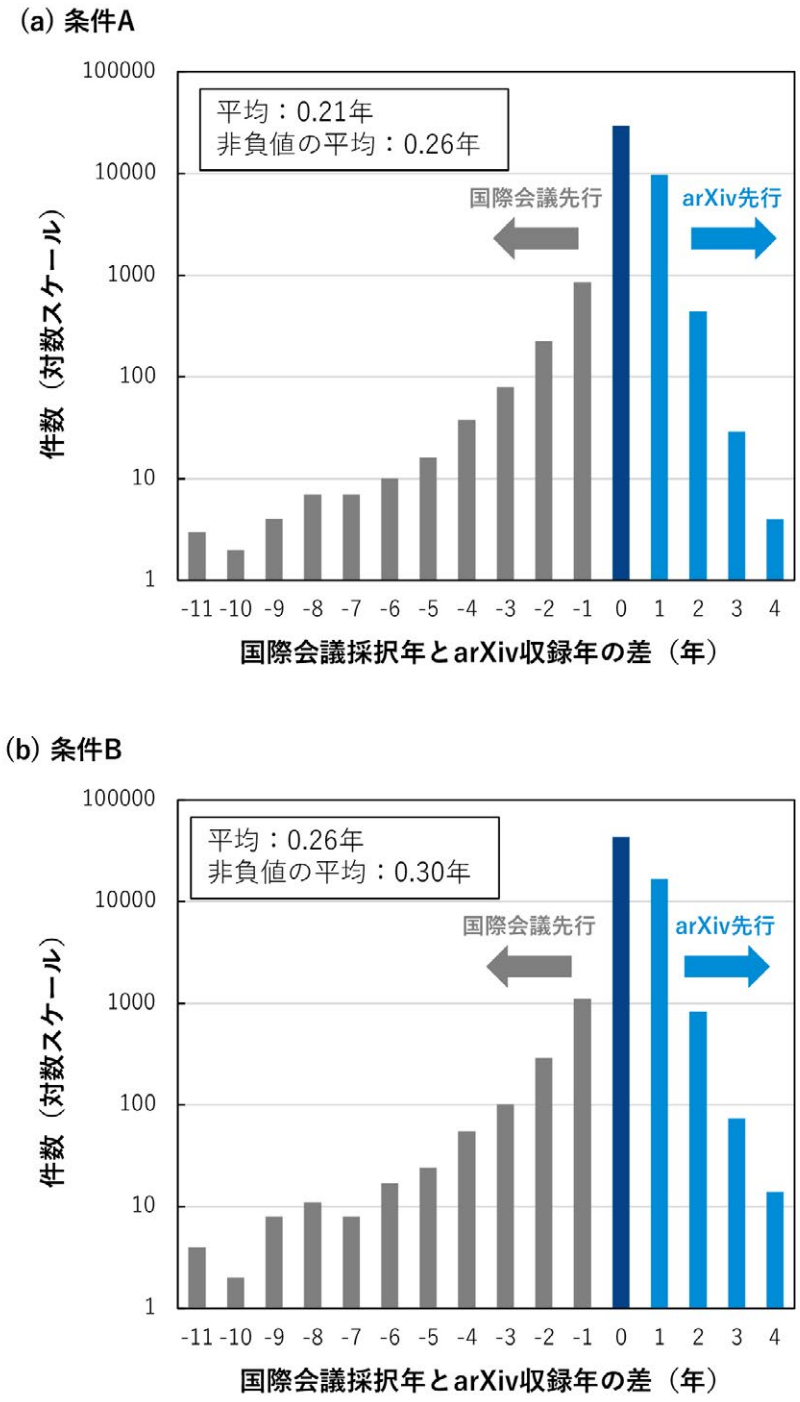


**図5-2　　国際会議採択年とarXiv掲載年の差の分布**

図5-3(b)はDOI情報から掲載が確認できるか、タイトルが一致するarXiv記事（条件B）の分布を標準化して示している。右図には比較対象として標準化された被引用数の分布を示した。

DOI一致または、タイトル一致したarXiv文献（62,398件）の位置を割合でプロット（左図）すると、図5-3(a)の掲載情報かつタイトル一致（条件A）の場合と比べ、被引用数の高い領域との一致度が低下した。

Transformerモデルに関しては一致度が向上した一方、敵対的生成ネットワークでは一致度が悪化した。

さらに、30の主要国際会議に限定せず、何らかの掲載情報を持つarXiv文献も優れた研究の指標となるという仮説を検証した。図5-3(c)の左図は、DOI情報、タイトルの一致、または何らかの掲載情報を持つarXiv記事（条件C）の分布を標準化して示している。右図には比較用の標準化された被引用数の分布を示した。条件Cに該当するarXiv文献（234,104件）に限定すると、被引用数の高い領域との相関がほぼ消失したように見える。

これらの結果を総合すると、標準化された被引用数で見た俯瞰マップとの相関の強さは、「30主要国際会議への掲載情報のあるarXiv > 30主要国際会議の議事録(Scopus)≧何らかの掲載情報のあるarXiv」という順序になることが明らかになった。この分析は、arXiv文献の価値を評価する際に、掲載情報の種類や範囲が重要な要素となることを示唆している。特に、主要国際会議への掲載情報を持つarXiv文献が、高い研究価値を示す指標として有効である可能性を示している。

## 5.2 引用回数の変化

被引用情報の分析において、標準化した俯瞰図上のメッシュ化マップを用いて、1回目と2回目に取得したデータを比較した。

図5-4(a)はarXivのみのデータで標準化した結果を示している。1回目と2回目の比較では、被引用の高低パターンは概ね一致しているが、大規模言語モデル、Transformerモデル、敵対的生成ネットワークなどの分野で標準化被引用数が増加し、色が橙色から赤へと変化している。この結果は、arXivでは約80日の取得日差でも目視可能な変化が反映されることを示している。

図5-4(b)は国際会議議事録のみのデータで標準化した結果を示している。1回目と2回目の比較では、ほとんど変化が見られない。arXivデータで観察されたような顕著な変化は確認できなかった。

図5-4(c)はarXivと国際会議議事録の両方のデータを合わせて標準化した結果を示している。 arXivでの変化と国際会議議事録の安定性が組み合わさり、arXivでの変化が緩和されて現れている。敵対生成ネットワークや大規模言語モデルで橙色から赤への変化が確認できるが、その程度はarXivのみの場合より小さい。

arXivと国際会議議事録は、大まかな注目領域で共通点を持つ（赤楕円の注目領域）。ただし、arXivでは特に大規模言語モデルの増加が顕著であるのに対し、国際会議議事録ではTransformerモデルや敵対生成ネットワークなど、やや技術的に遅れた領域で増加が見られる。

これらの分析結果から、arXivデータの先行性が一定程度確認できた。しかし、直近80日の変化と2010年以降の集計を比較すると、注目領域に大きな差異はないことも明らかになった。この知見は、短期的な変化と長期的なトレンドの関係を理解する上で重要な示唆を提供している。

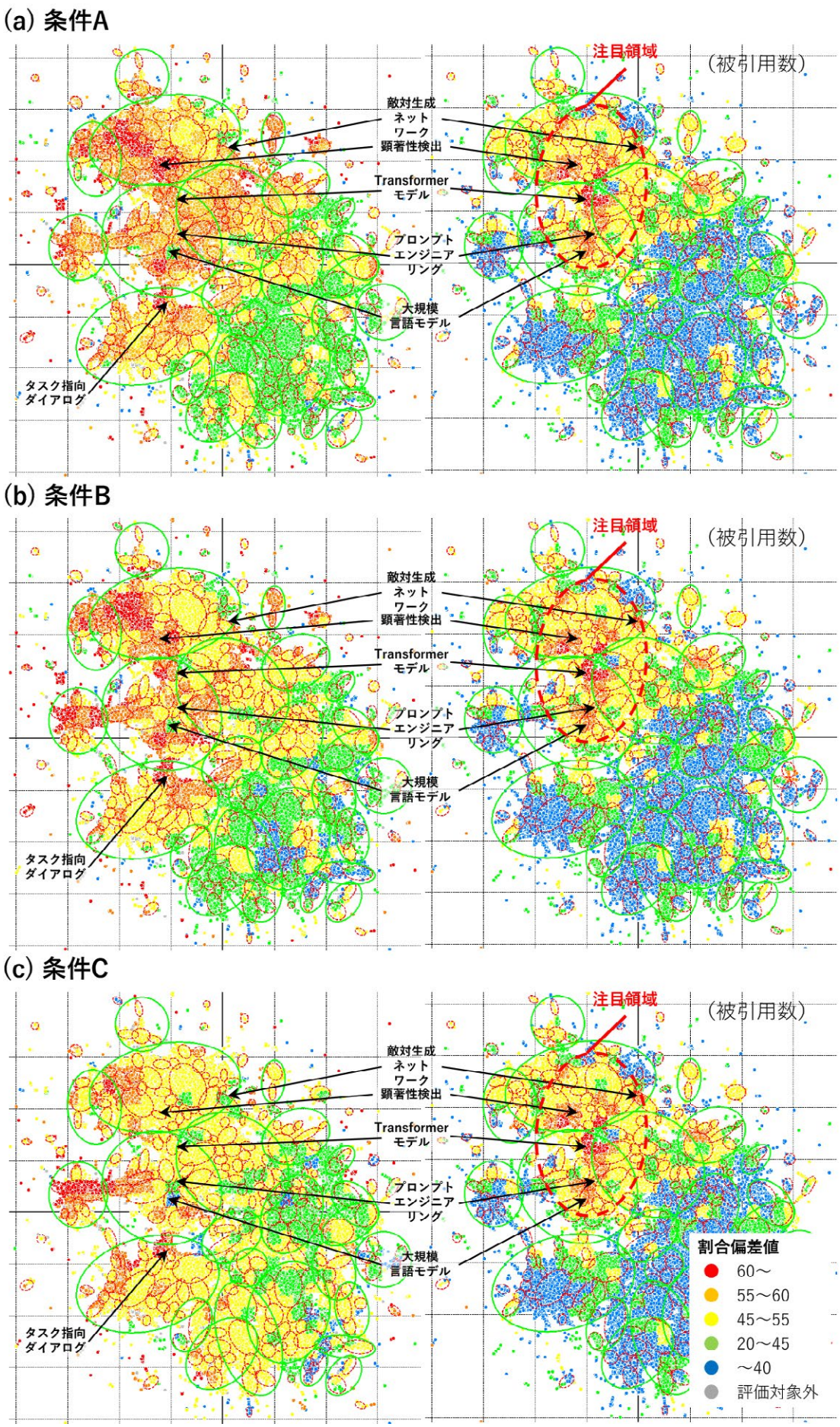


**図5-3　　国際会議採択年とarXiv掲載年の差の分布**

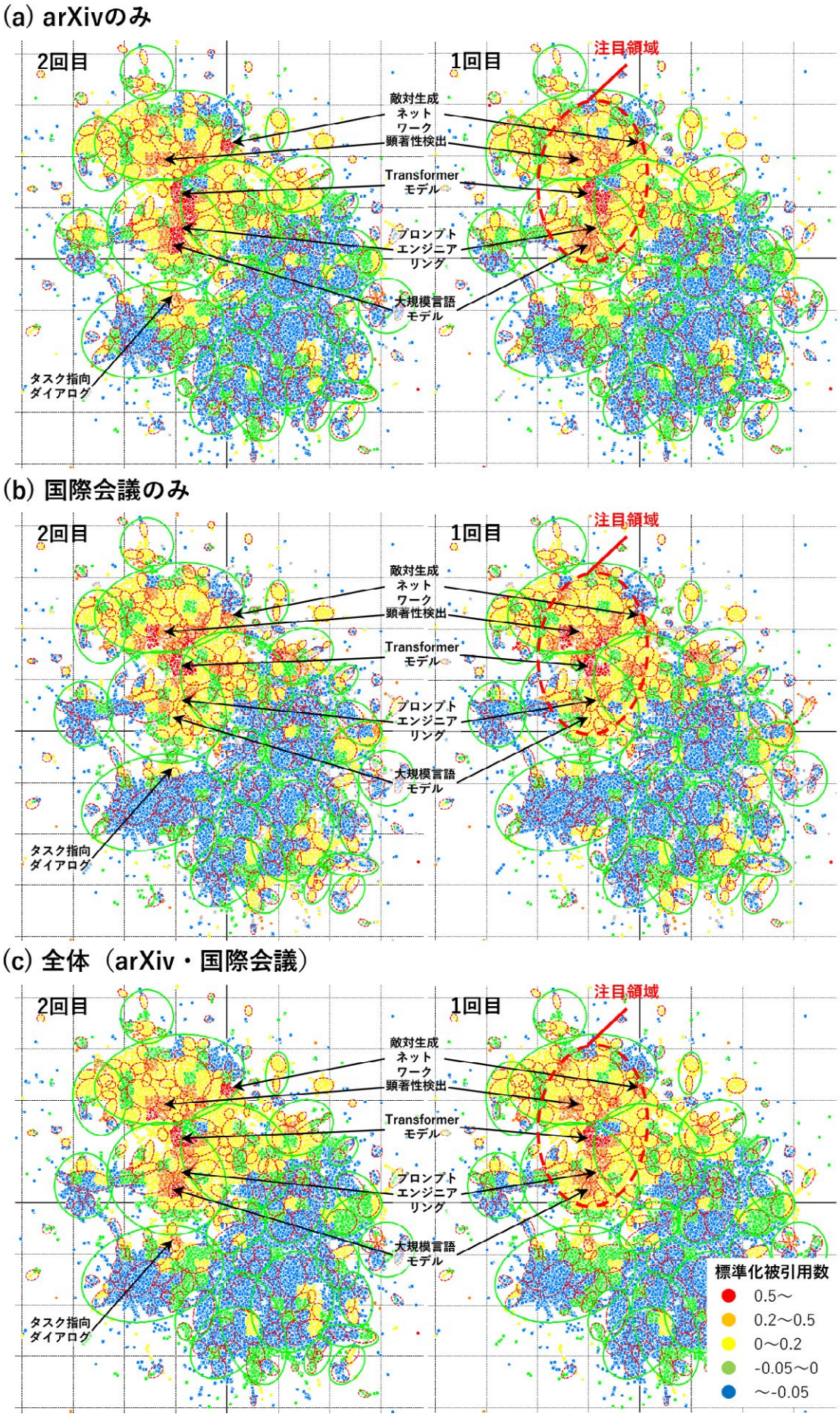


図5-4　　標準化被引用回数マップの比較

引用は積算過程的な現象として知られている。すなわち、ある時点で被引用数が多い文献は、その多寡に応じてさらに引用される傾向がある。このモデルに従うと、被引用数の時間発展は対数正規分布を示すことになる。

図5-5は、arXiv、国際会議議事録、およびその合計（全体）について、1回目に取得した被引用数と2回目に取得したときの増分の関係を示している。いずれの場合も、2回目の増分が1回目の被引用数にほぼ比例して増加する傾向が見られる。この傾向を1～3次関数でフィットした結果はほぼ同じで、いずれも1次関数的なプロットとなっている。フィット結果の詳細（誤差、決定係数など）は表5-1にまとめた。

当初の予想では、1～3次関数フィットの中で1次のものが最も高い決定係数（R^2）または自由度補正決定係数を示すと期待していた。しかし実際には、arXiv、国際会議議事録、全体のいずれの場合も、決定係数は0.15～0.16程度の低い値にとどまり、明確な判定はできなかった。この結果は、データにノイズが多く、フィットの精度が低いことが原因かもしれない。ただし、1次～3次フィットのいずれの場合でも1次係数の誤差が最も小さく、特に1次フィットの場合の1次係数の誤差は0.5%未満と最小であった。このことから、被引用数の増加傾向において1次の項が支配的であると考えられる。

この分析結果は、引用の積算過程的性質を裏付けるものであり、学術文献の影響力の時間的発展を理解する上で重要な知見を提供している。

**表5-1　　フィッティング係数**

| | 1次式 | 2次式 | 3次式 |
|---|---|---|---|
| 全体 | | | |
| 定数 | 6.50 E-01 | 9.65 E-01 | 1.33 E+00 |
| 誤差 (%) | 9.72 E+00 | 6.57 E+00 | 4.77 E+00 |
| 1次係数 | 4.59 E-02 | 3.60 E-02 | 2.35 E-02 |
| 誤差 (%) | 2.98 E-01 | 6.57 E-01 | 1.42 E+00 |
| 2次係数 | | 1.37 E-07 | 7.53 E-07 |
| 誤差 (%) | | 1.97 E+00 | 1.59 E+00 |
| 3次係数 | | | -4.11 E-12 |
| 誤差 (%) | | | -1.90 E+00 |
| R^2 | 0.15210235 | 0.15556876 | 0.15928148 |
| 補正R^2 | 0.152101 | 0.15556607 | 0.15927746 |
| arXiv | | | |
| 定数 | 7.73 E-01 | 9.81 E-01 | 1.22 E+00 |
| 誤差 (%) | 8.89 E+00 | 7.03 E+00 | 5.68 E+00 |
| 1次係数 | 3.98 E-02 | 3.31 E-02 | 2.47 E-02 |
| 誤差 (%) | 3.67 E-01 | 7.77 E-01 | 1.48 E+00 |
| 2次係数 | | 8.84 E-08 | 4.90 E-07 |
| 誤差 (%) | | 3.16 E+00 | 2.60 E+00 |
| 3次係数 | | | -2.60 E-12 |
| 誤差 (%) | | | -3.10 E+00 |
| R^2 | 0.14772647 | 0.14972256 | 0.15178427 |
| 補正R^2 | 0.14772447 | 0.14971859 | 0.15177832 |

| 国際会議 | | | |
|---|---|---|---|
| 定数 | 2.84 E-01 | 1.16 E+00 | 1.19 E+00 |
| 誤差（%） | 4.65 E+01 | 1.14 E+01 | 1.11 E+01 |
| 1次係数 | 6.03 E-02 | 3.48 E-02 | 3.38 E-02 |
| 誤差（%） | 4.91 E-01 | 1.44 E+00 | 2.26 E+00 |
| 2次係数 | | 4.13 E-07 | 4.72 E-07 |
| 誤差（%） | | 1.60 E+00 | 7.04 E+00 |
| 3次係数 | | | -4.54 E-13 |
| 誤差（%） | | | -5.49 E+01 |
| R^2 | 0.17132352 | 0.18705423 | 0.18706766 |
| 補正R^2 | 0.17131939 | 0.18704613 | 0.1870555 |

1次フィットによる期待値を上回る被引用数の増加を示す文献は、現在も注目される優れた研究であると考えられる。図5-6は、特に顕著な増加を示す文献を示している。青字はarXiv掲載、灰字は国際会議議事録収録を表し、各文献の研究内容と掲載年を付記した。著名な技術については技術名を、そうでない場合は研究タイトルを記載している。

全体的な傾向として、掲載年が古いほど横軸の値（1回目の被引用数）が大きくなるため、国際会議議事録収録の研究はグラフ右側に集中している。図5-6(a)の線形表示では、1回目の被引用数と掲載年の古さにほぼ比例関係が見られる。被引用数が多い順に、ResNet(2016)、Attention(2017:arXiv,国際会議議事録)、BERT(2018:arXiv,2019:国際会議議事録)、U-net(2017)、GPT-3(2020:arXiv,国際会議議事録)といった著名な技術が確認できる。Attention、BERT、GPT-3については、arXivと国際会議議事録の両方に収録があるが、arXivの方が早く、また件数も多い。

両対数表示（図5-6(b)）では、1回目被引用数の小さい範囲、つまり比較的新しい研究が多いと予想される領域では、ほぼarXiv掲載の研究に限定される。この範囲で件数の多いものとして、PyTorch(2019)、HuggingFace Transformers(2019)、DALL-E(2021)などの著名な研究が挙げられる。ただし、件数が少なくなると統計的揺らぎの影響か、横軸と年代順の関係が不明確になり、上振れの点が必ずしも優れた研究を示すとは限らない。

図5-7は、これらの文献の俯瞰マップ上での位置を示している。多くの文献が注目領域の周辺に分布していることが分かる。

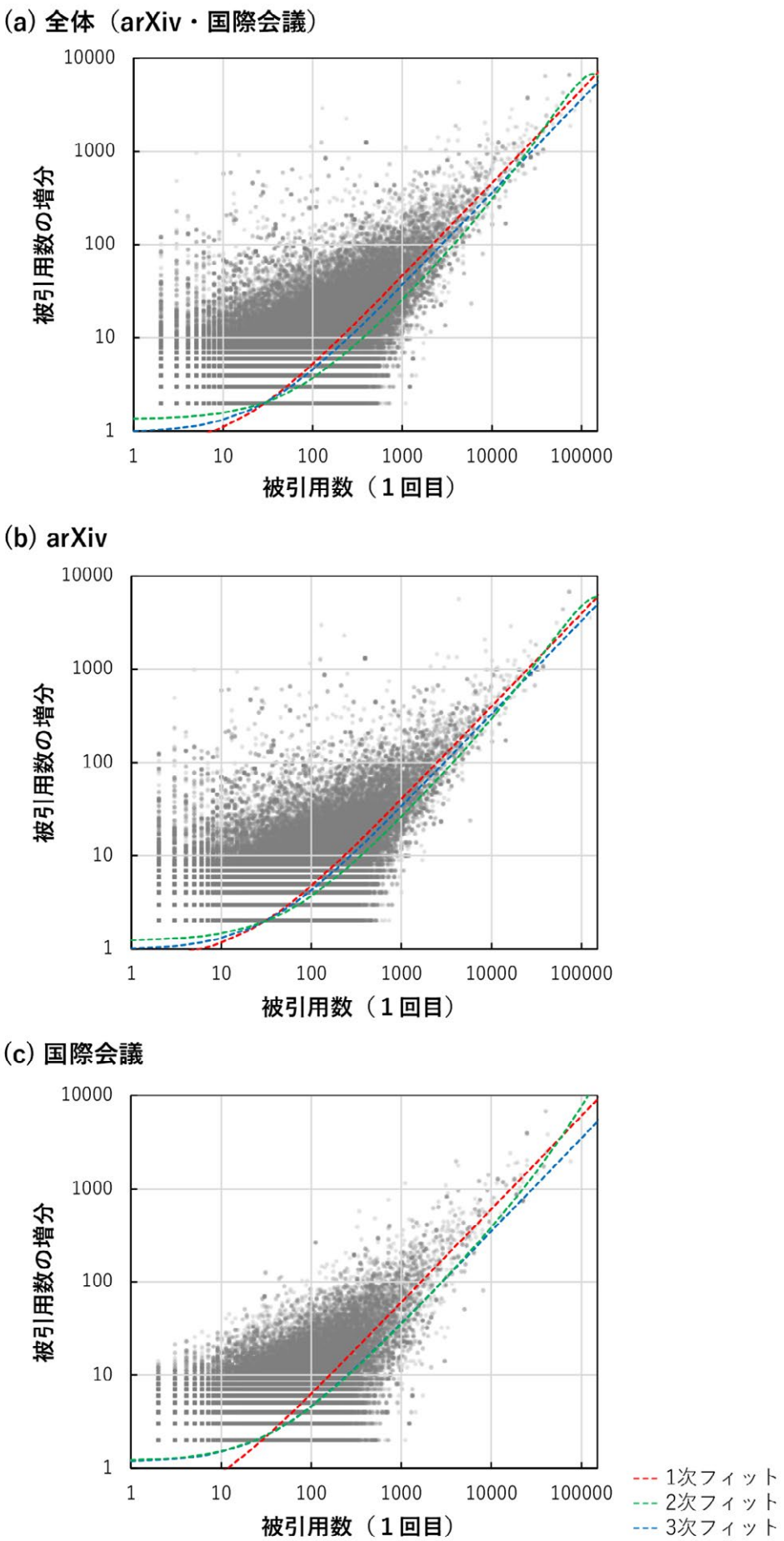


図5-5　被引用数とその増分の比較

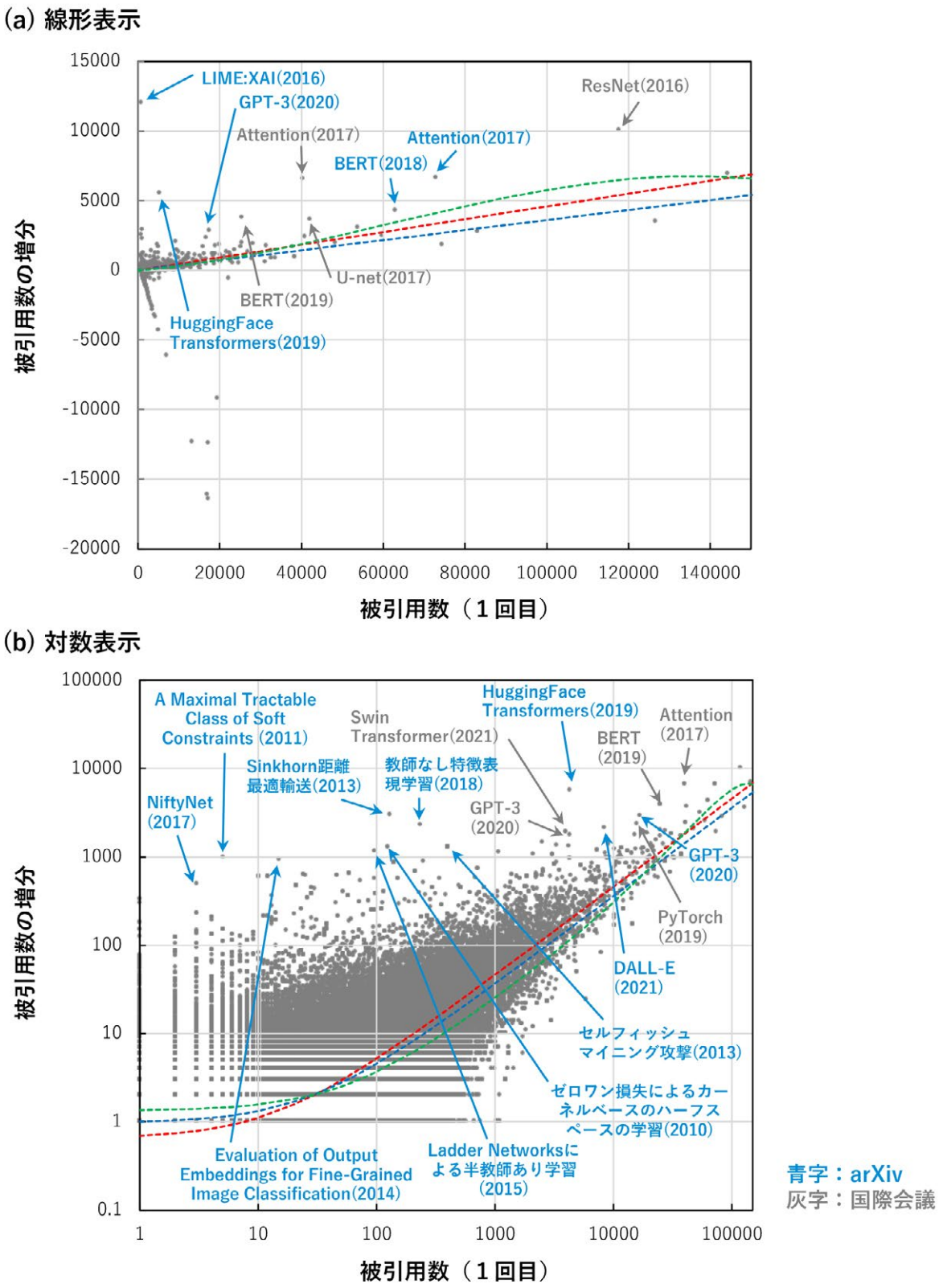


図 5-6　　期待値以上に被引用数が増加している文献

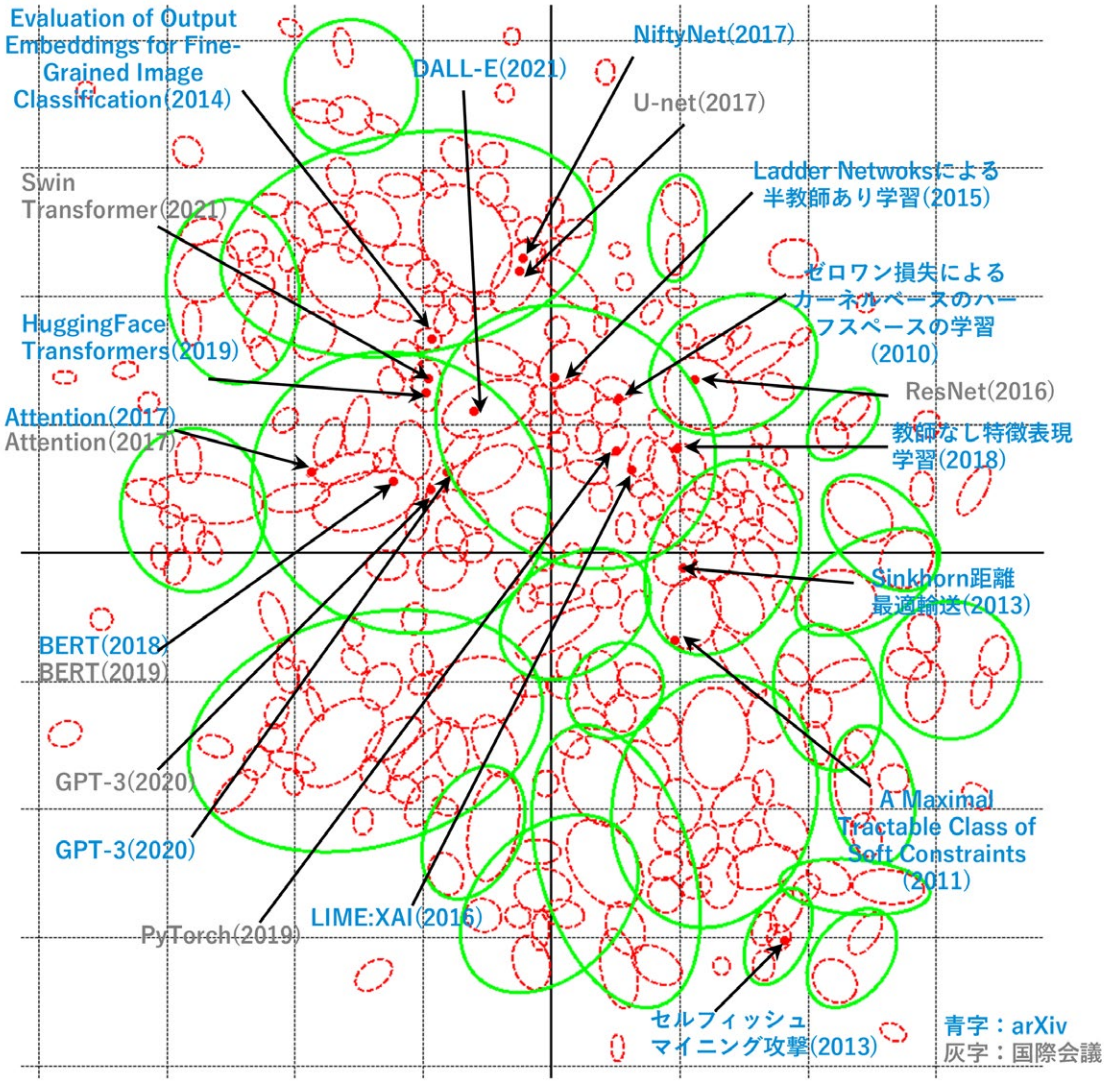


**図5-7　期待値以上に被引用数が増加している文献の俯瞰マップ上の位置**

## 5.3 キーフレーズレベル速報性

arXivの掲載情報と国際会議議事録を比較した結果、arXivの速報性は最大で1年程度であることが確認された。ただし、規模と認知度を考慮すると、実質的な影響は数年に及ぶ可能性がある。この知見を踏まえ、さらにキーフレーズレベルでの速報性判定を実施した。

分析対象のアブストラクトから、RAKE（Rapid Automatic Keyword Extraction）を用いてキーフレーズを抽出した。出現回数がおよそ1000件以上の上位80キーフレーズを表5-2に示す。

1位の「ソースコード」はノイズ的である一方、2位「敵対的生成ネットワーク」、3位「グラフニューラルネットワーク」、4位「深層学習」、5位「データ拡張」など、注目度の高い技術がほぼ抽出できている。9位には「大規模言語モデル」が3000件以上と多く、10位には「自動運転」が約3000件で入っている。70位の「トランスフォーマーモデル」も1000件以上あり、1000件を超えるキーフレーズは速報性の検証に適していると判断できる。

一方、10件程度しか出現しないキーフレーズは最新のものである可能性が高いが、ノイズ的な表現も多く含まれる。そこで本調査では、これらの中間をとり、100件以上出現したキーフレーズを用いてarXivと国際会議議事録の速報性比較を行った。

この方法により、技術トレンドの早期把握と、それぞれの情報源の特性をより詳細に分析することが可能となる。また、この比較は、研究者や情報分析者が最新の技術動向を効率的に追跡する上で有用な指針を提供するものと期待される。

図5-8は、arXivと国際会議議事録の全628,509件から抽出した、100件以上出現した8,089キーフレーズのうち、両者で初出年が異なる4,612キーフレーズ（arXiv 1,304、国際会議議事録3,308）の分析結果を示している。(a) は初出年別件数、(b) は出現年における件数割合、(c) は出現年以降の累計件数に対する割合を表している。

これらのグラフから、2014年までは国際会議議事録でキーフレーズが先に出現する傾向が見られる。この時期は国際会議議事録の発表件数が多く、深層学習黎明期における先端技術の紹介の場として国際会議が重要な役割を果たしていたことがうかがえる。

2015年以降、arXivが先端技術発表の主要な場となり、優勢に転じている。arXivの件数が急増する一方、国際会議議事録の増加は緩やかである。この規模の差がarXivの先行性を過大評価させる可能性も考えられるが、図5-8(b) の出現年における件数割合でも2015年以降のarXiv優勢は維持されている。

本調査では、ある程度確立したキーフレーズを指標とするため、100件以上出現したものに限定した。この基準は初出年が古いキーフレーズには緩く、新しいキーフレーズには厳しい傾向がある。この影響を緩和するため、図5-8(c) では出現年以降の累計件数で正規化を行った。この補正後も、2014年頃に国際会議からarXivへ先行性が移行したという傾向は維持されている。

図5-9は、arXivと国際会議議事録におけるキーフレーズの先行性を年数で示している。文献の掲載年と同様、キーフレーズにおいても大多数は1年の先行であるが、少数ながら複数年先行する例も存在する。arXivがやや長期の先行性を示す傾向が見られるものの、先行年の平均はarXivが1.86年、国際会議議事録が2.20年とほぼ2年程度である。

表5-3は、2014年以前に国際会議が先行した主要なキーフレーズ（1,000件以上出現）を示している。初出年の昇順、合計件数の降順でソートし、各キーフレーズについてarXivと国際会議議事録別に初出年、合計件数、直近3年件数、ピーク年を記載している。

**表5-2　キーフレーズの例（上位80個）**

| No. | キーフレーズ | 件数 | No. | キーフレーズ | 件数 |
|---|---|---|---|---|---|
| 1 | source code | 6120 | 41 | synthetic datasets | 1292 |
| 2 | generative adversarial network | 4393 | 42 | semantic information | 1272 |
| 3 | graph neural network | 4315 | 43 | COVID19 pandemic | 1265 |
| 4 | deep learning model | 4159 | 44 | neural machine translation | 1259 |
| 5 | data augmentation | 3899 | 45 | sample complexity | 1255 |
| 6 | machine learning model | 3546 | 46 | fewshot learning | 1225 |
| 7 | adversarial attack | 3391 | 47 | word embeddings | 1223 |
| 8 | model train | 3345 | 48 | medical imaging | 1221 |
| 9 | large language model | 3039 | 49 | name entity recognition | 1191 |
| 10 | autonomous driving | 2931 | 50 | project page | 1190 |
| 11 | training process | 2911 | 51 | widely adopt | 1176 |
| 12 | deep reinforcement learning | 2861 | 52 | domain shift | 1174 |
| 13 | latent space | 2851 | 53 | domain knowledge | 1172 |
| 14 | public datasets | 2650 | 54 | augment reality | 1163 |
| 15 | knowledge graph | 2533 | 55 | edge device | 1136 |
| 16 | adversarial training | 2480 | 56 | video frame | 1126 |
| 17 | contrastive learning | 2234 | 57 | fully exploit | 1125 |
| 18 | pretrained language model | 2192 | 58 | neural architecture search | 1124 |
| 19 | anomaly detection | 2175 | 59 | imitation learning | 1120 |
| 20 | autonomous vehicle | 2132 | 60 | widely apply | 1118 |
| 21 | baseline model | 2130 | 61 | key component | 1116 |
| 22 | data distribution | 2103 | 62 | specific task | 1096 |
| 23 | selfsupervised learning | 2070 | 63 | training phase | 1095 |
| 24 | comprehensive experiment | 2059 | 64 | continual learning | 1089 |
| 25 | federated learning | 2005 | 65 | fundamental task | 1079 |
| 26 | source domain | 1969 | 66 | real image | 1078 |
| 27 | knowledge distillation | 1953 | 67 | open source | 1054 |
| 28 | inference time | 1704 | 68 | future direction | 1053 |
| 29 | reward function | 1614 | 69 | promising solution | 1053 |
| 30 | single model | 1537 | 70 | transformer model | 1053 |
| 31 | federate learning | 1500 | 71 | instance segmentation | 1051 |
| 32 | vision transformer | 1481 | 72 | improve accuracy | 1040 |
| 33 | learning | 1435 | 73 | deep generative model | 1039 |
| 34 | training dataset | 1422 | 74 | task due | 1037 |
| 35 | diffusion model | 1405 | 75 | target task | 1027 |
| 36 | model architecture | 1405 | 76 | classification model | 1017 |
| 37 | graph convolutional network | 1392 | 77 | deep learningbased | 1000 |
| 38 | link prediction | 1382 | 78 | privacy concern | 996 |
| 39 | task | 1347 | 79 | computational resource | 993 |
| 40 | NLP task | 1318 | 80 | variational autoencoder | 992 |

2010年が初出年となっているキーフレーズが多いが、これは集計開始年によるもので、実際の初出はより早い可能性がある。そのため、国際会議議事録の先行年数は表示値より大きい可能性がある。2014年以前の国際会議議事録先行キーフレーズの半数以上が見かけ上2010年初出となっている。

深層学習黎明期には、「畳み込みニューラルネットワーク」や「学習済みモデル」など、多くのキーフレーズが国際会議議事録で先行して発表されている。「自律運転」「モバイルロボット」「拡散モデル」なども早期に出現している。

2011年以降の国際会議議事録先行キーフレーズには、「拡張現実」（2010年）、「アテンション機構」（2011年）、「大規模言語モデル」（2011年）などがある。2014年は国際会議議事録先行のキーフレーズがなく、先導交代の節目となっている。2015年以降、国際会議議事録先行のキーフレーズは「対照学習」など3つのみで、いずれも2015年初出である。これ以降、国際会議議事録が先行するケースは見られない。

表5-4は、arXivが先行した主要なキーフレーズ（1,000件以上出現）を示している。2014年以前にarXivが先行したキーフレーズを初出年の昇順、合計件数の降順でソートし、各キーフレーズについてarXivと国際会議議事録別に初出年、合計件数、直近3年件数、ピーク年を記載している。

2014年以前でも、少数ながらarXiv先行のキーフレーズが存在する。2010年初出では「データ拡張」「敵対的攻撃」など、国際会議議事録でも類似のキーフレーズがあったものが中心である。2014年にはarXiv初出のキーフレーズが増加し、この年からarXivの先導的役割が顕著になり始める。「知識蒸留」などがこの時期のarXiv初出キーフレーズの例である。

2015年以降は明らかにarXivが先行している。「グラフニューラルネットワーク」「変分オートエンコーダー」「事前自己学習」など、比較的新しい深層学習技術が多く含まれる。特筆すべきは「連合学習」で、3年もの先行期間があり、俯瞰図レベルで確認できるほど顕著な先行性を示している。

その他の注目すべきarXiv先行キーフレーズには、「デジタル・ツイン」、「説明可能AI」（2017年）、「トランスフォーマーモデル」（2017年）がある。2020年の「基盤モデル」も当然ながらarXiv先行となっている。

本分析結果は、2014年以降のarXivの重要性の急速な増大を明確に示している。 arXivは新たな深層学習技術や概念の初発表の場となり、学術コミュニティにおける情報伝播の中心的役割を担うようになった。この傾向は、研究者や技術者が最新の技術動向を把握する上で、arXivの重要性が高まっていることを示唆している。

同時に、この分析は深層学習分野における技術トレンドの変遷と情報発信の場の移行を明確に示している。特に、国際会議からarXivへと主要な情報源が移行していく過程が浮き彫りになっている。これにより、各情報源の特性や役割の変化を理解する上で重要な知見が得られた。

さらに、この分析は技術トレンドの変遷を追跡する上でも有用な情報を提供している。キーフレーズの出現時期や頻度の変化を通じて、深層学習分野の発展の歴史を概観することができる。これは、分野の過去の軌跡を理解し、将来の方向性を予測する上で貴重な資料となる。

これらの知見は、研究動向の把握や最新技術の追跡において、情報源の選択と評価に有用な指針を提供する。特に、arXivが近年の先端技術情報の重要な源泉となっていることが裏付けられた。この結果は、研究者や技術者が効率的に情報収集を行う上で、arXivの重要性を再認識させるものである。

総じて、この分析は深層学習分野の情報ダイナミクスの変化を包括的に捉え、学術情報の流通と技術進化の関係性に新たな視点を提供している。これは、今後の研究戦略や技術開発の方向性を考える上で、重要な示唆を与えるものである。

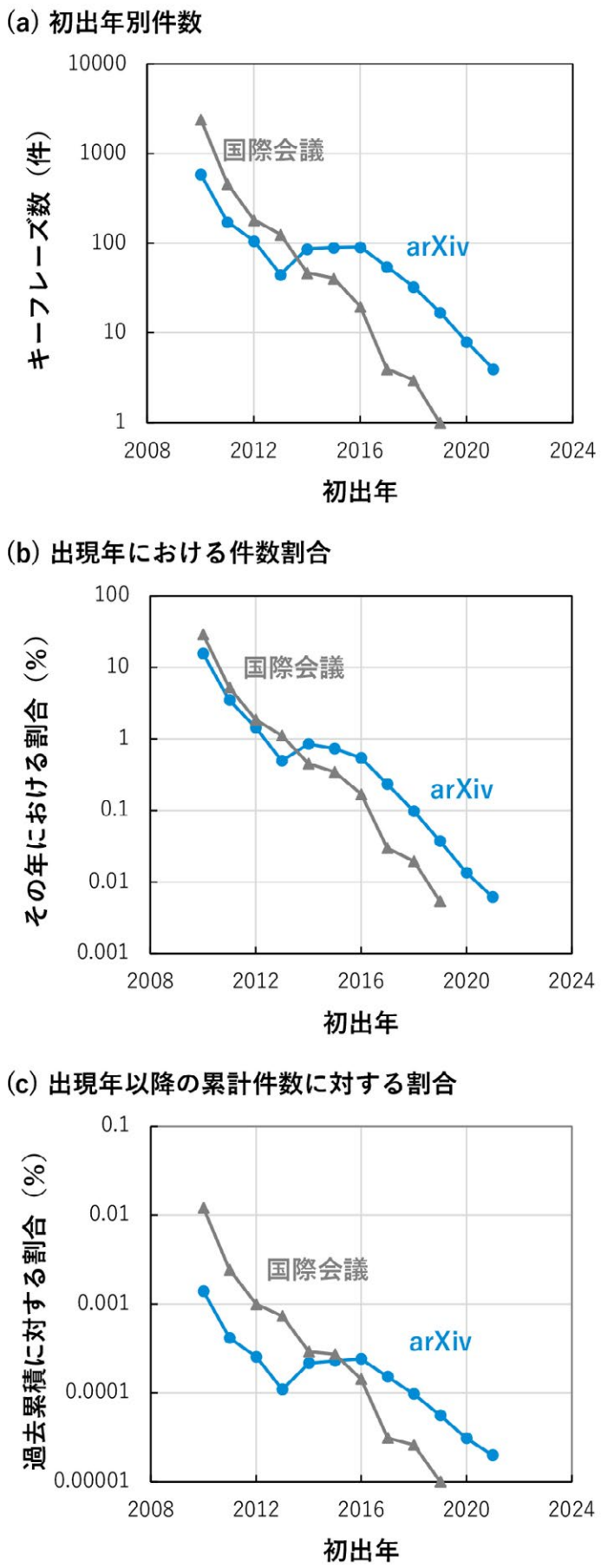


**図5-8　　キーフレーズ出現年の比較**

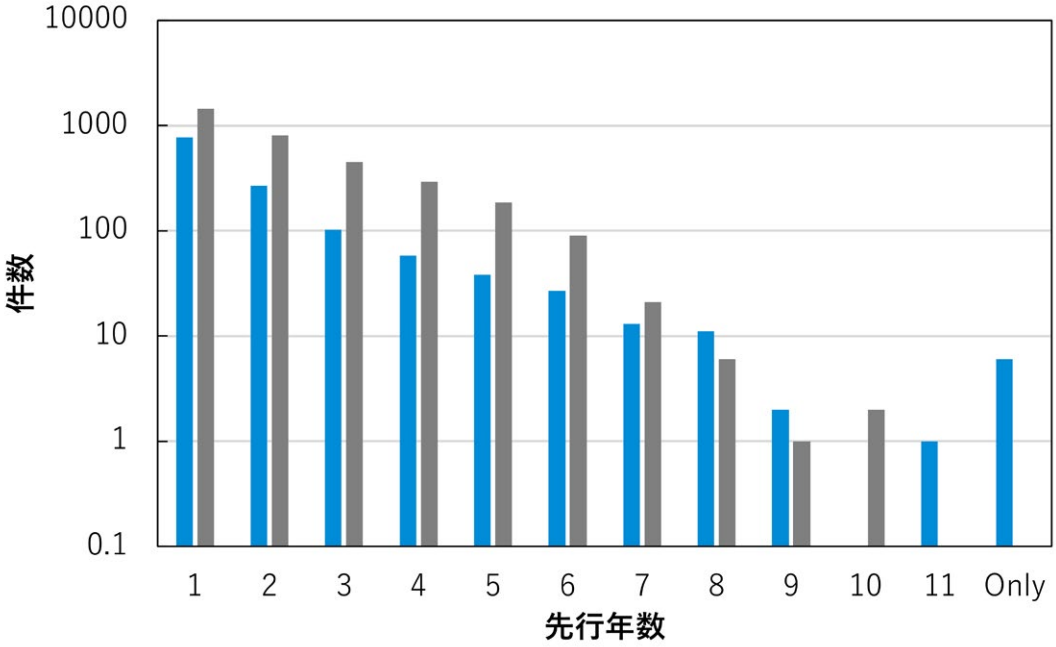


**図 5-9　　キーフレーズレベル先行性**

## 表5-3　国際会議が先行したキーフレーズ

| No. | キーフレーズ | 初出年 | | | 件数 | | | 直近3年の件数 | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | arXiv | 国際会議 | 差 | arXiv | 国際会議 | All | arXiv | 国際会議 | All |
| 1 | convolutional neural network | 2012 | 2010 | 2 | 8197 | 2757 | 10954 | 3184 | 690 | 3874 |
| 2 | benchmark datasets | 2011 | 2010 | 1 | 4439 | 2456 | 6895 | 2501 | 1112 | 3613 |
| 3 | downstream task | 2014 | 2010 | 4 | 3574 | 1289 | 4863 | 2889 | 995 | 3884 |
| 4 | pretrained model | 2015 | 2010 | 5 | 2909 | 1097 | 4006 | 2164 | 787 | 2951 |
| 5 | target domain | 2011 | 2010 | 1 | 2544 | 1434 | 3978 | 1460 | 645 | 2105 |
| 6 | machine learning model | 2011 | 2010 | 1 | 2997 | 549 | 3546 | 1863 | 280 | 2143 |
| 7 | semantic segmentation | 2012 | 2010 | 2 | 2405 | 1020 | 3425 | 1338 | 482 | 1820 |
| 8 | autonomous driving | 2014 | 2010 | 4 | 2199 | 732 | 2931 | 1457 | 426 | 1883 |
| 9 | point cloud | 2011 | 2010 | 1 | 1942 | 938 | 2880 | 1200 | 490 | 1690 |
| 10 | latent space | 2011 | 2010 | 1 | 2024 | 827 | 2851 | 1202 | 439 | 1641 |
| 11 | public datasets | 2011 | 2010 | 1 | 1727 | 923 | 2650 | 1102 | 462 | 1564 |
| 12 | strong baseline | 2012 | 2010 | 2 | 1375 | 845 | 2220 | 810 | 405 | 1215 |
| 13 | machine translation | 2011 | 2010 | 1 | 1409 | 809 | 2218 | 633 | 218 | 851 |
| 14 | autonomous vehicle | 2011 | 2010 | 1 | 1639 | 493 | 2132 | 957 | 237 | 1194 |
| 15 | baseline model | 2012 | 2010 | 2 | 1514 | 616 | 2130 | 934 | 303 | 1237 |
| 16 | data distribution | 2011 | 2010 | 1 | 1558 | 545 | 2103 | 939 | 264 | 1203 |
| 17 | comprehensive experiment | 2012 | 2010 | 2 | 1347 | 712 | 2059 | 921 | 381 | 1302 |
| 18 | model outperforms | 2011 | 2010 | 1 | 1262 | 767 | 2029 | 636 | 280 | 916 |
| 19 | source domain | 2012 | 2010 | 2 | 1269 | 700 | 1969 | 720 | 315 | 1035 |
| 20 | model training | 2015 | 2010 | 5 | 1392 | 487 | 1879 | 980 | 277 | 1257 |
| 21 | term of accuracy | 2011 | 2010 | 1 | 1315 | 516 | 1831 | 703 | 182 | 885 |
| 22 | question answering | 2011 | 2010 | 1 | 1202 | 618 | 1820 | 663 | 238 | 901 |
| 23 | inference time | 2013 | 2010 | 3 | 1233 | 471 | 1704 | 827 | 281 | 1108 |
| 24 | mobile robot | 2011 | 2010 | 1 | 696 | 741 | 1437 | 349 | 137 | 486 |
| 25 | model size | 2012 | 2010 | 2 | 1040 | 394 | 1434 | 703 | 219 | 922 |
| 26 | training dataset | 2011 | 2010 | 1 | 1085 | 337 | 1422 | 640 | 158 | 798 |
| 27 | learn model | 2011 | 2010 | 1 | 938 | 481 | 1419 | 447 | 172 | 619 |
| 28 | diffusion model | 2011 | 2010 | 1 | 1303 | 102 | 1405 | 1270 | 81 | 1351 |
| 29 | feature map | 2011 | 2010 | 1 | 921 | 463 | 1384 | 454 | 185 | 639 |
| 30 | user experience | 2011 | 2010 | 1 | 722 | 640 | 1362 | 401 | 179 | 580 |
| 31 | trained model | 2012 | 2010 | 2 | 956 | 292 | 1248 | 612 | 170 | 782 |
| 32 | action recognition | 2011 | 2010 | 1 | 722 | 513 | 1235 | 287 | 126 | 413 |
| 33 | pose estimation | 2011 | 2010 | 1 | 706 | 526 | 1232 | 378 | 153 | 531 |
| 34 | unsupervised manner | 2011 | 2010 | 1 | 756 | 452 | 1208 | 371 | 155 | 526 |
| 35 | name entity recognition | 2011 | 2010 | 1 | 817 | 374 | 1191 | 448 | 146 | 594 |
| 36 | domain shift | 2011 | 2010 | 1 | 813 | 361 | 1174 | 585 | 231 | 816 |
| 37 | augment reality | 2011 | 2010 | 1 | 744 | 419 | 1163 | 415 | 152 | 567 |
| 38 | bounding box | 2013 | 2010 | 3 | 749 | 395 | 1144 | 367 | 152 | 519 |
| 39 | fundamental task | 2011 | 2010 | 1 | 727 | 352 | 1079 | 435 | 161 | 596 |
| 40 | target object | 2011 | 2010 | 1 | 625 | 431 | 1056 | 344 | 146 | 490 |
| 41 | promising solution | 2011 | 2010 | 1 | 818 | 235 | 1053 | 568 | 121 | 689 |
| 42 | improve accuracy | 2011 | 2010 | 1 | 719 | 321 | 1040 | 395 | 105 | 500 |
| 43 | target task | 2011 | 2010 | 1 | 682 | 345 | 1027 | 428 | 165 | 593 |
| 44 | person re | 2011 | 2010 | 1 | 599 | 412 | 1011 | 168 | 93 | 261 |
| 45 | attention mechanism | 2014 | 2011 | 3 | 2520 | 1247 | 3767 | 1329 | 495 | 1824 |
| 46 | large language model | 2012 | 2011 | 1 | 2890 | 149 | 3039 | 2872 | 142 | 3014 |
| 47 | unsupervised domain adaptation | 2014 | 2011 | 3 | 779 | 398 | 1177 | 514 | 218 | 732 |
| 48 | deep generative model | 2014 | 2011 | 3 | 776 | 263 | 1039 | 405 | 110 | 515 |
| 49 | deep learning model | 2013 | 2012 | 1 | 3442 | 717 | 4159 | 2088 | 359 | 2447 |
| 50 | deep convolutional neural network | 2013 | 2012 | 1 | 1790 | 634 | 2424 | 461 | 90 | 551 |
| 51 | deep model | 2013 | 2012 | 1 | 903 | 403 | 1306 | 397 | 153 | 550 |
| 52 | deep learning technique | 2013 | 2012 | 1 | 1043 | 246 | 1289 | 543 | 93 | 636 |
| 53 | convolutional layer | 2014 | 2012 | 2 | 866 | 340 | 1206 | 318 | 94 | 412 |
| 54 | fewshot learning | 2015 | 2013 | 2 | 837 | 388 | 1225 | 547 | 240 | 787 |

| No. | キーフレーズ | 初出年 | | | 件数 | | | 直近3年の件数 | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | arXiv | 国際会議 | 差 | arXiv | 国際会議 | All | arXiv | 国際会議 | All |
| 1 | contrastive learning | 2017 | 2015 | 2 | 1610 | 624 | 2234 | 1470 | 591 | 2061 |
| 2 | monocular depth estimation | 2016 | 2015 | 1 | 235 | 107 | 342 | 154 | 63 | 217 |
| 3 | finetuned model | 2018 | 2015 | 3 | 236 | 45 | 281 | 202 | 31 | 233 |

## 表5-4　arXivが先行したキーフレーズ

| No. | キーフレーズ | 初出年 | | | 件数 | | | 直近3年の件数 | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | arXiv | 国際会議 | 差 | arXiv | 国際会議 | All | arXiv | 国際会議 | All |
| 1 | graph neural network | 2015 | 2016 | 1 | 3185 | 1130 | 4315 | 2411 | 843 | 3254 |
| 2 | visual question answering | 2015 | 2016 | 1 | 616 | 314 | 930 | 324 | 111 | 435 |
| 3 | variational autoencoders | 2015 | 2016 | 1 | 597 | 213 | 810 | 258 | 81 | 339 |
| 4 | style transfer | 2015 | 2016 | 1 | 415 | 186 | 601 | 231 | 97 | 328 |
| 5 | contrastive loss | 2015 | 2016 | 1 | 415 | 168 | 583 | 342 | 142 | 484 |
| 6 | attention module | 2015 | 2017 | 2 | 368 | 179 | 547 | 236 | 95 | 331 |
| 7 | faster RCNN | 2015 | 2016 | 1 | 315 | 115 | 430 | 116 | 23 | 139 |
| 8 | residual network | 2015 | 2016 | 1 | 278 | 111 | 389 | 92 | 30 | 122 |
| 9 | triplet loss | 2015 | 2016 | 1 | 241 | 128 | 369 | 112 | 50 | 162 |
| 10 | selfsupervised pretraining | 2015 | 2016 | 1 | 260 | 104 | 364 | 223 | 92 | 315 |
| 11 | attention weight | 2015 | 2016 | 1 | 219 | 109 | 328 | 119 | 56 | 175 |
| 12 | adversarial sample | 2015 | 2016 | 1 | 230 | 68 | 298 | 113 | 34 | 147 |
| 13 | augment data | 2015 | 2016 | 1 | 209 | 87 | 296 | 137 | 56 | 193 |
| 14 | forward pas | 2015 | 2016 | 1 | 201 | 86 | 287 | 123 | 44 | 167 |
| 15 | video captioning | 2015 | 2016 | 1 | 175 | 111 | 286 | 78 | 31 | 109 |
| 16 | deep metric learning | 2015 | 2016 | 1 | 178 | 98 | 276 | 87 | 37 | 124 |
| 17 | data protection regulation | 2015 | 2018 | 3 | 225 | 39 | 264 | 107 | 20 | 127 |
| 18 | attention layer | 2015 | 2016 | 1 | 178 | 79 | 257 | 112 | 38 | 150 |
| 19 | federated learning | 2016 | 2019 | 3 | 1743 | 262 | 2005 | 1423 | 236 | 1659 |
| 20 | graph convolutional network | 2016 | 2018 | 2 | 947 | 445 | 1392 | 523 | 216 | 739 |
| 21 | edge device | 2016 | 2017 | 1 | 962 | 174 | 1136 | 681 | 128 | 809 |
| 22 | continual learning | 2016 | 2017 | 1 | 817 | 272 | 1089 | 624 | 216 | 840 |
| 23 | conditional generative adversarial network | 2016 | 2017 | 1 | 319 | 112 | 431 | 137 | 34 | 171 |
| 24 | residual connection | 2016 | 2017 | 1 | 245 | 95 | 340 | 133 | 45 | 178 |
| 25 | digital twin | 2016 | 2018 | 2 | 314 | 18 | 332 | 280 | 12 | 292 |
| 26 | pretext task | 2016 | 2018 | 2 | 239 | 91 | 330 | 203 | 79 | 282 |
| 27 | graph convolution | 2016 | 2018 | 2 | 223 | 107 | 330 | 115 | 51 | 166 |
| 28 | sequential recommendation | 2016 | 2018 | 2 | 196 | 101 | 297 | 150 | 60 | 210 |
| 29 | blackbox setting | 2016 | 2017 | 1 | 208 | 81 | 289 | 122 | 44 | 166 |
| 30 | machine reading comprehension | 2016 | 2018 | 2 | 190 | 94 | 284 | 93 | 44 | 137 |
| 31 | policy network | 2016 | 2017 | 1 | 185 | 88 | 273 | 105 | 48 | 153 |
| 32 | graph convolutional neural network | 2016 | 2018 | 2 | 177 | 74 | 251 | 98 | 34 | 132 |
| 33 | pretrained language model | 2017 | 2018 | 1 | 1565 | 627 | 2192 | 1282 | 523 | 1805 |
| 34 | federate learning | 2017 | 2018 | 1 | 1279 | 221 | 1500 | 994 | 189 | 1183 |
| 35 | neural architecture search | 2017 | 2018 | 1 | 798 | 326 | 1124 | 443 | 186 | 629 |
| 36 | transformer model | 2017 | 2018 | 1 | 774 | 279 | 1053 | 609 | 191 | 800 |
| 37 | backdoor attack | 2017 | 2018 | 1 | 480 | 145 | 625 | 399 | 131 | 530 |
| 38 | fewshot setting | 2017 | 2018 | 1 | 367 | 136 | 503 | 317 | 112 | 429 |
| 39 | explainable AI | 2017 | 2018 | 1 | 345 | 66 | 411 | 263 | 51 | 314 |
| 40 | graph representation learning | 2017 | 2018 | 1 | 289 | 100 | 389 | 206 | 75 | 281 |
| 41 | blackbox attack | 2017 | 2018 | 1 | 245 | 100 | 345 | 130 | 49 | 179 |
| 42 | capsule network | 2017 | 2018 | 1 | 235 | 80 | 315 | 87 | 24 | 111 |
| 43 | proximal policy optimization | 2017 | 2018 | 1 | 248 | 43 | 291 | 160 | 24 | 184 |
| 44 | offline RL | 2017 | 2020 | 3 | 204 | 76 | 280 | 188 | 71 | 259 |
| 45 | generate adversarial | 2017 | 2018 | 1 | 204 | 74 | 278 | 113 | 34 | 147 |
| 46 | large pretrained language model | 2017 | 2018 | 1 | 211 | 64 | 275 | 179 | 52 | 231 |
| 47 | graph attention network | 2017 | 2018 | 1 | 200 | 72 | 272 | 147 | 48 | 195 |
| 48 | loss landscape | 2017 | 2018 | 1 | 197 | 67 | 264 | 123 | 41 | 164 |
| 49 | GNN model | 2018 | 2019 | 1 | 279 | 67 | 346 | 209 | 52 | 261 |
| 50 | OOD detection | 2018 | 2019 | 1 | 232 | 72 | 304 | 198 | 56 | 254 |
| 51 | physicsinformed neural network | 2018 | 2021 | 3 | 284 | 10 | 294 | 246 | 10 | 256 |
| 52 | counterfactual explanation | 2018 | 2019 | 1 | 239 | 49 | 288 | 194 | 40 | 234 |
| 53 | bidirectional encoder representation | 2018 | 2019 | 1 | 220 | 54 | 274 | 143 | 23 | 166 |
| 54 | GLUE benchmark | 2018 | 2019 | 1 | 178 | 91 | 269 | 123 | 72 | 195 |
| 55 | pretraining task | 2018 | 2019 | 1 | 182 | 75 | 257 | 149 | 59 | 208 |
| 56 | COVID19 pandemic | 2019 | 2020 | 1 | 1103 | 162 | 1265 | 757 | 153 | 910 |
| 57 | transformerbased model | 2019 | 2020 | 1 | 563 | 150 | 713 | 490 | 135 | 625 |
| 58 | vision transformer | 2020 | 2021 | 1 | 1143 | 338 | 1481 | 1138 | 338 | 1476 |
| 59 | neural radiance field | 2020 | 2021 | 1 | 465 | 113 | 578 | 460 | 113 | 573 |
| 60 | foundation model | 2020 | 2022 | 2 | 326 | 24 | 350 | 325 | 24 | 349 |

## 5.4 時系列予測に基づく新興テーマ探索

Prophetのforwardテストを用いて、今後成長が見込まれるキーフレーズを推定した。対象としたのは、2021年までの19年間に10件以下の出現で、2022年から2023年10月までの1年10ヵ月間に10件以上出現した2,973件のキーフレーズである。これらについて、2023年10月時点での今後12ヵ月の推移を予測し、トレンドの平均が1.0以上となった206件を「実際に成長しているキーフレーズ」として選定した。

選定された206キーフレーズに対し、2022年1月から2023年10月までの22ヵ月間について、その後12ヵ月のトレンドを計算した。そして、トレンド平均が1.0を超えた時期（予測される出現時期）を求めた。結果は2023年10月評価でのトレンド平均の降順にソートし、表5-5～5-9に示した。

分析の結果、2023年10月のトレンドが大きいキーフレーズほど、トレンド平均が1.0を超えた時期が早い傾向が見られた。また、トレンドが1.0に近いものは、ほぼ2023年10月に1.0を超えている傾向が確認された。抽出されたキーフレーズには、「in-context learning」「stable diffusion」「vision-language model」など、妥当性の高いものが含まれている。最近では「CoT（Chain of Thought）prompting」なども選出された。異なる表記で重複して出現しているものもあるが、49位の「NeRF（Neural Radiance Fields）」など、多くは適切なものと判断できる。

数値のみでは出現検出の精度が不明確なため、図5-10に実際のデータ比較を示した。206件中から、古い例として「in-context learning」「stable diffusion」、中程度の例として「visual prompt」「data drift」、新しい例として「graph anomaly detection」「CoT prompting」を選び、トレンド平均が1になる位置と2023年10月時点でのトレンド値を表示した。

いずれのケースもトレンドはほぼ単調に増加しており、2023年10月で1を超えていれば継続的な増加が見られた。これにより、遡及的に1を超えた時期を特定することが可能と考えられる。その時点でトレンドが1を超えた理由の説明は困難だが、データに基づく客観的な判定から、比較的納得できる結果が得られたと言える。

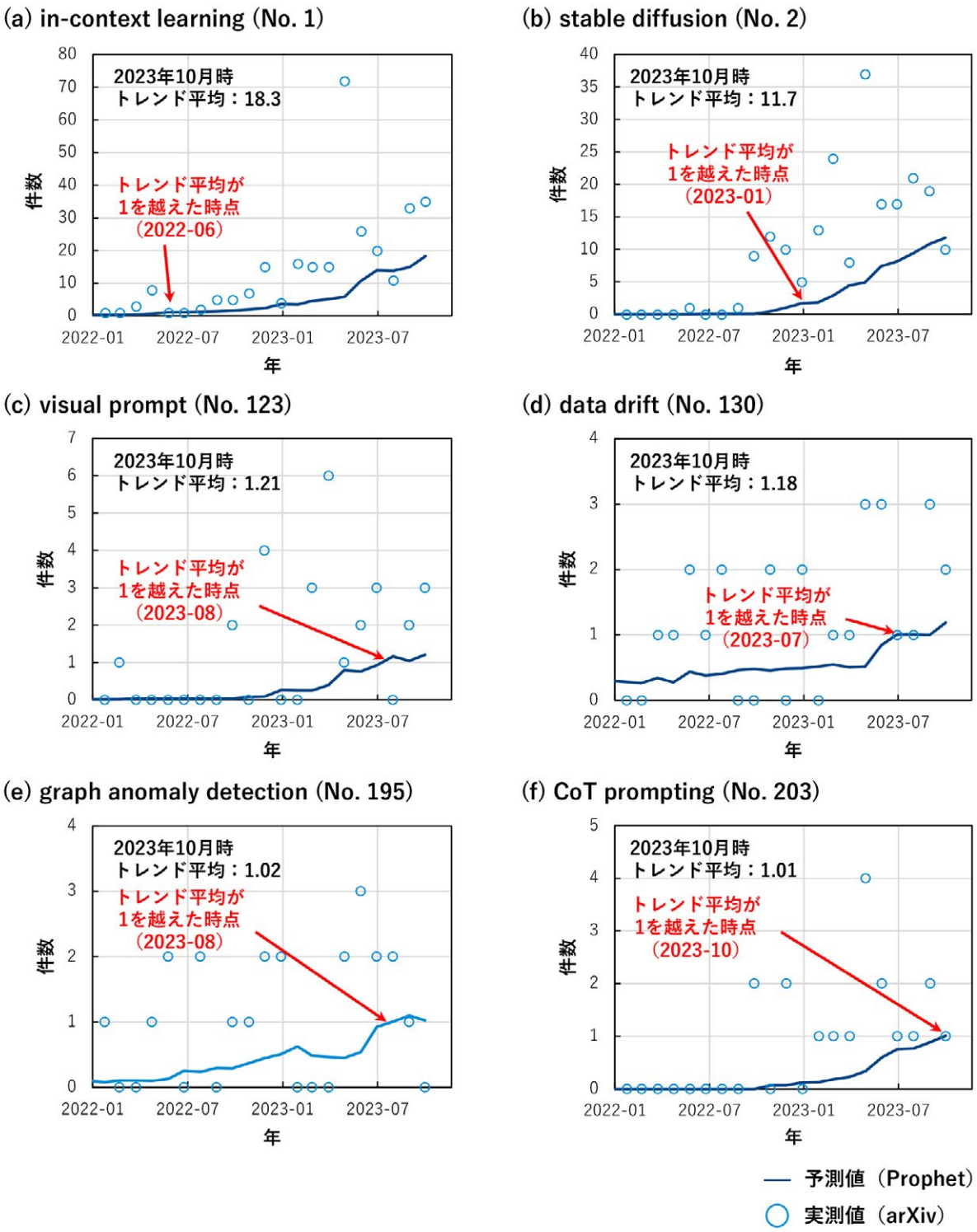


**図5-10　　時系列予測に基づく新興テーマ探索**

**表5-5　今後成長する可能性のあるキーフレーズ（1）**

| No. | キーフレーズ | 2023年10月時トレンド平均 | トレンド平均が1を超えた年月 |
|---|---|---|---|
| 1 | incontext learning | 18.2624 | 2022年6月 |
| 2 | stable diffusion | 11.7435 | 2023年1月 |
| 3 | visionlanguage model | 8.4758 | 2022年10月 |
| 4 | generative AI | 7.1884 | 2023年5月 |
| 5 | neural field | 6.0265 | 2022年7月 |
| 6 | mask image modeling | 5.8483 | 2022年6月 |
| 7 | prompt learning | 5.7916 | 2022年10月 |
| 8 | prompt engineering | 5.5978 | 2023年3月 |
| 9 | pretrained large language model | 5.3717 | 2023年2月 |
| 10 | latent diffusion model | 5.3668 | 2023年2月 |
| 11 | instruction tuning | 5.1103 | 2023年5月 |
| 12 | texttoimage model | 4.3668 | 2023年3月 |
| 13 | texttoimage diffusion model | 4.2134 | 2023年2月 |
| 14 | denoising diffusion model | 4.1078 | 2023年2月 |
| 15 | pretrained visionlanguage model | 4.1028 | 2022年12月 |
| 16 | full finetuning | 3.9139 | 2022年11月 |
| 17 | pretrained diffusion model | 3.8062 | 2023年3月 |
| 18 | testtime adaptation | 3.7518 | 2022年5月 |
| 19 | mask autoencoders | 3.6904 | 2022年9月 |
| 20 | generative pretrained transformer | 3.613 | 2023年4月 |
| 21 | graph transformer | 3.3902 | 2022年11月 |
| 22 | SSL model | 3.3504 | 2022年10月 |
| 23 | crossattention mechanism | 3.3128 | 2022年12月 |
| 24 | generative artificial intelligence | 3.2408 | 2023年7月 |
| 25 | offer valuable | 3.2163 | 2023年7月 |
| 26 | linear probing | 3.1438 | 2022年11月 |
| 27 | extend reality | 3.1272 | 2022年5月 |
| 28 | conditional diffusion model | 2.8619 | 2023年4月 |
| 29 | mask autoencoder | 2.7915 | 2023年1月 |
| 30 | fourier neural operator | 2.743 | 2022年12月 |
| 31 | capability of large language model | 2.7021 | 2023年7月 |
| 32 | texttoimage generative model | 2.6605 | 2023年7月 |
| 33 | input prompt | 2.6289 | 2023年6月 |
| 34 | diffusionbased generative model | 2.6192 | 2023年3月 |
| 35 | mathematical reasoning | 2.5843 | 2023年3月 |
| 36 | prompt template | 2.5596 | 2023年3月 |
| 37 | emergence of large language model | 2.5282 | 2023年7月 |
| 38 | GPT model | 2.4608 | 2023年6月 |
| 39 | prompting strategy | 2.435 | 2023年6月 |
| 40 | natural language prompt | 2.4058 | 2022年11月 |
| 41 | potential of LLMs | 2.3679 | 2023年7月 |
| 42 | LLMs | 2.3673 | 2023年7月 |
| 43 | chain of thought | 2.3665 | 2023年6月 |
| 44 | capability of LLMs | 2.3018 | 2023年6月 |
| 45 | reverse diffusion process | 2.2792 | 2023年5月 |
| 46 | fewshot prompting | 2.2091 | 2023年4月 |
| 47 | prompt design | 2.172 | 2023年6月 |
| 48 | cross attention | 2.1156 | 2022年10月 |
| 49 | NeRF model | 2.1087 | 2023年4月 |
| 50 | chainofthought prompting | 2.0686 | 2023年6月 |

**表5-6　今後成長する可能性のあるキーフレーズ（2）**

| No. | キーフレーズ | 2023年10月時トレンド平均 | トレンド平均が1を超えた年月 |
|---|---|---|---|
| 51 | visionlanguage pretraining | 2.0129 | 2022年6月 |
| 52 | diverse downstream task | 1.9948 | 2023年1月 |
| 53 | future advancement | 1.9202 | 2023年9月 |
| 54 | promptbased learning | 1.9131 | 2022年7月 |
| 55 | classifierfree guidance | 1.9032 | 2023年4月 |
| 56 | stable diffusion model | 1.8984 | 2023年6月 |
| 57 | GitHub copilot | 1.8522 | 2023年6月 |
| 58 | multimodal large language model | 1.8519 | 2023年9月 |
| 59 | scorebased diffusion model | 1.8396 | 2023年6月 |
| 60 | diffusionbased model | 1.8326 | 2023年5月 |
| 61 | success of large language model | 1.8258 | 2023年6月 |
| 62 | generative diffusion model | 1.8188 | 2023年7月 |
| 63 | heterophilic graph | 1.8023 | 2022年7月 |
| 64 | textual prompt | 1.7961 | 2023年5月 |
| 65 | online RL | 1.7926 | 2022年12月 |
| 66 | ability of LLMs | 1.7491 | 2023年8月 |
| 67 | leverage large language model | 1.7063 | 2023年7月 |
| 68 | dataset distillation | 1.6674 | 2023年6月 |
| 69 | opensource model | 1.6648 | 2023年9月 |
| 70 | multiview consistency | 1.6504 | 2023年6月 |
| 71 | SharpnessAware minimization | 1.6407 | 2023年1月 |
| 72 | finetuning paradigm | 1.6269 | 2023年3月 |
| 73 | crossattention module | 1.6019 | 2022年9月 |
| 74 | complex reasoning task | 1.6007 | 2023年7月 |
| 75 | contrastive pretraining | 1.575 | 2022年12月 |
| 76 | nuScenes benchmark | 1.5739 | 2023年4月 |
| 77 | sourcefree domain adaptation | 1.5708 | 2023年4月 |
| 78 | pretraining datasets | 1.5633 | 2023年7月 |
| 79 | emergent ability | 1.5493 | 2023年9月 |
| 80 | transformer backbone | 1.5411 | 2023年5月 |
| 81 | contrastive learning strategy | 1.5253 | 2023年6月 |
| 82 | operator learning | 1.5176 | 2023年4月 |
| 83 | contrastive languageimage pretraining | 1.5116 | 2023年5月 |
| 84 | code generation task | 1.4975 | 2023年6月 |
| 85 | foundational model | 1.4968 | 2023年10月 |
| 86 | decision transformer | 1.4901 | 2023年7月 |
| 87 | opensource LLMs | 1.4737 | 2023年10月 |
| 88 | nonfungible token | 1.4637 | 2023年2月 |
| 89 | current LLMs | 1.4541 | 2023年9月 |
| 90 | prompting technique | 1.4533 | 2023年8月 |
| 91 | zeroshot fashion | 1.448 | 2023年3月 |
| 92 | D4RL benchmark | 1.4405 | 2023年5月 |
| 93 | T2I model | 1.4031 | 2023年8月 |
| 94 | AIgenerated content | 1.4023 | 2023年9月 |
| 95 | deep operator network | 1.4003 | 2023年3月 |
| 96 | knowledgeintensive task | 1.3954 | 2023年6月 |
| 97 | outofdomain setting | 1.387 | 2023年1月 |
| 98 | crossmodal alignment | 1.3855 | 2023年4月 |
| 99 | vision transformer architecture | 1.3854 | 2023年6月 |
| 100 | visual encoder | 1.3792 | 2023年1月 |

**表5-7　今後成長する可能性のあるキーフレーズ（3）**

| No. | キーフレーズ | 2023年10月時トレンド平均 | トレンド平均が1を超えた年月 |
|---|---|---|---|
| 101 | soft prompt | 1.3736 | 2023年6月 |
| 102 | text guidance | 1.3694 | 2023年6月 |
| 103 | vision transformer model | 1.3634 | 2022年12月 |
| 104 | dense retrieval model | 1.355 | 2023年5月 |
| 105 | power of large language model | 1.347 | 2023年10月 |
| 106 | poison sample | 1.3392 | 2023年4月 |
| 107 | text encoders | 1.3379 | 2023年7月 |
| 108 | OpenAIs ChatGPT | 1.3302 | 2023年9月 |
| 109 | large visionlanguage model | 1.2967 | 2023年10月 |
| 110 | generative large language model | 1.2782 | 2023年10月 |
| 111 | BEV feature | 1.2768 | 2023年8月 |
| 112 | rapidly evolving field | 1.2703 | 2023年8月 |
| 113 | speech model | 1.2667 | 2023年6月 |
| 114 | neural collapse | 1.2658 | 2023年10月 |
| 115 | enhance efficiency | 1.2656 | 2023年9月 |
| 116 | individual client | 1.2627 | 2023年9月 |
| 117 | semisupervised object detection | 1.2575 | 2022年11月 |
| 118 | transformerbased pretrained language model | 1.2471 | 2022年11月 |
| 119 | segmentation datasets | 1.2392 | 2023年6月 |
| 120 | training resource | 1.2263 | 2023年9月 |
| 121 | client selection | 1.2181 | 2023年9月 |
| 122 | backdoored model | 1.2115 | 2023年5月 |
| 123 | visual prompt | 1.2098 | 2023年8月 |
| 124 | contrastive training | 1.2024 | 2023年7月 |
| 125 | algorithmic recourse | 1.2022 | 2023年10月 |
| 126 | novelview synthesis | 1.2017 | 2023年9月 |
| 127 | GPT3 model | 1.2002 | 2023年8月 |
| 128 | outofdistribution datasets | 1.1918 | 2023年6月 |
| 129 | backdoor defense | 1.1871 | 2023年6月 |
| 130 | data drift | 1.1834 | 2023年7月 |
| 131 | reasoning capability of LLMs | 1.1829 | 2023年10月 |
| 132 | potential of large language model | 1.1793 | 2023年10月 |
| 133 | generative AI model | 1.1792 | 2023年6月 |
| 134 | class name | 1.1788 | 2023年8月 |
| 135 | scorebased model | 1.1681 | 2023年6月 |
| 136 | segmentation head | 1.1665 | 2023年7月 |
| 137 | controllable text generation | 1.1595 | 2023年7月 |
| 138 | code generation model | 1.1581 | 2023年6月 |
| 139 | pretrained texttoimage diffusion model | 1.1561 | 2023年6月 |
| 140 | SSL framework | 1.1543 | 2022年8月 |
| 141 | camouflage object | 1.1492 | 2023年9月 |
| 142 | text instruction | 1.1459 | 2023年10月 |
| 143 | textvideo retrieval | 1.1452 | 2023年10月 |
| 144 | finetuning data | 1.1423 | 2023年6月 |
| 145 | pretrained backbone | 1.1374 | 2023年5月 |
| 146 | longtailed data | 1.1305 | 2023年6月 |
| 147 | instruction finetuning | 1.128 | 2023年10月 |
| 148 | generative flow network | 1.1217 | 2023年7月 |
| 149 | detection transformer | 1.1217 | 2022年7月 |
| 150 | arithmetic reasoning | 1.1173 | 2023年7月 |

**表5-8　今後成長する可能性のあるキーフレーズ（4）**

| No. | キーフレーズ | 2023年10月時トレンド平均 | トレンド平均が1を超えた年月 |
|---|---|---|---|
| 151 | SemanticKITTI dataset | 1.1152 | 2023年9月 |
| 152 | intermediate reasoning step | 1.115 | 2023年7月 |
| 153 | AI assistant | 1.1135 | 2023年10月 |
| 154 | model scaling | 1.1117 | 2023年10月 |
| 155 | enables LLMs | 1.1112 | 2023年10月 |
| 156 | diverse condition | 1.1073 | 2023年10月 |
| 157 | generalizable representation | 1.107 | 2023年10月 |
| 158 | graph machine learning | 1.1047 | 2022年8月 |
| 159 | sharing raw data | 1.1019 | 2023年5月 |
| 160 | exceptional capability | 1.0956 | 2023年10月 |
| 161 | prompt | 1.0951 | 2023年10月 |
| 162 | zeroshot prompting | 1.0936 | 2023年10月 |
| 163 | lowrank adaptation | 1.0929 | 2023年10月 |
| 164 | denoising step | 1.0927 | 2023年10月 |
| 165 | finetuning language model | 1.0915 | 2023年9月 |
| 166 | class token | 1.0876 | 2023年7月 |
| 167 | google bard | 1.0868 | 2023年10月 |
| 168 | outofdistribution scenario | 1.0866 | 2023年9月 |
| 169 | heterogeneous data distribution | 1.0861 | 2023年9月 |
| 170 | shortcut learning | 1.0808 | 2023年9月 |
| 171 | reconstruction attack | 1.0804 | 2022年11月 |
| 172 | neuromorphic datasets | 1.0803 | 2023年3月 |
| 173 | hypergraph neural network | 1.0744 | 2023年9月 |
| 174 | realtime deployment | 1.0709 | 2023年5月 |
| 175 | generating human | 1.0705 | 2023年8月 |
| 176 | XGBoost model | 1.0647 | 2023年10月 |
| 177 | transformerbased large language model | 1.0607 | 2023年10月 |
| 178 | NonFungible token | 1.0574 | 2023年2月 |
| 179 | finetuning phase | 1.0523 | 2023年1月 |
| 180 | molecular representation learning | 1.0519 | 2022年12月 |
| 181 | small language model | 1.0501 | 2023年10月 |
| 182 | digital human | 1.05 | 2023年8月 |
| 183 | diverse modality | 1.0496 | 2023年9月 |
| 184 | texttoimage generation model | 1.0482 | 2023年8月 |
| 185 | LLMs ability | 1.0426 | 2023年10月 |
| 186 | face limitation | 1.042 | 2023年10月 |
| 187 | selfattention operation | 1.0415 | 2023年10月 |
| 188 | enables multiple client | 1.0409 | 2023年9月 |
| 189 | language model pretrained | 1.0321 | 2023年8月 |
| 190 | transformerbased neural network | 1.0308 | 2023年6月 |
| 191 | heterogeneous client | 1.0286 | 2023年9月 |
| 192 | potential of ChatGPT | 1.0256 | 2023年8月 |
| 193 | crossattention layer | 1.0256 | 2023年7月 |
| 194 | training graph | 1.025 | 2023年6月 |
| 195 | graph anomaly detection | 1.0209 | 2023年8月 |
| 196 | object rearrangement | 1.0189 | 2023年10月 |
| 197 | sharpnessaware minimization | 1.0177 | 2023年1月 |
| 198 | personalize FL | 1.0171 | 2023年9月 |
| 199 | remarkable effectiveness | 1.0145 | 2023年10月 |
| 200 | finegrained alignment | 1.0118 | 2023年10月 |

**表5-9　今後成長する可能性のあるキーフレーズ（5）**

| No. | キーフレーズ | 2023年10月時トレンド平均 | トレンド平均が1を超えた年月 |
|---|---|---|---|
| 201 | open radio access network | 1.0078 | 2023年10月 |
| 202 | integration of artificial intelligence | 1.0072 | 2023年10月 |
| 203 | CoT prompting | 1.0069 | 2023年10月 |
| 204 | mask autoencoding | 1.0039 | 2023年7月 |
| 205 | training LLMs | 1.0036 | 2023年10月 |
| 206 | training large language model | 1.0002 | 2023年10月 |

# 参考文献




［1］ 科学技術振興機構研究開発戦略センター.人工知能研究の新潮流２～基盤モデル・生成AIのインパクト. No. CRDS-FY2023-RR-023, 2023.

［2］ 科学技術振興機構研究開発戦略センター.（研究開発の俯瞰報告書）システム・情報科学技術分野（2023年）. No. CRDS-FY2022-FR-04, 2023.

［3］ 科学技術振興機構研究開発戦略センター.次世代AIモデルの研究開発. No. CRDS-FY2023-SP-03,2024.

［4］ 林和弘,小柴等. arXivに着目したプレプリントの分析.科学技術・学術政策研究所DISCUSSION PAPER187, 2020.

［5］ arXiv. https://arxiv.org/.

［6］ bioRXiv. https://www.biorxiv.org/.

［7］ ChemRxiv. https://chemrxiv.org/.

［8］ medRxiv. https://www.medrxiv.org/.

［9］ PsyArXiv. https://osf.io/preprints/psyarxiv.

［10］ JXiv. https://jxiv.jst.go.jp/index.php/jxiv.

［11］ 林和弘.オープンアクセスの進展が生み出す学術ジャーナルと論文の変容とオープンサイエンス（令和２年１月21日）.文部科学省科学技術・学術審議会情報委員会ジャーナル問題検討部会（第７回）（令和２年10月27日）
https://www.mext.go.jp/kaigisiryo/content/20201026-mxt_jyohoka01-000010663_2.pdf.

［12］ MEXT-NISTEPプレプリント調査・検討チーム.プレプリントをめぐる近年の動向及び今後の科学技術行政への示唆.文部科学省科学技術・学術審議会情報委員会ジャーナル問題検討部会（第７回）（令和２年10月27日）https://www.mext.go.jp/content/20201026-mxt_jyohoka01-000010684_2.pdf.

［13］ arXiv info: Submissions by Category since 2009+.
https://info.arxiv.org/about/reports/submission_category_by_year.html.

［14］ A. Dosovitskiy, et al. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. arXiv:2010.11929, 2020.

［15］ S. Khan, et al. Transformers in Vision: A Survey. ACM Computing Surveys, Vol. 54, No. 200, pp.1-41, 2022.

［16］ Ashish Jaiswal, et al. A Survey on Contrastive Self-supervised Learning. arXiv:2011.00362, 2020.

［17］ Kaiming He, et al. Masked Autoencoders Are Scalable Vision Learners. arXiv:2111.06377, 2021.

［18］ David Foster. Generative Deep Learning: Teaching Machines to Paint, Write, Compose, and Play.O'reilly Media Inc., 2019.

［19］ Ian Goodfellow, et al. Generative Adversarial Nets. In Proceedings of the 28th Conference on NeuralInformation Processing Systems（NIPS 2014）, pp. 2672-2680, Montréal, Canada, December 8-13 2014.

［20］ Diederik P. Kingma and Max Welling. Auto-Encoding Variational Bayes. In Proceedings of the 2ndInternational Conference on Learning Representations（ICLR 2014）, Banff, Canada, April 14-162014.

［21］ Simon J.D. Prince Ivan Kobyzev and Marcus A. Brubaker. Normalizing Flows: An Introduction and Review of Current Methods. IEEE Transactions on Pattern Analysis and Machine Intelligence,Vol. 43, pp. 3964-3979, 2021.

［22］ Florinel-Alin Croitoru, et al. Diffusion Models in Vision: A Survey. arXiv:2209.04747, 2022.

［23］ Aditya Ramesh. Zero-Shot Text-to-Image Generation. arXiv:2102.12092, 2021.

［24］ Aditya Ramesh. Hierarchical Text-Conditional Image Generation with CLIP Latents. arXiv:2204.06125, 2022.

［25］ Chitwan Saharia, et al. Photorealistic Text-to-Image Diffusion Models with Deep Language Understanding.arXiv:2205.11487, 2022.

［26］ Jiahui Yu, et al. Scaling Autoregressive Models for Content-Rich Text-to-Image Generation. arXiv:2206.10789, 2022.

［27］ Huiwen Chang, et al. Muse: Text-To-Image Generation via Masked Generative Transformers.arXiv:2301.00704, 2023.

［28］ 坪井祐太,海野裕也,鈴木潤.深層学習による自然言語処理.講談社, 2017.

［29］ 岡崎直観,荒瀬由紀,鈴木潤,鶴岡慶雅,宮尾祐介. IT Text自然言語処理の基礎.オーム社, 2022.

［30］ 中澤敏明.機械翻訳の新しいパラダイム：ニューラル機械翻訳の原理.情報管理, Vol. 66, No. 5, pp.299-306, 2017.

［31］ 西田京介,斉藤いつみ.深層学習におけるアテンション技術の最新動向.電子情報通信学会誌, Vol. 101,No. 6, pp. 591-596, 2018.

［32］ Ashish Vaswani, et al. Attention Is All You Need. In Proceedings of the 31st Conference on Neural Information Processing Systems（NIPS 2017）, Long Beach, CA, USA, December 4-9 2017.

［33］ Jacob Devlin, et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.arXiv:1810.04805, 2018.

［34］ Yinhan Liu, et al. RoBERTa: A Robustly Optimized BERT Pretraining Approach. arXiv:1907.11692,2019.

［35］ Zhenzhong Lan, et al. ALBERT: A Lite BERT for Self-supervised Learning of Language Representations.arXiv:1909.11942, 2019.

［36］ Yu Sun, et al. ERNIE: Enhanced Representation through Knowledge Integration. arXiv:1904.09223,2019.

［37］ Li Dong, et al. Unified Language Model Pre-training for Natural Language Understanding and Generation. arXiv:1905.03197, 2019.

［38］ Jiasen Lu, et al. ViLBERT: Pretraining Task-Agnostic Visiolinguistic Representations for Vision-and-Language Tasks. arXiv:1908.02265, 2019.

［39］ Weijie Su, et al. VL-BERT: Pre-training of Generic Visual-Linguistic Representations. arXiv:1908.08530, 2019.

［40］ Tom Brown, et al. Language Models are Few-Shot Learners. In Proceedings of the 34th Conference on Neural Information Processing Systems（NeurIPS 2020）, December 6-12 2020.

［41］ Shaden Smith, et al. Using DeepSpeed and Megatron to Train Megatron-Turing NLG 530B, A Large-Scale Generative Language Model. arXiv:2201.11990, 2022.

［42］ Long Ouyang, et al. Training language models to follow instructions with human feedback. arXiv:2203.02155, 2022.

［43］ Rishi Bommasani, et al. On the Opportunities and Risks of Foundation Models. arXiv:2108.07258,2021.
［44］ Norman P. Jouppi, et al. In-Datacenter Performance Analysis of a Tensor Processing Unit. In Proceedings of 44th Annual International Symposium on Computer Architecture（ISCA ’17）, pp.1-12, Toronto, Canada, June 24-28 2017.
［45］ Yu-Hsin Chen, Joel Emer, and Vivienne Sze. Eyeriss: A Spatial Architecture for Energy-Efficient Dataflow for Convolutional Neural Networks. In Proceedings of 43rd Annual International Symposium on Computer Architecture（ISCA ’16）, pp. 367-379, Seoul, South Korea, June 18-22 2016.
［46］ Thomas Norrie, et al. Google’s Training Chips Revealed:TPUv2 and TPUv3. In Proceedings of 2020 IEEE Hot Chips 32 Symposium（HCS）, pp. 1-70, Palo Alto, CA, USA, August 16-18 2020.
［47］ Chris Nicol, et al. Wave Computing:A Dataflow Processing Chip for Training Deep Neural Networks.In Proceedings of 2019 IEEE Hot Chips 29 Symposium（HCS）, pp. 20-22, Cupertino, CA, USA,August 18-20 2019.
［48］ Jeremy Fowers, et al. A Configurable Cloud-Scale DNN Processor for Real-Time AI. In Proceedings of 45th Annual International Symposium on Computer Architecture（ISCA ’18）, pp. 1-14, LosAngeles, CA, USA, June 2-6 2018.
［49］ Shouyi Yin, et al. An Ultra-High Energy-Efficient Reconfigurable Processor for Deep Neural Networks with Binary/Ternary Weights in 28nm CMOS. In Proceedings of 2018 IEEE Symposium onVLSI Circuits, pp. 37-38, Honolulu, HI, USA, June 18-22 2018.
［50］ 中原啓貴．Fpgaを用いたエッジ向けディープラーニングの研究開発動向．人工知能，Vol. 33, No. 1, pp.31-38, 2018.
［51］ Kodai Ueyoshi, et al. QUEST:A 7.49TOPS multi-purpose log-quantized DNN inference engine stacked on 96MB 3D SRAM using inductive-coupling technology in 40nm CMOS. In Proceedings of 2018 IEEE International Solid-State Circuits Conference（ISSCC）, pp. 216-218, San Francisco,CA, USA, February 11-15 2018.
［52］ Taro Fujii, et al. New Generation Dynamically Reconfigurable Processor Technology for Accelerating Embedded AI Applications. In Proceedings of 2018 IEEE Symposium on VLSI Circuits, pp.41-42, Honolulu, HI, USA, June 18-22 2018.
［53］ 株式会社PreferredNetworks. PFNの深層学習用スーパーコンピュータMN-3、省電力性能ランキングGreen500で再び世界１位を獲得（2021年6月28日プレス発表）. https://www.preferred.jp/ja/news/pr20210628/.
［54］ David Ha and Jürgen Schmidhuber. World Models. arXiv:1803.10122, 2018.
［55］ S. M. Ali Eslami, et al. Neural scene representation and rendering. Science, Vol. 360, No. 6394,pp. 1204-1210, 2018.
［56］ Pin-Chu Yang, et al. Repeatable Folding Task by Humanoid Robot Worker using Deep Learning.IEEE Robotics and Automation Letters, Vol. 2, No. 2, pp. 397–403, 2016.
［57］ 尾形哲也．ディープラーニングがロボットを変える．日刊工業新聞社，2017.
［58］ Wenshuai Zhao, Jorge Peña Queralta, and Tomi Westerlund. Sim-to-Real Transfer in Deep Reinforcement Learning for Robotics: a Survey. arXiv:2009.13303, 2020.

（ウェブサイトは2024年5月31日時点）

**作成メンバー**


| | | | | |
|---|---|---|---|---|
| 総括責任者 | 木村 康則 | 上席フェロー | CRDS | システム・情報科学技術ユニット |
| リーダー | 嶋田 義皓 | フェロー | CRDS | システム・情報科学技術ユニット |
| メンバー | 青木 孝 | フェロー | CRDS | システム・情報科学技術ユニット |
| | 福島 俊一 | フェロー | CRDS | システム・情報科学技術ユニット |
| | 茂木 強 | フェロー | CRDS | システム・情報科学技術ユニット |
| | 吉脇 理雄 | フェロー | CRDS | システム・情報科学技術ユニット |


**調査報告書** **CRDS-FY2024-RR-04**

# プレプリントサーバーarXivを利用した AI分野の研究動向俯瞰調査

# Panoramic Survey of Trends in AI Research Field using arXiv


令和 6 年 8 月　August 2024
ISBN 978-4-88890-932-7

**国立研究開発法人科学技術振興機構　研究開発戦略センター**
**Center for Research and Development Strategy, Japan Science and Technology Agency**

〒102-0076 東京都千代田区五番町 7 K's 五番町
電話　03-5214-7481
E-mail crds@jst.go.jp
https://www.jst.go.jp/crds/



本書は著作権法等によって著作権が保護された著作物です。
著作権法で認められた場合を除き、本書の全部又は一部を許可無く複写・複製することを禁じます。
引用を行う際は、必ず出典を記述願います。
なお、本報告書の参考文献としてインターネット上の情報が掲載されている場合には、本報告書の発行日の１ヶ月前の日付で入手しているものです。
上記日付以降後の情報の更新は行わないものとします。

This publication is protected by copyright law and international treaties.
No part of this publication may be copied or reproduced in any form or by any means without permission of JST,
except to the extent permitted by applicable law.
Any quotations must be appropriately acknowledged.
If you wish to copy, reproduce, display or otherwise use this publication, please contact crds@jst.go.jp.
Please note that all web references in this report were last checked one month prior to publication.
CRDS is not responsible for any changes in content after this date.

ISBN 978-4-88890-932-7

FOR THE FUTURE OF
SCIENCE AND
SOCIETY

CRDS

https://www.jst.go.jp/crds/