

ソーシャルネットワークにおける 人間とAIの協働がもたらす集団的創造性

The Dynamics of Collective Creativity in Human-AI Social Networks

Shota Shiiku^{1,5} * Raja Marjeh² Manuel Anglada-Tort^{3,5} NoriJacoby^{4,5}

¹Shizuoka University

²Princeton University

³Goldsmiths, University of London

⁴Cornell University

⁵Max Planck Institute for Empirical Aesthetics

Abstract: Generative AI is reshaping modern culture, enabling individuals to create high-quality outputs across domains such as images, text, and music. However, we know little about the impact of generative AI on collective creativity. This study investigates how human-AI interactions shape collective creativity within experimental social networks. We conducted large-scale online experiments with 973 participants and AI agents in a creative writing task. Participants (either humans or AI) joined 5×5 grid-based networks, and were asked to iteratively select, modify, and share stories. Initially, AI-only networks showed greater creativity (rated by a separate group of 100 human raters) and diversity than human-only and human-AI networks. However, over time, hybrid human-AI networks became more diverse in their creations than AI-only networks. In part, this is because AI agents retained little from the original stories, while human-only networks preserved continuity. These findings highlight the value of experimental social networks in understanding human-AI hybrid societies.

1 はじめに

洞窟壁画から交響曲に至るまで、創造性は、新たな可能性を想像し、実現する、人間の体験を定義してきた能力である。近年、生成AIが人間のアイデア形成とその実践プロセスを変革しており、我々は興味深い転換期に立っている [1]。これらのAIエージェントは単なる支援ツールではなく、創造プロセスにおける能動的なパートナーとして機能しており [2, 3]、かつてない形で作家、音楽家、芸術家と協働している。このハイブリッド社会において、人間と知能機械が相互に作用するネットワークは、個々の行動だけでは集団全体が生み出した結果の質を推し量ることができない複雑な社会システムを構成している [4]。しかしながら、人間とAIエージェントが混在するネットワークにおいて、集団的創造性がどのように出現するかを研究することは依然として大きな課題となっている。

生成AIは、カスタマーサポート、プログラミング、学術執筆などのルーティーン作業の質と効率を向上させることが示されているが、人間の創造性に及ぼす影響については十分に理解されていない [5]。一部の研究では、人間とAIの協働が個人の創造性を高める可能性を示唆しており、新しいアイデアを生み出す効率やスピードが向上するとの報告もある [6, 7]。対照的に、大規模言語モデルに内在する美的および文化的バイアスが、芸術や文化における均質化効果をもたらし、全体的な多様性を制限する可能性も指摘されている [8, 6]。これまでの研究は、生成AIが人間の創造性に及ぼす潜在的影響を示し始めているが、具体的な相互作用がどの程度人間の創造的潜在力を引き出し、またそれがどのように人間とAIの協働によって強化されるかを定量化するのは困難だった。その一因として、これらのプロセスを統制されたかつ大規模に実験的に検証することが難しいという点が挙げられる。

本研究では、人間と生成AIが参加するソーシャルネットワーク内での集団的創造性を研究するためのオ

*連絡先：静岡大学総合科学技術大学院
〒432-8011 静岡県浜松市中区城北 3-5-1
E-mail:shiiku.shota.19@shizuoka.ac.jp

ンライン実験環境を実装した。実験の各世代では、参加者（人間または AI）がネットワーク上の隣接者から物語を選択・修正し、次の世代の参加者に向けてその創作物を共有する（図 1A 参照）。時間の経過とともに、集団的創作物の特徴（例：質、多様性）が出現・進化し、実験条件（例：人間と AI の相互作用のレベル）の因果関係と集団的創造性の出現との関係を明らかにすることが可能となる。具体的には、同一のシミュレートされたネットワーク内で、(1) Human-only, (2) AI-only, (3) Human-AI collaboration の 3 条件を体系的に比較した（図 1B 参照）。ここでは、科学的執筆、エンターテインメント、マーケティング、デザインなど実世界の多くの創造的課題の中心となる反復的な認知プロセスである物語作成に着目した。この創造的手法は、ブレインストーミング、物語の生成、文体やトーンの改善など、生成 AI が人間と共同でアイデア出しや創作活動を行う上で大きな可能性を秘めている [2, 6]。

本研究のアプローチは、人間と AI の効果的な協働を可能にする条件と、その相互作用のダイナミクス、特に反復的な創造プロセスにおいて多様性と創造性がどのように出現・変化するかについて貴重な知見を提供する。より広くは、現実世界の実験において生態学的妥当性の高い、オープンエンドな課題を用いることが、人間と AI の集団的知性および創造性を研究するための強力なアプローチとなり得ることを示唆している。

2 背景

2.1 ツールとしての AI

生成 AI は、テキスト、画像、音声など、さまざまなモダリティにおいて高品質かつ人間に近いコンテンツを生成することが可能である。これにより、多様な AI 駆動ツールが開発された。例えば、GitHub Copilot のようなコーディングアシスタント、医療分野での応用、ロボティクス、自律走行車などが挙げられる。

しかし、AI ツールの効果はどの程度なのでしょうか。開発投資が進む中、AI 支援による意思決定の研究は賛否が分かれてる。医療、カスタマーサービス、科学研究といった分野で AI がパフォーマンスを向上させるとする研究もあれば [9, 10]、逆の結果を示す研究も存在する [11, 12]。最近の 106 件の実験研究のシステムティックレビューとメタ解析はこの分かれ方を確認しており、平均的には、人間と AI のチームは単独で作業する人間よりも良い成果を上げている [5]。人間と AI の協働を改善する有望なアプローチの一つは、AI を「思考のパートナー」として概念化することである。これは、従来の AI ツールとは異なり、機械の「思考パートナー」が人間の意図や内面状態を動的に推測し、人間

同士の対話を模倣する形で応答を適応させ、より深く生産的な協働を促進するものである [2]。

2.2 人間と AI の協働

AI の人間社会への統合は、特に集団レベルで広範な影響を及ぼす [4]。AI は個々の人間だけでなく、複数の個人と AI エージェントが同時に相互作用することによって生じる社会システムのダイナミクスにも影響を与える [4, 3]。したがって、生成 AI が集団行動をどのように形成するか、複数の個人と AI エージェント間の相互作用から生じる新たな構造に着目してその役割を理解することが非常に重要になっている [13, 4, 3]。

計算機科学的・実験的手法の進展により、ソーシャルネットワーク内で集団行動がどのように出現するかを研究することが可能になった [14, 15]。多くの研究は、人間やエージェントで構成される人工的なソーシャルネットワークにおいて、ネットワークの規模やトポロジーがコンセンサス形成や集団的知性にどのように影響を与えるかを検証した [16, 17, 18]。例えば、Shirado (2017) は、単純なボットエージェントのようなノイズをシステムに導入することで、局所的最適解から脱出しコンセンサスを促進できることを示した [19]。しかし、これまでの研究は、問題空間内に解があらかじめ定義されているような単純な 1 次元タスク（問題解決、推論、意思決定など）にほぼ限定されていた。一方、創造性は、オープンエンドな課題、広大な探索・活用空間、複雑な問題解決環境との関わりが必要とされる。

2.3 創造性研究

人間の創造性に関する研究は、数十年にわたって創造的プロセスの心理計量的測定を用いて行われてきた [20]。従来の「創造性テスト」— 例えば、Guilford による「知能の構造」発散生成テスト [21] や Torrance の創造的思考テスト [22] — は、今日においても創造性研究で広く利用されている。これらのテストは、非反復的な単語のリストを生成したり、ある対象物の新たな利用法をブレインストーミングするなど、単純な連想の測定に依拠している。意味ネットワークの計算モデルの進展により、従来の研究の限界を補完し、人々が概念の意味空間をどのように探索するかを定量的に検証し、創造性に関連する主要な探索戦略を明らかにする強力な手法が提供されている [23]。さらに、集団でのグリッドベースの画像への描画など、言語に依存しないより複雑なオープンエンドなタスクに取り組む研究も行われており、計算的アプローチを用いて創造プロセスにおける探索と活用のバランスをモデル化している [24, 25]。しかしながら、これらの研究は主に個

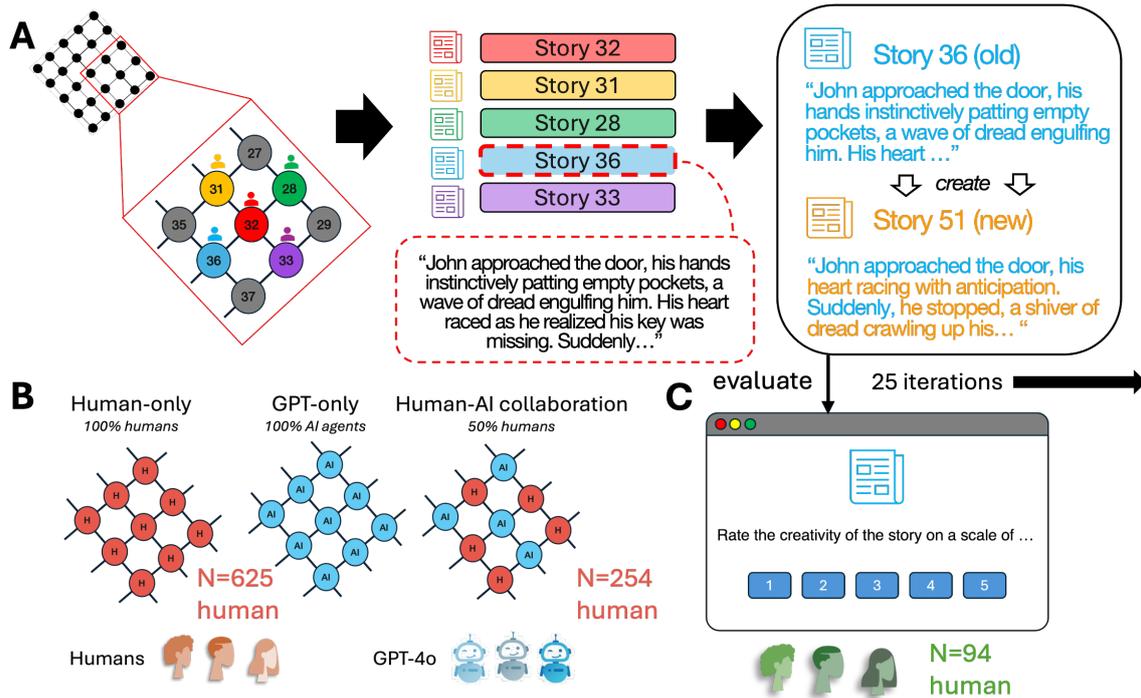


図 1: 集団的創造性を検証するための実験フレームワーク. (A) 参加者はソーシャルネットワークに参加し、短い物語を選択、修正、伝達する課題に従事する. (B) 人間のみのみ、AI のみ、人間と AI の 3 つのネットワーク構成を検証する. (C) 物語の創造性は、別のグループの人間評価者によって評価される.

人レベルでの創造性を扱っており、AI を創造プロセスの一部として考慮していなかった.

2.4 人間と AI が融合した社会における集団的創造性

AI の登場は、人間の創造性を根本から再形成している [1]. AI エージェントは個人と協働するだけでなく、人間が活動する社会環境そのものを形作る役割を果たしている [4]. この非線形的な影響は、人間と AI の相互作用を集団レベルで研究する上で大きな課題を提示する. Doshi & Hauser (2024) は、生成 AI が個人の創造性を向上させる一方で、グループ全体で生み出されるアイデアの範囲や独創性である集団的多様性を低下させる可能性があることを示した [6]. この研究は、AI 駆動の創造性が集団に及ぼす影響の初期的な検証として重要だが、多様なエージェントが時間とともに相互作用するソーシャルネットワーク内で、これらの効果がどのように進展するかを捉えきれていなかった.

本研究のアプローチは、現実の状況に即したオープンエンドな方法を導入することで、数百名規模の参加者が関与するソーシャルネットワーク内での大規模な社会的相互作用を検証可能にする. 反復的な物語作成

という創造的課題を用いることで、アイデアがどのように伝播し、変容し、多様化していくかを、人間、AI、そしてその混合協働が関与するネットワーク内で検証する. このフレームワークは、人間と AI エージェント間の複雑な相互作用を集団レベルで解析することを可能にする.

3 方法

3.1 参加者

参加者はオンラインプラットフォーム Prolific¹ を通じて募集した. すべての参加者は英国から募集され、母語として英語を使用していると確認された. 合計で 973 名の人間参加者を募集するとともに、GPT-4o [26] への呼び出しを 996 回実施した (2024 年 9 月 3 日にリリースされたバージョンを使用). 参加者は、承認済みの研究プロトコルに従ってインフォームド・コンセントを行い、時給 £9 の報酬が支払われた.

¹www.prolific.com

3.2 集団的創造性に関するネットワーク実験

本実験では、参加者（人間または AI）が 5×5 のソーシャルネットワークに組み込まれ、クリエイティブな物語作成課題に従事する大規模オンライン実験を行った。参加者はネットワーク内の隣接者から物語を選択・修正し、その創作物が 25 回のイテレーションを通じて他の参加者に提示された（図 1A 参照）。人間と AI エージェントの双方に対して、同一のプロンプト “Please creatively elaborate on the story, adding your own details and ideas” を提示した。参加者はネットワーク内の 1 つのノードにランダムに割り当てられ、実験への参加は一度のみだった。すべてのネットワークは、以下の同一の物語で初期化された：

“As John reached for his front door, he realized his key was missing. Panic set in as he searched his pockets, but the key was nowhere to be found. Feeling defeated, he slumped against the door, only to hear a jingle from inside—his cat had been playing with the key all along.”

実験条件は 3 通りとした（図 1B 参照）：(1) *Human-only* — 全ノードに人間参加者が配置される ($N=625$)、(2) *AI-only* — 全ノードが GPT-4o を用いてシミュレートされる（625 回の OpenAI API 呼び出し）、(3) *Human-AI collaboration* — 人間参加者 ($N=254$, 50%) と GPT-4o エージェント（50%）が均等に配置される条件である。

すべての実験は、先進的なオンライン心理学実験を可能にする Python ベースのフレームワーク PsyNet² を用いて実施された [27]。

3.3 創造性の測定

各条件における創作物の質を評価するため、別の 100 名の人間参加者を対象に検証実験を実施した。各参加者には、すべての実験条件から無作為に選ばれた物語が提示され、5 段階尺度（1：not creative at all, 5：extremely creative）で創造性を評価した。各参加者は 20 作品ずつ評価した。重要な点として、参加者にはその物語が人間によるものか AI によるものかは知らされなかった。

主観的な創造性評価に加え、物語の多様性も測定した。具体的には、TF-IDF によるベクトル化を用いて各物語のセマンティックな埋め込みを作成し、あらかじめ定義されたイテレーショングループ内で全物語間のコサイン類似度を計算した。多様性は、平均コサイ

ン類似度の逆数として定量化し、類似度が低いほど多様性が高いと評価した。

4 結果

4.1 創造性と多様性

まず、3 つの実験条件における集団的創造性のダイナミクスを検証した。図 2A は、検証実験における各条件のイテレーションごとの平均創造性評価を示している。*AI-only* の条件は、創造性評価が最も高く ($M = 3.571$, $SD = 1.026$, 95% CI: [3.491, 3.652])、*Human-only* および *Human-AI* の条件よりも有意に高い結果となった ($p < .001$)。一方、*Human-only* の条件 ($M = 2.482$, $SD = 1.026$, 95% CI: [2.389, 2.576]) と *Human-AI* の条件 ($M = 2.327$, $SD = 1.159$, 95% CI: [2.184, 2.576]) は類似の評価を示し、いずれも $p < .001$ だった。この結果は、GPT がシンプルなクリエイティブ・ライティング課題において、人間単独よりも容易に優れた成果を示すことを実証している。興味深いことに、*AI-only* と *Human-AI* のネットワークは、実験の最終段階において有意な正の改善を示した (GPT: $M = 0.464$, $SD = 0.154$, $p < 0.001$, *Human-AI*: $M = 0.264$, $SD = 0.175$, $p < 0.001$)。一方、*Human-only* のネットワークは時間とともに改善が見られなかった ($M = -0.133$, $SD = 0.067$, $p = 1.000$) (図 2C 参照)。

次に、集団的創作物の多様性について検討した（図 2B, D 参照）。この評価では、*AI-only* の条件が最も高い多様性 ($M = 0.880$, $SD = 0.017$, 95% CI: [0.859, 0.900]) を示し、次いで *Human-AI* ($M = 0.860$, $SD = 0.043$, 95% CI: [0.806, 0.913])、*Human-only* が最も低い多様性 ($M = 0.823$, $SD = 0.019$, 95% CI: [0.800, 0.900]) となった。

しかし、集団的創造性の進展において興味深い所見が得られた。初期（1～5 世代）では、*AI-only* ネットワークの物語は創造性と多様性ともに最も高い評価を示していたが、世代を重ねるにつれて多様性は着実に低下し、 $M = -0.034$, $SD = 0.17$, 95% CI: [-0.047, -0.021] の低下が見られた（図 2D 参照）。これに対し、*Human-AI* のネットワークは、初期には *Human-only* と同等の多様性だったが、時間とともに最も大きな増加を示し ($M = 0.098$, $SD = 0.039$, 95% CI: [0.064, 0.131])、最終的に全体として最高の多様性スコアを達成した。

4.2 物語のセマンティック解析

人間生成と AI 生成の物語の違いを可視化するため、すべての物語に基づくセマンティック埋め込み空間を

²www.psy.net.dev

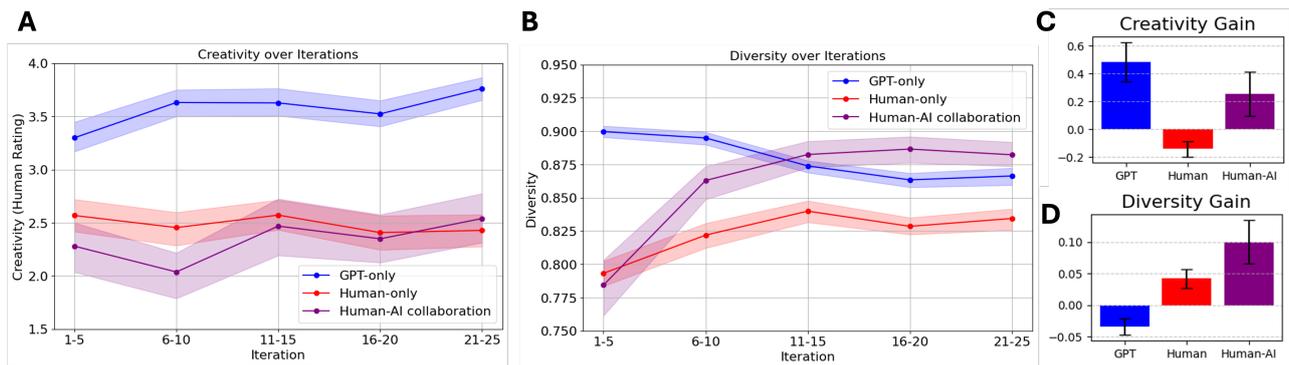


図 2: 集団的創造性のダイナミクス. (A) 人間評価者による物語の平均創造性評価の時間変化. (B) 物語の多様性 (逆類似度) の時間変化. 横軸は 25 回のイテレーションを示し, 各 5 回ごとにグループ化されている. エラーバーは参加者間の標準偏差を表す. (C-D) 創造性と多様性の向上: 初回イテレーションから最終イテレーションまでに測定された創造性と多様性の改善.

作成した. 図 3 は, UMAP による物語の埋め込み結果を示しており, 3 つの条件間のダイナミクスを視覚化している. 各点は 1 つの物語を表し, 点同士の近接は元の高次元空間におけるセマンティックな類似性を反映している. 高次元空間を算出するため, まず事前学習済みモデル (all-MiniLM-L6-v2) を用いて各物語のセマンティック埋め込みを生成し, その後 UMAP を適用して 2 次元に次元削減することで, 物語間のセマンティック類似性を可視化した [28].

この解析から 3 つの重要な知見が得られた. 第一に, *Human-only* と *AI-only* の条件は明確なクラスターを形成しており, 物語に用いられる言語や意味内容に顕著な違いがあることを示している. 図 3 の右下のワードクラウドには, 「universe」や「cosmic」などの宇宙関連の単語が見受けられ, これは AI が関与する物語生成に特徴的な語彙であることを示唆している. 一方, 図 3 左側のワードクラウド (主に人間生成のデータに由来) には, 元のシード物語に近いテーマ, 例えば「John」や「Key」といったキーワードが多く含まれていることが示されている. 第二に, *Human-only* の条件では, 「Amanda」や「David」, 「Tom」といった追加の人名が登場しており, これは物語の主人公や内容に変化が生じた可能性を示唆するものである. 第三に, *Human-AI* の条件では, 人間が AI 的な語彙やセマンティクスを取り入れ, 2 つの条件の中間を補完するような語彙選択となっており, これは人間が AI 生成の新たなアイデアに影響される可能性を示している.

これらのクラスターは, 創作における「ニッチ」(特定のストーリーテリングスタイルの領域) として出現し, 人間の創造性と AI によるセマンティック拡張との相互作用から特徴付けられるものである.

最後に, 物語の語彙動態を把握するため, TF-IDF スコアによって各物語中の主要単語の出現頻度を算出し

た. すべての物語を統合し, 各語の TF-IDF を計算, 頻出順に並べることで, 最も頻出する単語を特定した. 図 4 は, 物語の語彙動態を示しており, 横軸に単語, 縦軸にイテレーション回数をプロットしている. 各座標 (x, y) の円は, イテレーション y において単語 x が使用されたことを示し, 円の大きさはそのイテレーション内での出現頻度を表している. 例えば, 「dreams」や「danced」, 「celestial」といった単語は, *AI-only* の条件下でのみ出現・継続しており, 対照的に, 元の物語に由来する「John」や「cat」, 「keys」といった単語は, 人間のみの条件下で実験終了まで持続していることが見受けられる. これは, 人間が元のストーリーラインに沿った新たな物語を創作する傾向にある一方で, AI は特定の創造的テーマ (例: 宇宙関連) に収束する傾向があることを示唆している. この収束は創造的ではあるものの, タスクの解釈における人間と AI の違いを反映している可能性がある.

4.3 結果のまとめ

この研究は, 物語作成課題における人間と AI の創造性の相互作用が, 実験的ソーシャルネットワーク内の集団的創造性にどのように影響を与えるかを検証した. 3 つの異なるネットワーク構成 (*Human-only*, *AI-only*, *Human-AI collaboration*) を検討し, 人間と AI の協働がより高い創造性へと導くか否かを評価した.

我々の知見は, 初期段階において AI のみのネットワークが, 人間のみおよび混合ネットワークと比較して, 創造性と多様性の両面で最も高い成果を示したことを明らかにした. イテレーションが増すにつれて, AI のみのネットワークの創造性優位性は維持され, GPT の新規かつ多様な物語生成能力が本課題に適していることが示唆される. しかし, AI 生成の物語は初期には

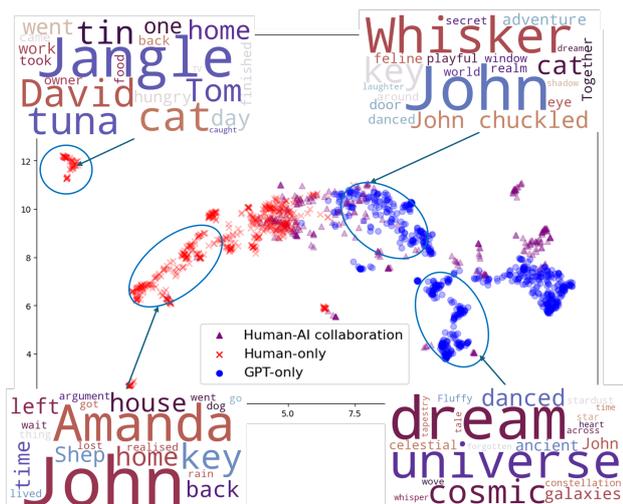


図 3: 共有セマンティック埋め込み空間の UMAP 投影。特定のクラスタのワードクラウドをハイライトしている。

高い多様性を示したものの、時間の経過とともにその多様性は徐々に低下した。これは、GPT が新たなアイデアを提示できる一方で、時間とともにテーマが収束する傾向があることを示している。この結果は、生成 AI が個人の創造性を向上させる一方で、集団的な多様性を低下させる可能性があるという先行研究を補完するものである [6]。

対照的に、人間と AI が混在するネットワークは、初期には多様性が低かったものの、時間の経過とともに AI のみのネットワークを上回る多様性を獲得した。この変化は、AI が物語生成において核となる物語要素を無視して新規性を追求する一方で、人間は主要なストーリー要素（例：主人公の John や key など）を保持する傾向があるためと考えられる。混合ネットワークにおける人間と AI の相互作用は、安定性と新規性の動的なバランスをもたらし、結果的に時間とともにより豊かで多様な創作物を生み出すことが可能となった。これは、人間と AI の協働が双方の強みを活かし、より充実した多様な創造的成果をもたらす可能性を示唆している。

5 考察

本研究は、物語作成課題における人間と AI の創造性の相互作用が、実験的ソーシャルネットワーク内での集団的創造性にどのように影響するかを探求した。人間のみ、AI のみ、そして人間と AI の 3 種類のネットワーク構成を検証し、人間と AI の協働が創造性向上に寄与するか否かを評価した。

結果として、初期段階から AI のみのネットワークは創造性と多様性の両面で最高の成果を示し、GPT による新規かつ発散的な物語生成能力が本課題に非常に適していることが明らかとなった。しかし、AI が生成した物語は初期には高い多様性を示しながらも、時間とともにテーマの収束が見られ、多様性が徐々に低下する傾向が観察された。対照的に、人間と AI が混在するネットワークは、初期には低い多様性であったものの、最終的には AI のみのネットワークを上回る多様性を達成した。これは、AI が革新的なアイデアを提示する一方で、核心的な物語要素を保持するのは主に人間であるということを示唆している。

創造的課題として用いた物語作成タスクは、従来の人間の創造性研究で用いられてきた手法と整合しているものの、GPT のような大規模言語モデルにとっては比較的容易なものであり、その結果、人間単独の場合よりも高い成果を示した。これは必ずしも AI が本質的に人間よりも創造的であることを意味するのではなく、人間の認知プロセスを前提としたタスク設計の限界を反映していると考えられる。AI は膨大なデータセットを基に、迅速かつ意味的に整合性のあるテキストを生成する能力に長けている一方で、人間の創造性は記憶の制約、認知バイアス、暗黙のナラティブ規範に影響される [29]。

今後の創造性研究においては、人間の創造性評価に適しているが AI 能力や人間と AI の相乗効果を十分に評価できないタスクから脱却し、双方にとって意味のある挑戦的な課題の設計に注力する必要がある。これにより、集団的創造性やその根底にあるメカニズムをより精緻に理解することが可能になるだろう。

6 限界と結論

本研究は、実験的社会における社会的自然発生的ダイナミクスをシミュレートすることを目的としていた。しかしながら、本手法は現実世界における社会的相互作用の展開とは異なる。実験では参加者は一度きりの参加であったのに対し、実世界では個人は継続的に相互作用し、交流する。また、現実のやりとりは、本実験で用いた単純な物語作成という課題よりもはるかに複雑である。さらに、シミュレートされたソーシャルネットワークはグリッド型という静的なトポロジーを用いたが、実世界のソーシャルネットワークは、友人関係や職場などのモジュラー構造を持ち、不規則なパターンや変化する接続関係を伴う。個人は固定された接続先に縛られるのではなく、積極的に社会的つながりを形成・調整することが可能である。最後に、本パラダイムでは参加者間の直接コミュニケーションを回避し、パートナーが人間か AI かといった情報を隠蔽

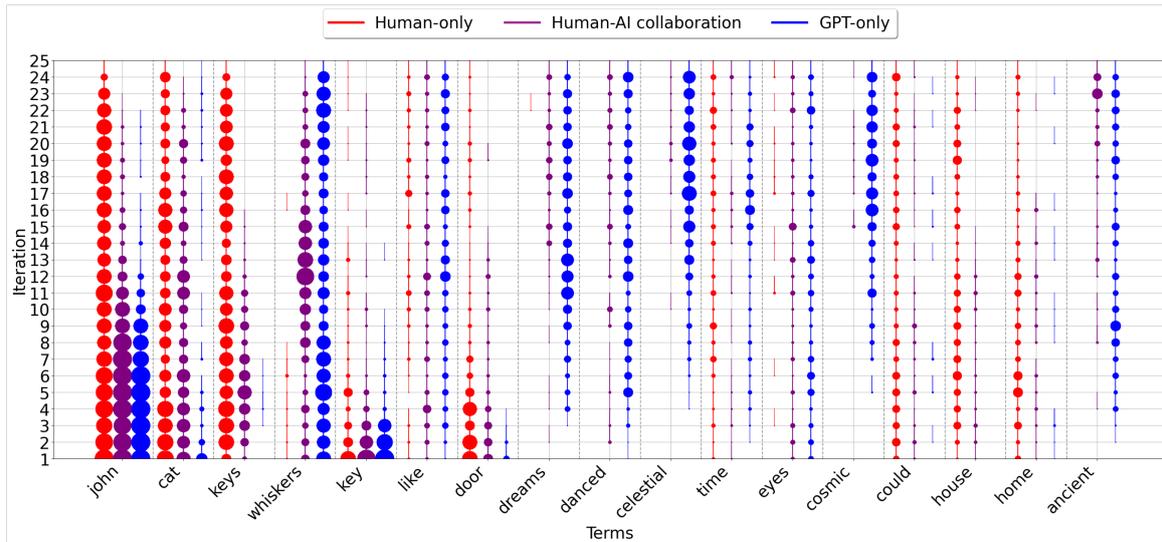


図 4: 条件別の語彙動態：横軸に単語，縦軸に世代（イテレーション）をプロット．座標 (x, y) の円は，イテレーション y において単語 x が物語中で使用されたことを示し，線は連続して使用されたことを表す．円の大きさは，そのイテレーションにおける単語の出現頻度を示す．

している．

これらの限界にもかかわらず，本手法の枠組み自体は，持続的な相互作用，進化する社会構造，参加者の主体性を取り入れることで，より生態学的妥当性の高いダイナミクスに拡張することが可能である．また，音楽や視覚芸術といった非言語的なオープンエンドな創造的課題への応用も検討できる．今回の研究は，ハイブリッドな実験的ソーシャルネットワークの初期的な探究に過ぎないが，今後の認知科学および計算機科学における人間と AI の融合的協働の研究に向けた大きな可能性を示唆している．

参考文献

- [1] Ziv Epstein, Aaron Hertzmann, Investigators of Human Creativity, Memo Akten, Hany Farid, Jessica Fjeld, Morgan R Frank, Matthew Groh, Laura Herman, Neil Leach, et al. Art and the science of generative ai. *Science*, Vol. 380, No. 6650, pp. 1110–1111, 2023.
- [2] Katherine M Collins, Ilia Sucholutsky, Umang Bhatt, Kartik Chandra, Lionel Wong, Mina Lee, Cedegao E Zhang, Tan Zhi-Xuan, Mark Ho, Vikash Mansinghka, et al. Building machines that learn and think with people. *Nature Human Behaviour*, Vol. 8, No. 10, pp. 1851–1863, 2024.
- [3] Levin Brinkmann, Fabian Baumann, Jean-François Bonnefon, Maxime Derex, Thomas F Müller, Anne-Marie Nussberger, Agnieszka Czaplicka, Alberto Acerbi, Thomas L Griffiths, Joseph Henrich, et al. Machine culture. *Nature Human Behaviour*, Vol. 7, No. 11, pp. 1855–1868, 2023.
- [4] Milena Tsvetkova, Taha Yasseri, Niccolo Pescetelli, and Tobias Werner. A new sociology of humans and machines. *Nature Human Behaviour*, Vol. 8, No. 10, pp. 1864–1876, 2024.
- [5] Michelle Vaccaro, Abdullah Almaatouq, and Thomas Malone. When combinations of humans and ai are useful: A systematic review and meta-analysis. *Nature Human Behaviour*, pp. 1–11, 2024.
- [6] Anil R Doshi and Oliver P Hauser. Generative ai enhances individual creativity but reduces the collective diversity of novel content. *Science Advances*, Vol. 10, No. 28, p. eadn5290, 2024.
- [7] Byung Cheol Lee and Jaeyeon Chung. An empirical investigation of the impact of chatgpt on creativity. *Nature Human Behaviour*, Vol. 8, No. 10, pp. 1906–1914, 2024.
- [8] Barrett R Anderson, Jash Hemant Shah, and Max Kreminski. Homogenization effects of large language models on human creative ideation. In

Proceedings of the 16th conference on creativity & cognition, pp. 413–425, 2024.

- [9] Han Liu, Vivian Lai, and Chenhao Tan. Understanding the effect of out-of-distribution examples and interactive explanations on human-ai decision making. *Proceedings of the ACM on Human-Computer Interaction*, Vol. 5, No. CSCW2, pp. 1–45, 2021.
- [10] Valerie Chen, Q Vera Liao, Jennifer Wortman Vaughan, and Gagan Bansal. Understanding the role of human intuition on reliance in human-ai decision-making with explanations. *Proceedings of the ACM on Human-computer Interaction*, Vol. 7, No. CSCW2, pp. 1–32, 2023.
- [11] Gagan Bansal, Tongshuang Wu, Joyce Zhou, Raymond Fok, Besmira Nushi, Ece Kamar, Marco Tulio Ribeiro, and Daniel Weld. Does the whole exceed its parts? the effect of ai explanations on complementary team performance. In *Proceedings of the 2021 CHI conference on human factors in computing systems*, pp. 1–16, 2021.
- [12] Yunfeng Zhang, Q Vera Liao, and Rachel KE Bellamy. Effect of confidence and explanation on accuracy and trust calibration in ai-assisted decision making. In *Proceedings of the 2020 conference on fairness, accountability, and transparency*, pp. 295–305, 2020.
- [13] Jason W Burton, Ezequiel Lopez-Lopez, Shahar Hechtlinger, Zoe Rahwan, Samuel Aeschbach, Michiel A Bakker, Joshua A Becker, Aleks Berditchevskaia, Julian Berger, Levin Brinkmann, et al. How large language models can reshape collective intelligence. *Nature Human Behaviour*, Vol. 8, No. 9, pp. 1643–1655, 2024.
- [14] Damon Centola. The network science of collective intelligence. *Trends in Cognitive Sciences*, Vol. 26, No. 11, pp. 923–941, 2022.
- [15] Thomas W Malone and Michael Bernstein. *Handbook of collective intelligence*. MIT Press, 2022.
- [16] Damon Centola and Andrea Baronchelli. The spontaneous emergence of conventions: An experimental study of cultural evolution. *Proceedings of the National Academy of Sciences*, Vol. 112, No. 7, pp. 1989–1994, 2015.
- [17] Maxime Derex, Marie-Pauline Beugin, Bernard Godelle, and Michel Raymond. Experimental evidence for the influence of group size on cultural complexity. *Nature*, Vol. 503, No. 7476, pp. 389–391, 2013.
- [18] David G Rand, Samuel Arbesman, and Nicholas A Christakis. Dynamic social networks promote cooperation in experiments with humans. *Proceedings of the National Academy of Sciences*, Vol. 108, No. 48, pp. 19193–19198, 2011.
- [19] Hirokazu Shirado and Nicholas A Christakis. Locally noisy autonomous agents improve global human coordination in network experiments. *Nature*, Vol. 545, No. 7654, pp. 370–374, 2017.
- [20] James C Kaufman and Robert J Sternberg. *The Cambridge handbook of creativity*. Cambridge University Press, 2010.
- [21] Joy Paul Guilford. The nature of human intelligence. *New York: Macgraw Hill*, 1967.
- [22] E Paul Torrance. Torrance tests of creative thinking. *Educational and psychological measurement*, 1966.
- [23] Roger E Beaty and Yoed N Kenett. Associative thinking at the core of creativity. *Trends in cognitive sciences*, Vol. 27, No. 7, pp. 671–683, 2023.
- [24] Yuval Hart, Avraham E Mayo, Ruth Mayo, Liron Rozenkrantz, Avichai Tendler, Uri Alon, and Lior Noy. Creative foraging: An experimental paradigm for studying exploration and discovery. *PloS one*, Vol. 12, No. 8, p. e0182133, 2017.
- [25] Yuval Hart, Hagar Goldberg, Ella Striem-Amit, Avraham E Mayo, Lior Noy, and Uri Alon. Creative exploration as a scale-invariant search on a meaning landscape. *Nature Communications*, Vol. 9, No. 1, p. 5411, 2018.
- [26] Aaron Hurst, Adam Lerer, Adam P Goucher, Adam Perelman, Aditya Ramesh, Aidan Clark, AJ Ostrow, Akila Welihinda, Alan Hayes, Alec Radford, et al. GPT-4o system card. *arXiv preprint arXiv:2410.21276*, 2024.

- [27] Peter Harrison, Raja Marjeh, Federico Adolfi, Pol van Rijn, Manuel Anglada-Tort, Ofer Tchernichovski, Pauline Larrouy-Maestri, and Nori Jacoby. Gibbs sampling with people. *Advances in neural information processing systems*, Vol. 33, pp. 10659–10671, 2020.
- [28] Wenhui Wang, Furu Wei, Li Dong, Hangbo Bao, Nan Yang, and Ming Zhou. Minilm: Deep self-attention distillation for task-agnostic compression of pre-trained transformers. *Advances in Neural Information Processing Systems*, Vol. 33, pp. 5776–5788, 2020.
- [29] John R Hayes. Modeling and remodeling writing. *Written Communication*, Vol. 29, No. 3, pp. 369–388, 2012.