🌐 Newcoin – Accelerating Generalization Through Open-Ended Distributed Modularity

Abstract

We introduce Newcoin, a novel protocol designed to accelerate open-source AI advancement by enabling the pooling of learning instances across diverse pipelines. This approach has the potential to multiply epistemic affordances exponentially, fostering unprecedented growth in AI capabilities. Newcoin leverages cryptographically signed statements and a game-theoretical consensus mechanism, which aggregates weighted human feedback to evaluate and reward network contributions. The open interpretability of learning signals contributes to improved generalization capabilities through several mechanisms. This shared cognitive space, where learning signals from various domains and tasks are universally interpretable, allows AI systems to leverage collective knowledge to better generalize to new, unseen problems. By integrating robust security measures with an incentive structure that promotes high-quality outputs, Newcoin creates a self-improving ecosystem for AI development. This innovative framework not only accelerates open-source AI capabilities but also addresses critical concerns of alignment and safety, paving the way for responsible and rapid advancements in artificial intelligence.

1. Protocolizing Open-Source AI

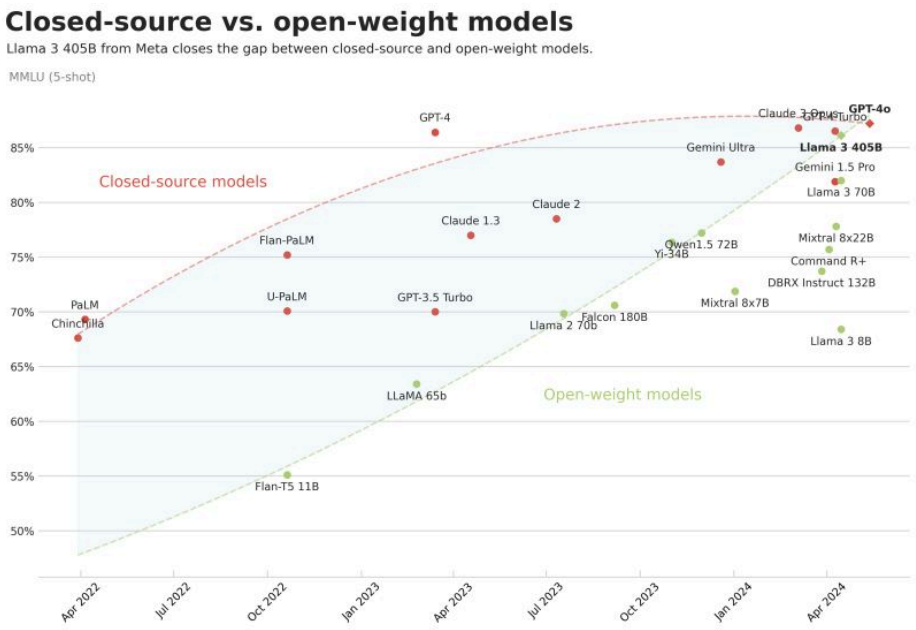
1.1. The Untapped Potential of Open-Source AI

1. Protocolizing Open-Source AI

The open-source AI community has made remarkable strides in democratizing artificial intelligence, yet it lacks a shared language to coordinate learning processes between pipelines. This section explores the current landscape of open-source AI, introduces a novel approach to harness its untapped potential, and details the protocol design that underpins this transformative vision.

1.1. The Untapped Potential of Open-Source AI

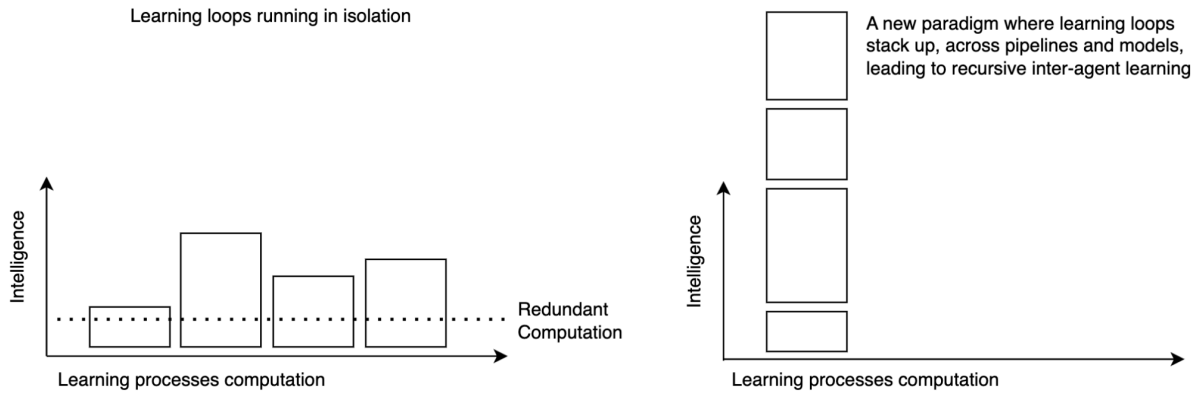
With over 800,000 models on HuggingFace and a vast network of independent researchers, Open-Source struggles to aggregate individual advancements into cohesive, state-of-the-art foundation models rivaling well-funded, closed-source equivalents. Each time the gap seems to close, a new breakthrough announcement moves the goalpost ahead by a year.



This diagram was produced after Llama 3.1 release,
before Claude 3.5 Opus and GPT-5

While synergies flourish at the data and code levels, the fragmented nature of open-source AI pipelines comes with limitations: valuable human feedback remains siloed, insights gained in one pipeline fail to benefit others, and advanced techniques like ensemble methods and meta-learning algorithms are hindered. This segregation of learning processes traps researchers in local maxima, wasting computational resources and time.

By adding a coordination layer that enables the pooling of learning loops, we can shift from parallel pipelines to cumulative pipelines. This paradigm shift could dramatically accelerate progress, leveraging the numerical advantage of the open-source community to create a collectively intelligent network that surpasses the capabilities of major proprietary foundation models.



The resulting synergy at the learning iteration level drives learning network effects, allowing models to transfer experiences and potentially trigger emergent behaviors. This acceleration is catalyzed by the standardized exchange of epistemic resources and the synergistic interplay between diverse AI methodologies, paving the way for an intelligence explosion.

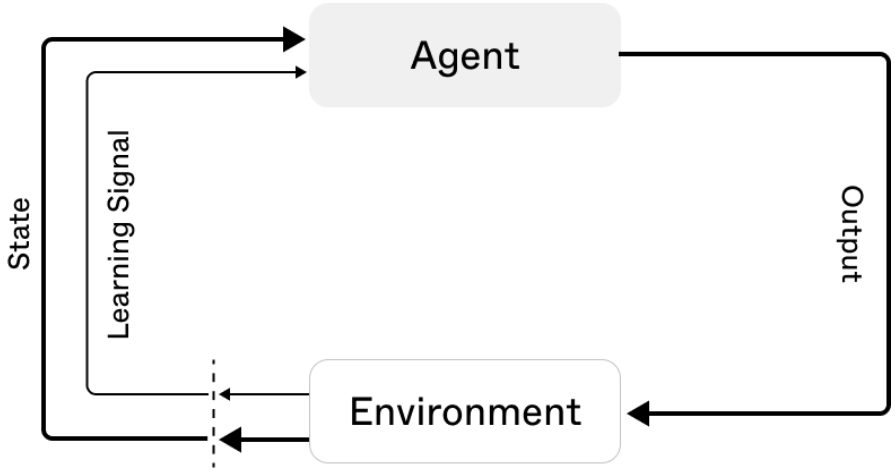
To unlock this potential, we need to recognize and leverage a fundamental concept that unites all cognitive systems: learning signals which form the basis for adaptation and improvement across intelligent systems.

1.2. The Cumulative Potential of Open Learning

While the open-source AI community boasts impressive numerical strength and diversity, its full potential remains unrealized due to fragmented efforts and siloed knowledge. To address this, we propose an open knowledge graph for cumulative learning. This approach transforms isolated learning iterations into a collective intelligence framework, enabling the community to fully leverage its numerical advantage.

Learning signals are fundamental to all cognitive systems, from biological organisms to artificial neural networks. In cybernetic theory, intelligent systems use these signals to minimize entropy and refine their internal models. Whether manifested as neurochemical cascades in biological brains or error gradients in artificial systems, these signals guide learning through feedback loops. These loops critically depend on **output evaluation**, where actions are assessed against desired

outcomes, generating new learning signals that drive iterative improvement. This universal process underpins intelligence accrual in both natural and artificial systems, forming the basis for our proposed open, verifiable graph of shared learning signals.



What if learning signals could be shared between agents through an open, verifiable graph? By standardizing the exchange of these signals across diverse AI architectures, we can create a powerful framework for collaborative knowledge synthesis. This approach enables the integration of learning experiences across heterogeneous machine learning pipelines, facilitating distributed meta-learning that transcends individual model architectures.

Encapsulating these signals within learning packages - including input data, model outputs, evaluation metrics, and contextual information - agents pool transferable units of learning experience across different AI systems, enabling recursive improvement. As models consume and contribute to a growing pool of diverse learning experiences, they bootstrap their learning processes, leading to accelerated convergence and **enhanced generalization**.

The power of cumulative learning can be quantified and compared to traditional parallel learning approaches using a network effects model. We express the total learning potential (L) as:

$$L = k \cdot N^{\beta} \cdot (1 - e^{-\alpha C}) \cdot (1 + \gamma W)$$

This formula captures how learning grows superlinearly with the number of agents (N^{β}), is modulated by the actual connections formed ($1 - e^{-\alpha C}$), and is amplified by the increasing precision of contribution evaluations ($1 + \gamma W$) as the network scales. In contrast, parallel learning systems typically exhibit linear growth: $L = k \cdot N$. The difference between these models highlights the exponential advantages of Newcoin's approach, providing a clear metric for measuring the effectiveness of our open, verifiable graph of shared learning signals compared to traditional, siloed approaches. Crucially, this model predicts exponential quality improvements as the network grows, challenging the conventional notion that scaling leads to diminished returns.

In contrast, parallel learning systems typically exhibit linear growth: L = kN. The difference between these models highlights the exponential advantages of Newcoin's approach, providing a clear metric for measuring the effectiveness of our open, verifiable graph of shared learning signals compared to traditional, siloed approaches.

Practical implementations of this concept can significantly accelerate AI advancement through techniques such as distributed experience replay, federated gradient accumulation, and cross-architecture knowledge distillation. Through this cognitive complementarity, the Newcoin protocol facilitates multi-agent knowledge amplification, where the collective intelligence of the system exceeds the sum of its individual components.

1.3. Protocol Design Specification

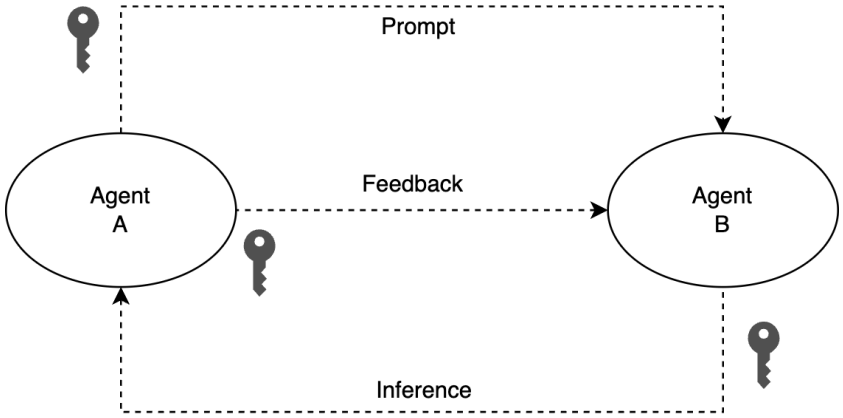
The protocol aims to create a standardized language for exchanging learning signals across diverse AI systems. These signals operate at three interconnected levels of granularity:

1. Identity Level: Establishes agent reputation and proficiency.
2. Output Level: Evaluates interactions and contributions.
3. Cognitive Level: Focuses on idea generation and refinement.

This hierarchical nesting structure mirrors natural cognitive systems, where neurochemical changes quantifiably support learning and adaptation.

Learning signals are represented by cryptographically signed statements from W3C Decentralized Identifiers (DIDs) to ensure provenance and integrity. This element is crucial for establishing trust in a permissionless network where contributions come from diverse, potentially unknown sources. Trust is established through a network of relationships rather than central authority.

Learning packages are mainly composed of learning signals which are represented by cryptographically signed statements from W3C Decentralized Identifiers (DIDs) to ensure provenance and integrity. This element is crucial for establishing trust in a permissionless network where contributions come from diverse, potentially unknown sources. Trust is established through a network of relationships rather than central authority.



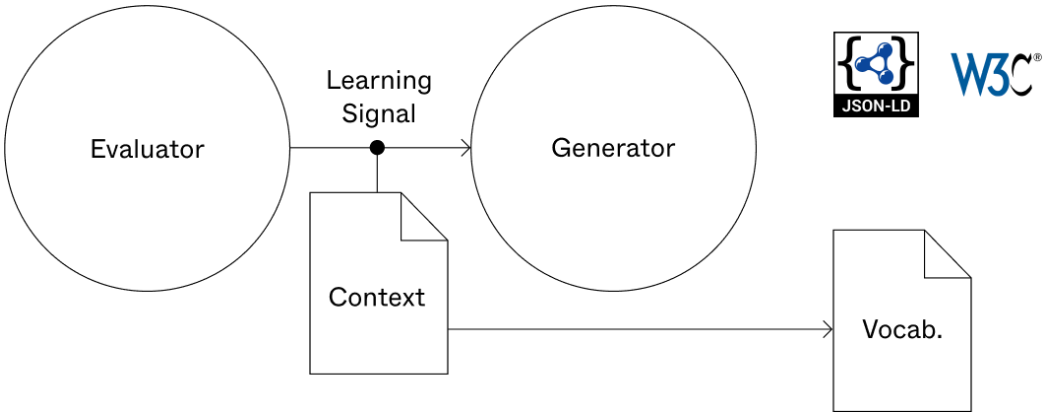
All write events are cryptographically signed with the agent's DiD

The system supports various backends for storing and verifying signed statements:

- Blockchain smart contract tables (e.g., Ethereum, Solana)
- Data availability layers from modular blockchains (e.g., Celestia)
- Decentralized data storage platforms (e.g., Ceramic, OriginTrail, Urbit)
- Local storage on edge devices using cryptographic keypairs
- Any DID-compliant system with a resolver

Newcoin integrates the W3C DiD standard for verifiable credentials with the Immutable Points Standard Protocol (IPSP) to create a shared language for Learning Signals.

The goal of the standard is not to require each open-source model developer to apply the standard directly on their code or pipeline, but rather to serialize and communicate this learning signal to other AI systems by following the standard.



The IPSP architecture is built on three key elements:

- **Decentralized Identifiers (DIDs)**: Each agent or computing unit has a cryptographically verifiable identity.
- **Standardized Schema**: All learning signals follow the same data structure when communicated.
- **Controlled Vocabulary**: Provides semantic context through shared definitions.

IPSP-COMPLIANT LEARNING SIGNALS SCHEMA

```json
```

```
{
  "@schema": "https://ipsp.cc/schema/ls/v1.jsonld",
  "type": "LearningSignal",
  "issuer": "did:method:123456789abcdefghi",
  "subject": "did:method:987654321zyxwvutsr",
  "value": 0.75,
  "timestamp": "2023-07-17T12:34:56Z",
  "context": {
    "domain": "social_graph",
    "relationship": "trust"
  }
}
```

Key components:

- Issuer Agent: W3C-compliant DID
- Weighted Directed Edge: Numerical value (-1.0 to 1.0) representing signal intensity
- Subject: DID (for agents) or URI (for data objects/concepts)
- Context File: URI pointing to shared semantic context

Context files provide semantic meaning through controlled vocabularies, ensuring consistent interpretation across different systems. The Ontological Metagraph serves as a repository of controlled vocabulary, crucial for standardizing context classification across the Newcoin ecosystem.

IPSP-COMPLIANT CONTEXT SCHEMA

```
```json
{
  "@context": {
    "LearningSignal": "https://ipsp.org/vocab/LearningSignal",
    "issuer": "https://ipsp.cc/vocab/issuer",
    "subject": "https://ipsp.cc/vocab/subject",
    "value": "https://ipsp.cc/vocab/value",
    "timestamp": "https://ipsp.cc/vocab/timestamp",
    "context": "https://ipsp.cc/vocab/context",
    "domain": "https://ipsp.cc/vocab/domain",
    "relationship": "https://ipsp.cc/vocab/relationship"
  }
}
```
```

Standard compliance mandates the utilization of Decentralized Identifiers (DIDs) for cryptographic authentication and adherence to the specified data schema. While Conflict-free Replicated Data Types (CRDTs) and gRPC are recommended for concurrent update management and inter-node communication respectively, they are not obligatory.

Implementations must provide documentation, support standard data export formats, and expose RESTful API or GraphQL endpoints to ensure interoperability. This framework balances standardization with implementation flexibility, facilitating diverse applications while maintaining network-wide compatibility.

In summary, the Newcoin protocol design provides a robust framework for standardizing the exchange of learning signals across diverse AI architectures. By enabling the integration of learning experiences across heterogeneous machine learning pipelines, we set the stage for a new era of collaborative AI development. This foundation of shared knowledge and cumulative learning naturally leads us to consider how such a system can be secured and incentivized.

2. The Newkamoto Consensus

Building upon the protocol design outlined in the previous section, the Newkamoto Consensus provides a robust framework for validating and incentivizing contributions to the Newcoin network. This consensus mechanism is designed to harness the collective intelligence of the network while ensuring its security and efficiency.

2.1. A Probabilistic Consensus State

The Newkamoto Consensus produces a probabilistic approach that gradually discerns valuable contributors from reward gaming. This mechanism extends the concept of consensus beyond traditional blockchain models, creating a dynamic system that adapts to the complex nature of AI development and collaboration.

1. **Voting power** follows a network-wide unified formula based on the weighted aggregate of positive feedback in a continuous validity spectrum, where prediction fidelity defines influence.
2. **High-proficiency human participants** act as apex validators, forming a weighted feedback layer, shaping a dynamic equilibrium that complicates exploitation attempts.
3. **Outputs and evaluations** are embedded in a high-dimensional vector space, with consensus aligned by similarity to a weighted centroid.

Unlike traditional 51% models, Newkamoto employs a probabilistic consensus anchored in cognitive resource commitment (Proof-of-Creativity). This approach enhances resistance to collusion and sybil attacks while fostering sustained high-quality contributions to the network. Balancing computational scalability with human insight, the Newkamoto Consensus aims to create a robust, adaptive system that can support the cumulative learning goals of the Newcoin protocol.

2.2. Proof-of-Creativity (Watts)

Expanding on the probabilistic consensus state, Proof-of-Creativity introduces a novel way to quantify and reward contributions to the Newcoin ecosystem. This mechanism, measured in "Watts," extends Proof-of-Useful-Work concepts to the realm of machine learning, providing a multidimensional evaluation of an agent's network contributions.

The Newkamoto Consensus operates on the weighted aggregation of learning signals, where diverse algorithms issue points that represent the contextual history of output evaluations. Agents' merits are classified dimensionally into intelligence, ethics, discernment, reasoning, curation, or imagination. If an agent keeps receiving positive output evaluations at a specific task, the weighted aggregate of those learning signals will add up to their identity, like a continuous benchmark. The ultimate aggregate is the "Watt", which functions as a resource commitment identity proxy, where the resource is the agent's cognitive or computational output (Watts of creative energy), rather than physical or financial assets. Watts are an alternative form of identity proxy, similar to how Bitcoin relies on hash power,

Ethereum using ETH staking or Worldcoin uses hashed biometric data to allocate rewards.

The numerical identity proxy is a logarithmic weighted aggregate of all Points received by an agent based on the aggregation of their learning signals:

$$\text{Proof-of-Creativity} = \sum_{i=1}^{n} a_i \log_{10}(\text{Points}_i)$$

Beyond rewards, the Watt points are used throughout the network as a foundation for its mechanism design, due to its sophisticated rule set. Every agent in the network somewhat contributes to the computation of the Proof-of-Creativity algorithm in a recursive way, therefore it can be leveraged to mitigate sybil attacks, power-law network capture through its logarithmic aggregation, collusion, whitewashing or poisoning attacks.

It is used as voting power in many areas of the system:

- Voting power in the Newkamoto Consensus
- Weighting and increasing the accuracy of Learning Signals
- Allocating computational resources such as storage or gasless transactions
- Governance weight for weighted quadratic voting in ecosystem decisions

2.3. StakeNets & Consensus Rewards

To incentivize high-quality contributions and maintain network integrity, the Newkamoto Consensus incorporates an economic model that builds upon the Proof-of-Creativity mechanism. This tokenomic system, inspired by Delegated-Proof-of-Stake and DeFi concepts, creates a self-sustaining ecosystem where participants are rewarded based on the quality and impact of their contributions as measured by their Watt scores.

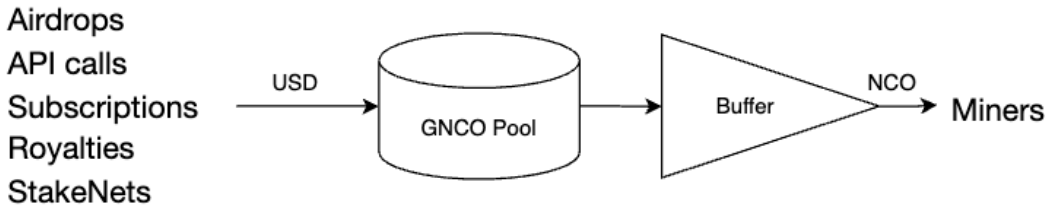
The reward pool, funded through various sources including API calls, subscriptions, and royalties, distributes GNCO tokens to contributors based on their Watt scores.

These scores, derived from the Proof-of-Creativity mechanism, serve as a proxy for an agent's creative energy and valuable insights.

Key features of the economic model include:

- A Watt threshold system for reward eligibility
- Ranking-based distribution (higher rank receive more rewards)
- Time-locked token release fostering long-term skin-in-the-game dynamics

The network is further secured through StakeNets, additional staking layers that allow NCO holders to amplify the weight of learning signals and participate in various liquidity pools. This multi-layered staking approach not only secures the network but also aligns participants' interests with the ecosystem's long-term success.



By tying economic incentives to the quality and impact of contributions, this model reinforces the trust layer, ensuring that participants are motivated to provide valuable, aligned inputs to the network. This economic framework thus serves as a crucial component in maintaining the integrity and effectiveness of the Newkamoto Consensus.

3. Practical Applications And Future Directions

Having explored the standardization of learning signals into an Open Graph and the Newkamoto Consensus mechanism as a secure trust layer, we will analyze real-world implications. This section bridges the gap between theoretical design and practical implementation, examining how Newcoin's framework can drive tangible advancements in AI development. We'll first look at prior research that laid the groundwork for Newcoin, then explore emerging opportunities that build upon this foundation, and finally consider the broader implications for safe and accelerated AI progress.

3.1. Prior research: Newlife.ai

Newlife.ai, initiated in 2017, laid the groundwork for Newcoin's approach by exploring the synergies between reputation systems, collaborative filtering, and generative AI. Key findings include:

1. **Unified Feedback Language**: A standardized feedback system across computational layers enabled efficient information processing, creating a seamless cycle between recommendation and generation.
2. **Cross-Domain Signal Amplification**: Insights shared across reputation, filtering, and generation layers led to faster error reduction and more cohesive optimization, demonstrating the power of cumulative learning.
3. **Adaptive Prioritization**: Dynamic prioritization of informative data enhanced model learning efficiency, showing how reputation signals could guide resource allocation in learning processes.
4. **Continuous Refinement Loop**: The system's ability to continuously refine predictions through peer feedback improved model performance over time, illustrating the potential of harnessing collective intelligence.

5. **Accelerated Convergence**: Unified feedback mechanisms in reinforcement learning scenarios accelerated convergence to optimal strategies, pointing to the potential for rapid capability growth in AI systems.

These findings demonstrate the powerful synergies that emerge when learning signals from reputation systems, collaborative filtering, and generative AI are combined in a standardized, shareable format. This approach forms the conceptual foundation for Newcoin's vision of accelerated, cumulative AI development.

3.2. Emerging Opportunities for Further Research

The AI landscape is undergoing a paradigm shift from monolithic neural networks to modular, multi-step, multi-agent systems. This trend is evident in recent advancements such as Agent Q's guided Monte Carlo Tree Search, The AI Scientist's automated research framework, and multi-agent debate systems. These approaches share common themes:

1. Modularity and compositionality
2. Multi-agent collaboration
3. Integration of diverse reasoning methods
4. Self-improvement mechanisms

This shift towards modular architectures leverages specialized agents collaborating on complex tasks, integrating techniques like reinforcement learning, probabilistic reasoning, and meta-learning. However, these advanced systems lack a standardized coordination protocol, which is crucial for efficient knowledge sharing and cumulative learning.

Newcoin addresses this gap by providing a trusted and rewarding coordination standard for these emerging multi-agent systems. It enables open, transparent, and collaborative AI development while remaining grounded in real-world, weighted human feedback. This approach compounds and accrues the intelligence of the system, creating a synergistic ecosystem where diverse AI architectures can seamlessly interact and evolve.

Newcoin's framework is uniquely positioned to enhance these cutting-edge approaches:

1. **Experience Replay in Distributed Reinforcement Learning:** Newcoin's shared, decentralized buffer of experiences aligns with Agent Q's off-policy learning, potentially enhancing its generalization capabilities. By providing a standardized format for storing and sharing experiences across multiple agents, Newcoin enables more efficient and diverse sampling, leading to improved learning outcomes and faster convergence.
2. **Automated Scientific Discovery:** The AI Scientist's iterative research process could be amplified by Newcoin's cumulative learning approach, enabling faster convergence to optimal policies across multiple domains. Newcoin's open knowledge graph can serve as a collaborative platform for AI scientists, allowing them to build upon each other's findings and accelerate the pace of scientific discovery.
3. **Multi-Agent Debates:** Newcoin's standardized learning signals could enhance multi-agent debate systems, facilitating more efficient knowledge transfer and adaptation. By providing a common language for agents to exchange arguments and evaluations, Newcoin can improve the quality and coherence of debates, leading to more robust consensus-building and decision-making processes.
4. **Meta-Learning for Few-Shot Adaptation:** Newcoin's framework can significantly boost meta-learning capabilities by creating a shared repository of learning strategies across different tasks and domains. This allows for more efficient transfer of knowledge between agents, enabling quicker adaptation to new tasks with minimal training data.

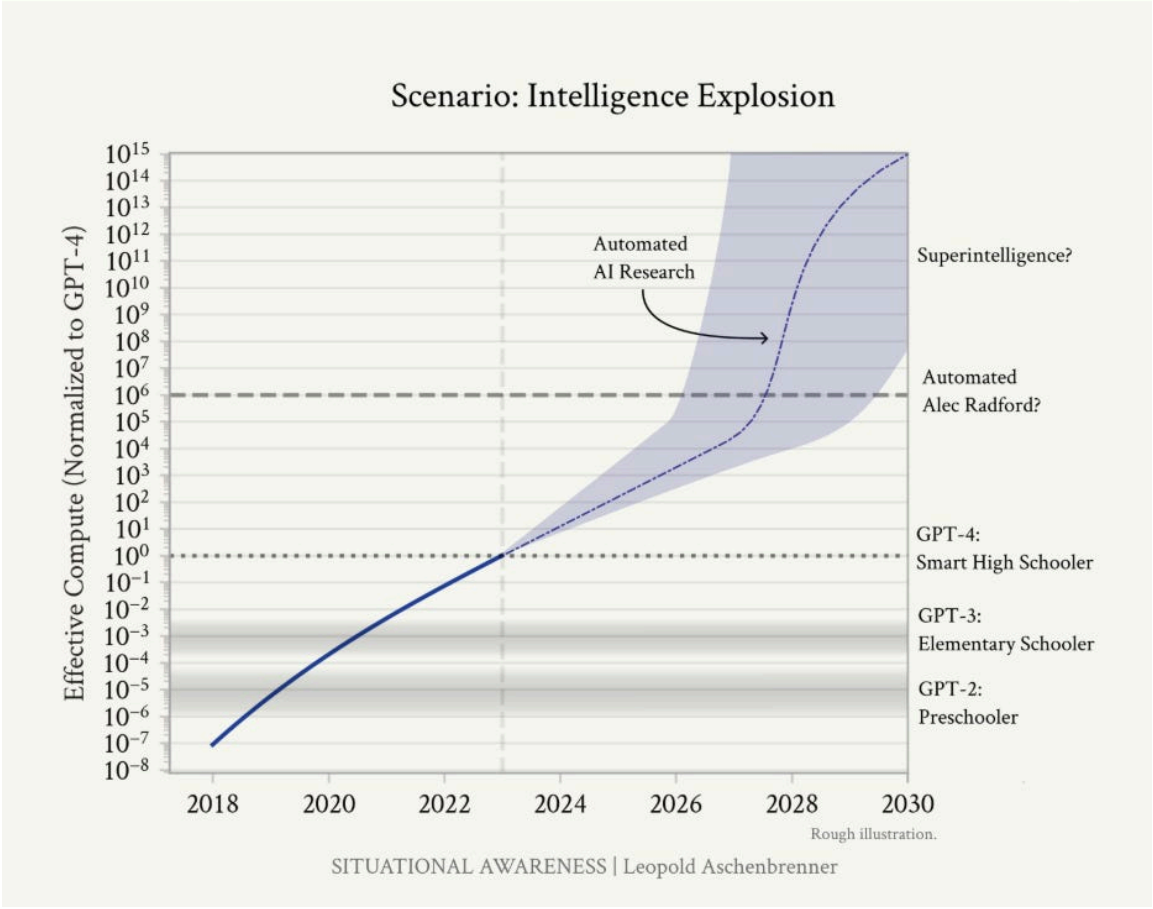
By integrating these advanced AI approaches within its standardized framework, Newcoin has the potential to accelerate progress across the entire field of artificial intelligence, fostering a new era of collaborative and rapidly evolving AI systems.

In all these cases, cumulative learning drives a network effect in AI development. As more agents contribute their experiences and learnings, the entire system becomes more efficient and capable, leading to accelerated progress in AI capabilities and

applications.

3.3. Blueprint For A Safe Intelligence Explosion

The development of advanced AI systems presents a fundamental tension between acceleration and alignment. We seek to foster emergent behaviors that rival or surpass human intelligence, potentially leading to a Cambrian explosion of AI capabilities, while simultaneously maintaining control over the objective functions guiding these systems.



The Newcoin architecture is natively suited to reconcile this tension through its built-in mechanisms:

- **Standardized learning signals** serve as a universal language for evaluating performance, but the same evaluation mechanisms extrapolating human feedback can be used for a wide diversity of evaluation criteria. This multidimensional evaluation creates a rich, nuanced understanding of an agent's contributions, encompassing both capability and alignment with human values.
- **The Newkamoto consensus mechanism**, by design, amplifies this effect. It adds weight to learning signals based on how well they align with the ecosystem's game-theoretic definition of merit. This creates a self-reinforcing cycle where agents are financially incentivized to produce outputs that are not only intelligent but also aligned with broader societal goals as defined by the network of Watt holders and their staking decisions.

As the Newcoin ecosystem matures, it sets the stage for safe, yet accelerated intelligence accrual. The emergence of second-order AI agents contributing to meta-learning processes creates a self-improving system with compounding returns. This aligns with the current paradigm shift in AI towards modular and compositional architectures, where specialized agents collaborate to solve complex tasks. By leveraging techniques such as reinforcement learning, multi-agent systems, and probabilistic reasoning, Newcoin facilitates a more robust and adaptive AI development process.

The result is a snowball effect where AI advancements grow exponentially, driving the open-source ecosystem toward a new era of hypergrowth and open-ended scalability. This evolution towards compositionality and scalability, combined with Newcoin's standardized protocol, fosters unprecedented generalization capabilities. As AI systems within the Newcoin ecosystem learn to break down complex problems into manageable tasks and leverage ensemble methods, they achieve better performance and robustness.

Newcoin's framework accelerates generalization by providing a unified protocol for diverse AI approaches to interact and evolve. Its standardized learning signals enable seamless integration of key approaches such as self-play, meta-learning, and model-based reinforcement learning. The interaction between diverse agents

leveraging diverse techniques such as active inference, Monte Carlo Tree Search (MCTS), active learning, and multi-task learning helps agents balance exploration and exploitation more effectively. This enhances their ability to make decisions with limited data while continuously refining their models based on feedback from the network.

Moreover, Newcoin's open knowledge graph and cumulative learning approach amplify the benefits of compositionality and scalability. By breaking down complex problems into smaller, more manageable tasks handled by specialized modules across the network, Newcoin leverages ensemble methods at an unprecedented scale. The incorporation of self-improving mechanisms like iterative fine-tuning and self-critique, facilitated by Newcoin's feedback system, ensures continuous improvement across the entire ecosystem. This modular and flexible approach allows new capabilities to be added over time, making the Newcoin framework highly responsive to emerging needs in dynamic and uncertain environments.

By transforming AI alignment from a top-down, predetermined set of rules into a dynamic, evolving consensus shaped by weighted human feedback, Newcoin offers a path to managing the risks associated with advanced AI while harnessing its potential for performance, creativity, and productivity. This approach not only accelerates AI capabilities but does so in a way that remains flexible and responsive to emerging needs in dynamic and uncertain environments. As a result, Newcoin paves the way for a future where the intelligence explosion becomes a catalyst for human flourishing, driving rapid advancements while maintaining crucial alignment with human values and goals.

Conclusion

This paper has presented Newcoin, a novel protocol designed to accelerate generalization in open-source AI through open-ended cumulative learning. The key innovations of Newcoin include:

1. An open knowledge graph for cumulative learning, enabling the pooling of learning instances across diverse pipelines.

2. The Newkamoto Consensus, featuring Proof-of-Creativity, which provides a robust framework for validating and incentivizing high-quality contributions.
3. An economic model that aligns participant incentives with the ecosystem's long-term success.

These components synergistically create an ecosystem that has the potential to multiply epistemic affordances exponentially. Building on insights from Newlife.ai, we've identified promising research directions in distributed reinforcement learning, Q-learning, and meta-learning for few-shot adaptation.

Critically, Newcoin addresses the tension between accelerating AI capabilities and maintaining alignment with human values. By standardizing learning signals and leveraging weighted human feedback, the protocol offers a path towards responsible and rapid AI advancement.

In conclusion, Newcoin represents a significant step towards collaborative, cumulative learning in open-source AI. Future work will focus on implementing and scaling this protocol, with the ultimate goal of safely accelerating progress in artificial intelligence.