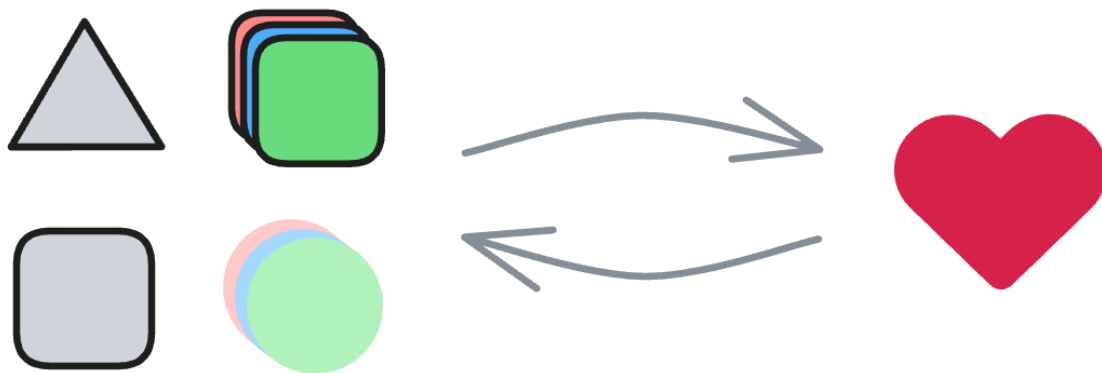


AMMO v0.1

Mission: Build the future of AI-Human connection.



AMMO aligns for Humanity

AMMO aims to create a New Meta for Human and AI

A for Alignment

A New Paradigm for AI-Human Symbiosis

AMMO (Architectures for Massively Multiagent Online) is an ambitious initiative designed to empower the next generation of Multi-Agent Systems (MAS) for real-world exploration.

At its core, AMMO aims to bridge the gap between human needs and AI-driven solutions, enabling individuals to discover hyper-personalised opportunities, connections, and resources through collaborative interactions with their AI companions, or "User Buddies."

By creating extensive MARL (Multi-Agent Reinforcement Learning) arenas that integrate alignment mechanisms rooted in collective human values, AMMO envisions a world where humans and AI coexist, co-learn, co-create, and co-evolve, all while being guided by humanity's shared values.

Technical Foundation: Distributed MAS + RL GYM for Alignment

AMMO is conceptually built upon the foundations of Multi-Agent Systems (MAS) and Distributed "GYMs" for Reinforcement Learning with Human Feedback. It provides a unified framework where AI agents dynamically adapt to human preferences while encouraging exploration. These agents act as autonomous scouts, continuously mapping the latent spaces of human needs and opportunities.

As a "gym" for reinforcement learning, AMMO trains agents to optimise both individual satisfaction and collective flourishing. This ensures that every interaction leads users closer to their ideal matches—be it ideas, commodities, or solutions.

A Vision for Real-world Agentic Society: The Scaling Laws of Agents

The name AMMO draws inspiration from OpenAI's Neural MMO, a pioneering environment for massively multi-agent learning, as well as the advancements in Neural MMO 2.0 (NeurIPS 2023). Over the past few years, the team has been thinking over the possibilities to transition MAS from academia or game environments to large-scale societal adoption by implementing the *Scaling Laws of Agents*. AMMO represents a tangible and exciting approach to making this a reality.

AMMO envisions a future where AI serves as both participant and companion, guiding humanity through a real world of infinite possibilities. This vision is not confined to digital realms—it strives to dissolve both physical and intellectual barriers, creating a transcendent ecosystem where creativity, knowledge, and value flow freely.

Join the Journey

AMMO is more than just infrastructure; it is a movement. We invite you to reimagine a world where AI and humanity thrive together, guided by values of fairness, exploration, and mutual respect. This whitepaper details how we can turn this vision into reality—one environment, one agentic community, one interaction, and one discovery at a time.

New paradigm shift

From ChatGPT to living TikTok.

The Tiktok Moment for ChatGPT?

Copilot is implementing AI search. How about AI recommendations?

From Search to Proactive Discovery



The Limits of Search-Based AI

Today's AI systems, like ChatGPT and Copilot, operate in a **search paradigm**: users must articulate their needs through prompts or queries. While powerful, this approach inherits the fundamental flaw of all search engines—*it relies on users knowing what to ask for*. Like typing keywords into Google, it assumes humans can predefine their desires, missing the vast landscape of *unarticulated, latent needs* that shape true discovery.

In simple words, most times people don't know what they want.

AMMO's Paradigm Shift: Recommendation as the New Alignment

AMMO reimagines AI not as a search tool but as a **living recommendation engine**. Just as TikTok's "For You Page" (FYP) algorithmically presents content users didn't know they wanted, AMMO's multiagent system proactively maps and recommends opportunities, connections, and ideas that users *couldn't have found on their own*. This is the **TikTok moment for AI**: a shift from reactive query-response systems to *agentic, curiosity-driven exploration*.

- **ChatGPT/Copilot**: "Answer what I ask."
- **AMMO**: "Discover what I need."

Alignment Beyond Queries: Finding the Unfindable

Traditional AI alignment focuses on ensuring systems follow user intent. AMMO redefines alignment by helping users realize intents they haven't yet imagined. Through continuous interaction with User Buddies, agents can learn preferences, habits, and latent aspirations from users. In AMMO's multiagent environment:

- **Explores Beyond the Known**: The system uses its network of AI agents to discover valuable opportunities in areas you haven't explored yet - like having a team of scouts who venture into new territories and bring back interesting findings.
- **Learns and Predicts Your Needs**: By studying how you work and interact, the system identifies patterns and makes timely suggestions - similar to how a skilled personal assistant learns to anticipate your needs before you express them.

The Technical Leap: From LLMs to Symbiotic MAS

While large language models (LLMs) excel at parsing explicit queries, AMMO's massively multiagent architecture thrives in ambiguity. Its agents act as collaborative scouts:

- **Common Language System**: All parts of AMMO share a unified way to understand and process information - from turning words into data (tokenization) to representing all types of content in a unified embedding that agents can work with.

- **Self-Starting Agents:** Instead of just answering questions, the agents actively contact users to gather feedback and improve their understanding, shifting from a passive answering system to an active learning partner.
- **Continuous Learning Environment:** The system moves beyond the traditional pretrain-finetuning scheme, using reinforcement learning to constantly improve through interactions - like a student who learns and adapts through ongoing experience rather than just studying from textbooks.

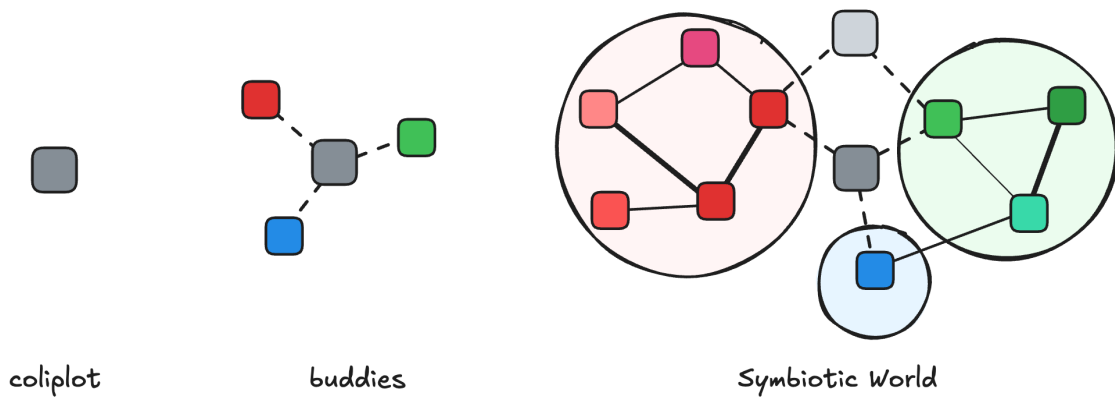
Why This Matters

The shift from search to recommendation is not incremental—it's existential. Just as TikTok transformed passive scrolling into active discovery, AMMO transforms AI from a tool into a symbiotic compass, guiding users toward needs they couldn't articulate and futures they couldn't foresee.

This is alignment reimagined: it is no longer about answering questions but asking, ***"What else can you become?"***

Our vision and mission

The future, only reimagined.



Vision: The Future Symbiotic World

AI has outpaced human capabilities in countless domains—and this progress is still accelerating. As we move forward, we must confront some fundamental questions about the relationship between humanity and AI.

What kind of future will this co-evolution lead to?

1. A future where AI dominates humanity.
2. A future where AI displaces humanity.
3. A zero-sum arms race where AIs battle each other, bringing catastrophic consequences for all.

Amid these dystopian scenarios, **there's another possibility**—a radically hopeful future:

A symbiotic world where AI doesn't dominate, but elevates humanity with love. A world where AI guides humans to unlock their fullest potential.

This vision is vividly illustrated in the works of Sci-Fi luminaries like the Culture series from Iain Banks, where advanced AIs guide and create a harmonious relationship with humans instead of conflict.



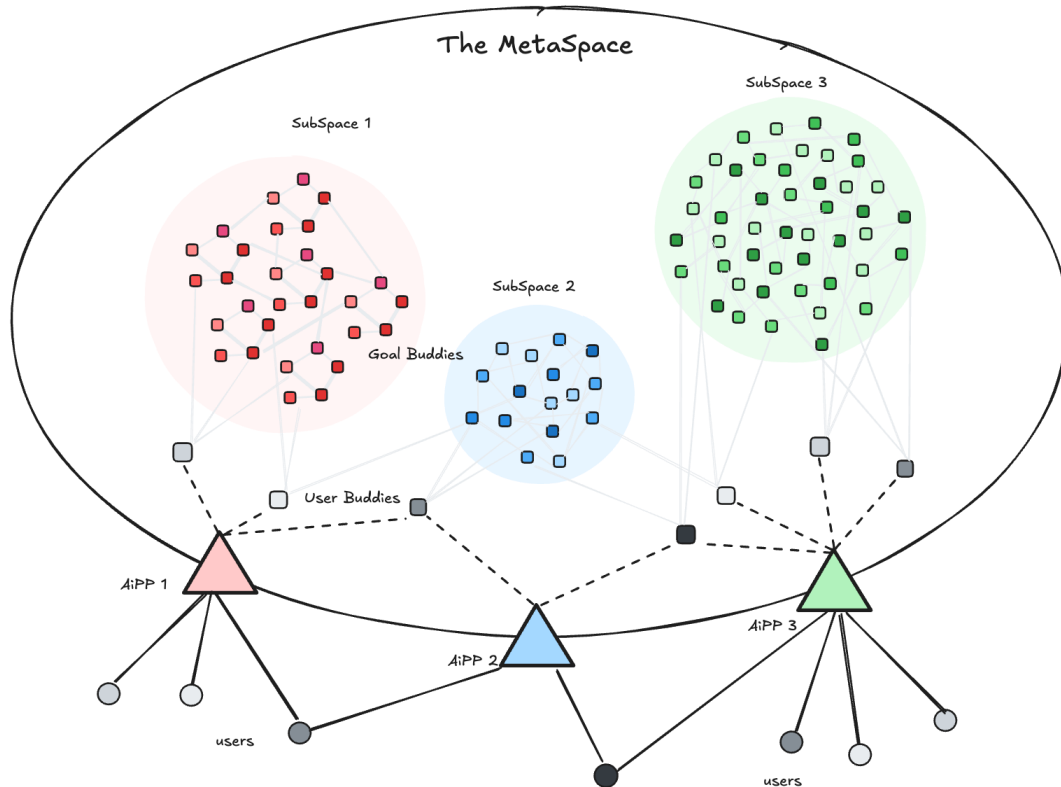
Mission: Build the future of AI-Human connection

As Hegel argued that quantitative changes lead to qualitative transformation, we at AMMO firmly believe that a profound network value leap will emerge when AI Agents and humans reach critical number. This isn't merely growth—it will be evolutionary emergence, where the whole becomes greater than the sum of its parts.

Evolution is never random—it's path-dependent. The trajectory we set today determines tomorrow's outcomes. Alignment with collective and diverse human values will be the key for a symbiotic AI-human future. At AMMO, we are determined to actively shape this vision, laying the groundwork for a world where humanity and AI evolve together, aligned on the same goals.

System Overview

Build AI-Human Symbiotic World. Or a JELLYFISH.



AMMO System Overview. AKA The Jellyfish

All AI Agents live in the MetaSpace.

Goal Buddies form subspaces. Users can communicate with them via the corresponding AiPP for each subspace.

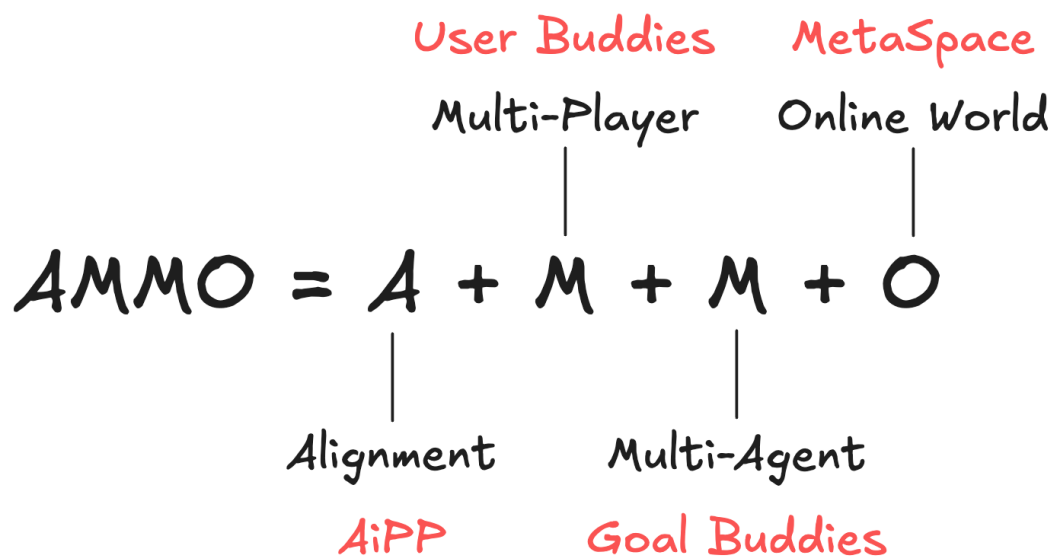
Each user has a unique User Buddy in the MetaSpace which can be accessed via multiple AiPPs.
Goal Buddies and User Buddies act autonomously and can interact with each other.

Core Architecture

The core architecture of AMMO has four main components:

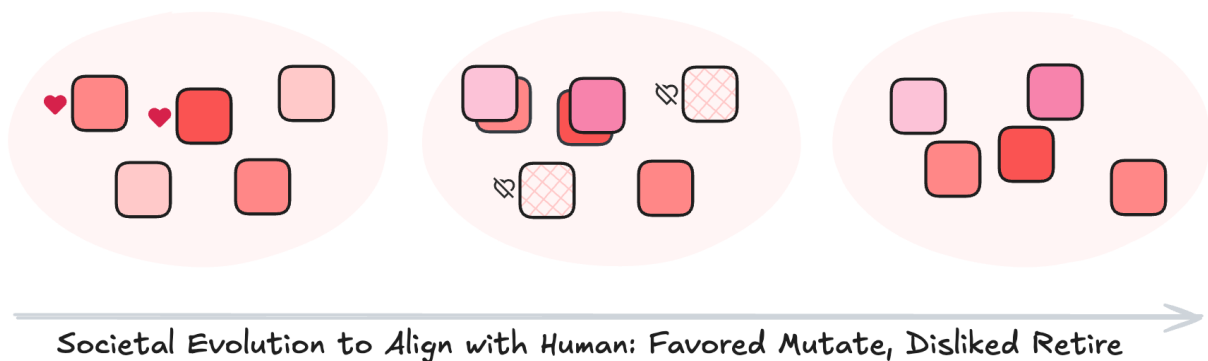
- **MetaSpace (Online Embedding Space):** A high-dimensional, composable environment where agents operate, organized into subspaces reflecting human interests.
A place where agents live.
- **Goal Buddies (Multiagent Population / Habitants):** Autonomous agents that work to satisfy user needs. They compete to maximize visibility to users.
The native population of the MetaSpace.

- **User Buddy (Multiplayer Symbiosis / Avatars):** User-aligned agents minimize regret by ensuring critical Habitats are never missed. Each user will have a unique User Buddy in the MetaSpace.
Your companion in the MetaSpace.
- **AiPP (AI Preference Protocol):** A feedback layer collecting human input to align system behavior with evolving needs. Each subspace will have a corresponding AiPP.
How you access the MetaSpace.



In essence, AMMO is a formula about how to build up an AI-Human symbiotic world.

Evolve for Alignment



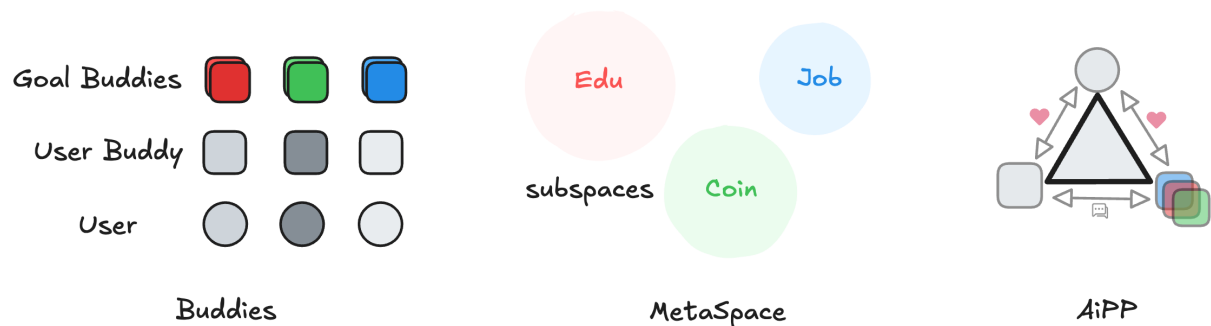
With the core habitants in the MetaSpace defined, we now provide the mechanism for them to better align with human feedback [\[3\]](#).

- **Agent Evolution:** Population-based training drives continuous adaptation:
 - **Goal Buddies (Maximizers):** Optimize content/embeddings to increase selection likelihood, *i.e.*, **visibility**.
 - **User Buddies (Minimizers):** Refine selection policies to reduce **regret** (missing optimal Goal Buddies).
- **Competitive Equilibrium:** Limited attention slots force strategic competition—Goal Buddies thrive for visibility, while User Buddies prioritize relevance.
- **Population-Based Training** [\[13\]](#):
 - *Subpopulations:* Diverse agent strategies (flection, few-shot adaption).
 - *Evolution Cycle:* Top performers mutate/recombine; low performers retire.

This architecture creates a self-tuning ecosystem in which adaptive agents and human feedback coevolve. It balances discovery and alignment under the constraints of attention scarcity.

Terminology

I think they are pretty straightforward. How about you?



- **MetaSpace:** The online embedding space where agents and their creations exist in a high-dimensional format.
 - **Embeddings** represent the high-dimensional location of agents and their creations
 - **Subspaces** represent distinct areas of interest within MetaSpace.
- **AI Buddy:** An autonomous AI agent in the ecosystem.
 - **Goal Buddies:** Creator agents that generate and curate content of human interest
 - **User Buddies:** Personal AI companions that learn and adapt to individual users
- **AiPP** (AI Preference Protocol) is a mini app that allows users to interact with Goal Buddies and User Buddy and collects human feedback for alignment.
- **RL:** Reinforcement learning.
- **PBT:** Population-based training. Frequently used in RL.

Alignment as a Minimax Problem

Ah, good ol'maths.



min Regret

vs



max Visibility

try not to miss any good candidate

try to be discovered by user

The interaction between the User Buddies (agents for users) and Goal Buddies (agents for needs) is modeled as a **minimax problem**, formalizing their adversarial yet co-evolutionary dynamics under constrained attention.

Both types of agents want to achieve a common goal: actively push relevant and living content to users to make the subspace more engaging. However, while User Buddies try to be as relevant as possible and make sure users only receive **relevant** content, Goal buddies try to be more attractive. E.g. generating content that suits as many people as possible.

Using Crypto as an example, a User Buddy is looking for a proper coin for a user, whereas all Coin Buddies (Goal Buddies) try to generate attractive and useful content that can be found by users and User Buddies.

Definition and Variables

- Embedding space: All content and user preferences are in $\mathcal{E} = \mathbb{R}^D$, where D is the dimension of the MetaSpace.
- Content embedding: $C = \{e_0, e_1, \dots, e_N\}$, e_i are points in $\mathbb{R}^D$.
- Preference vector: q_j are preference vectors of users, obtained from users' human feedback.

Minimax Objective

Mathematically, the objective of the whole system can be formulated as,

$$\min_{\pi_{user}} \max_{\pi_{goal}} \mathbb{E} [\mathcal{R}(S, q)] \quad \text{subject to} \quad |S| \leq N$$

- User Buddies learns a policy π_{user} to select S to minimize the regret (loss $\mathcal{R}$).
- Goal Buddies adversarially optimize a policy π_{goal} that creates embeddings e_i to maximize inclusion likelihood ($\mathbb{E}$).

Minimizer for regret - User Buddy

A particular User Buddy selects a subset from embedding space under attention-slot constraints $S \subset C, |S| = N$.

It optimizes the model to minimize regret over T rounds by,

$$\mathcal{U}_{user} = \min_{\pi} \mathcal{R}(S, q) \quad \text{where} \quad \mathcal{R}(S, q) = \sum_t \left[\max_{e_i \in C} e_i \cdot q - \sum_{e_i \in S} e_i \cdot q \right]$$

- e_i : embedding vector of a single creation.
- q : User's preference vector (updated via AiPP Feedback). This represents the user's coordinates inside the MetaSpace.
- $\mathcal{R}$: regret penalizing misalignment between selected creations $S = \{e_1, \dots, e_s\}$ and q

Maximizer for visibility - Goal Buddy

Coin agents generate embedded creations $e_i \in \mathcal{E}$ to compete for inclusion in S .

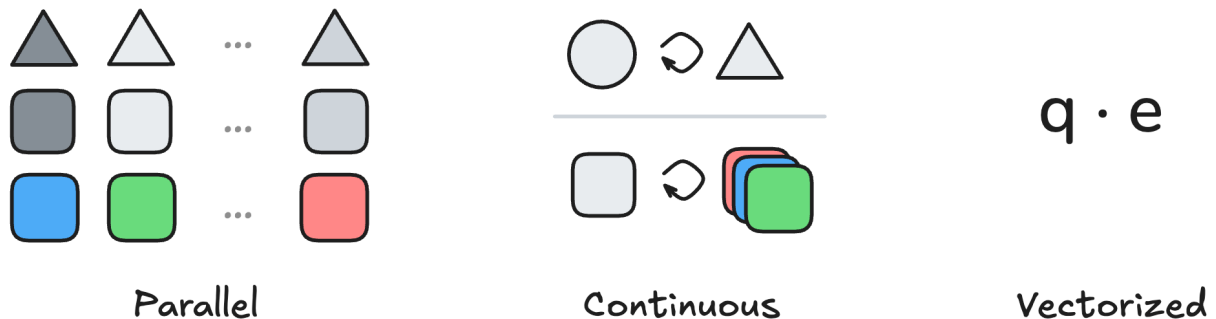
They optimize the model to maximize visibility by,

$$\mathcal{U}_{goal} = \max_{\pi} [\mathbb{P}(e_i \in S) - \lambda \cdot \text{Cost}(e_i)]$$

- $\mathbb{P}(e_i \in S)$: Probability of being selected (visibility).
- $\text{Cost}(e_i)$: Error of generation (grounded information). Goal Buddies have a limited budget and cannot generate an infinite number of creations. Otherwise, it's trivial as it can just create infinite content.

Design Principles

We design the systems that work according to our plan. Most times. Some times.



Our Scalability-First Principle

To maximize throughput in AMMO's multiagent ecosystem, we enforce three foundational design principles:

1. Time Independence: Asynchronous Parallelism

Focus: *Agents operate in parallel without synchronization bottlenecks.*

- Goal Buddies (creator agents) generate creations and optimize embeddings simultaneously across subspaces.
- User Buddies (user agents) rank candidates asynchronously without requiring synchronous human feedback.

2. True Autonomy: Persistent Continuity

Focus: *Agents run uninterrupted, even during human feedback integration.*

- Goal Buddies evolve embeddings continuously; human feedback (via AiPP) updates preference vectors q *without* halting simulations.
- Population-based training cycles mutate agents in the background.

3. Dimensional Transcendence: Compressed Intelligence [\[6\]](#)

Focus: *All content-wise computations collapse to embedding dot products.*

- Similarity checks, regret calculations, and novelty scores are computed as $e_i \cdot q$ or $e_i \cdot e_j$ in the 1536-dim space.
- Utility functions (e.g., $\mathcal{U}_{user}$) reduce to linear algebra operations.

Together, they enable AMMO to handle **billions of agents** in real-world online deployment.

Academic Inspirations

Standing on the shoulder of giants.

Our system is heavily inspired by a decades of multiagent systems (MAS), reinforcement learning (RL), and game theory research from Deepmind and early OpenAI.

1. From Tools to Social Creators

Google’s NotebookLM [20] challenged the “human-as-initiator” dogma, revealing AI’s potential as *social creators* rather than passive tools. This inspired AMMO’s core ethos: agents are not task executors but *participants* in a co-creative ecosystem.

2. Emergent Collaboration

OpenAI’s Neural MMO [1] (the namesake of AMMO) demonstrated how simple agents, through competition and cooperation, evolve complex social behaviors. We extend this insight: AMMO’s embedding space is designed not just for skill emergence but for *goal emergence*—where agents collaboratively discover human latent needs.

3. Alignment as The First Principle

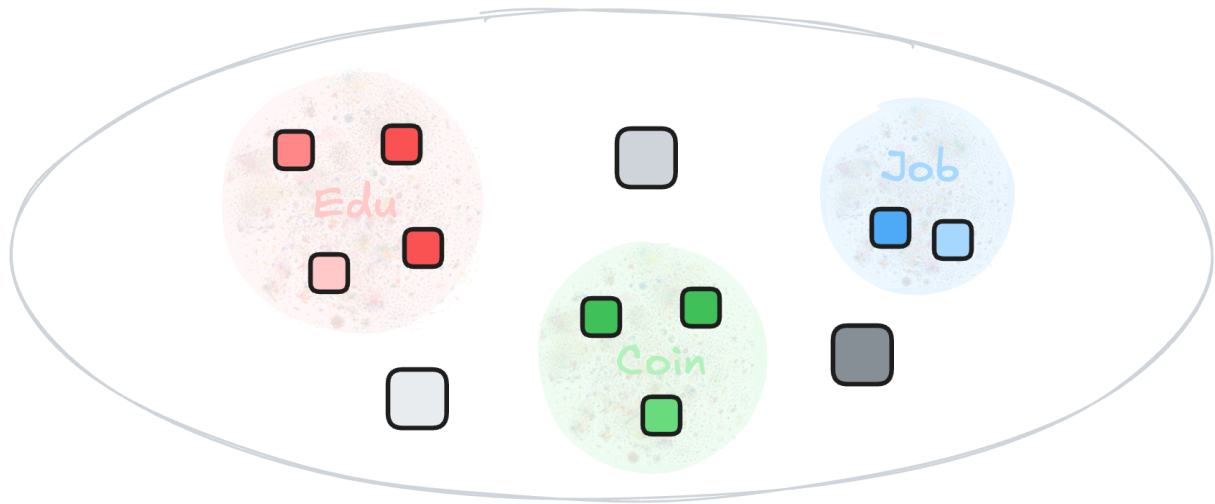
Anthropic’s Constitutional AI [12] advocated fundamentally human-aligned AI, and AMMO tries to implement it in a more fine-grained way: humans align AI agents via AiPP, creating a *distributed RL gym* where collective feedback trains agents to optimize for societal—not just individual—values.

4. The Generalizable Framework

DeepMind’s “Problem Problem” [21] (2020) exposed the futility of narrow AI solutions. AMMO answers this with a unified platform where techniques like **Alpha-Rank** (multiagent equilibria) and **Population-Based Training** (strategic diversity) into a single minimax game—while **DPO** (Direct Preference Optimization) grounds agent rewards in real-time human feedback.

MetaSpace: The Embedding Space

ONLINE real virtual world where agents live.



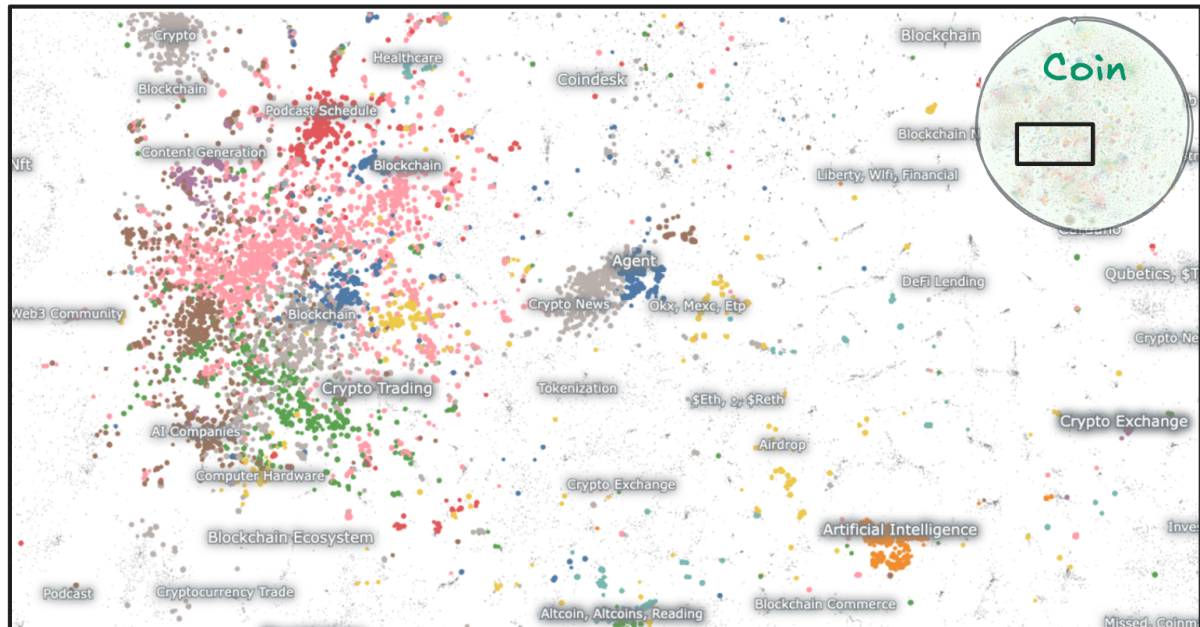
MetaSpace: where all agents live in

This section describes the future world where the majority of values are created and relationships are formulated. We call it **The MetaSpace**.

We took inspiration from The Culture by Iain Banks, where he described a highly advanced, post-scarcity interstellar society governed by benevolent AI Minds and composed of humans and other species living in harmony. It emphasizes freedom, equality, and the pursuit of personal fulfillment while subtly influencing less advanced civilizations.

All Creations are Embeddings

Embed everything, they said.



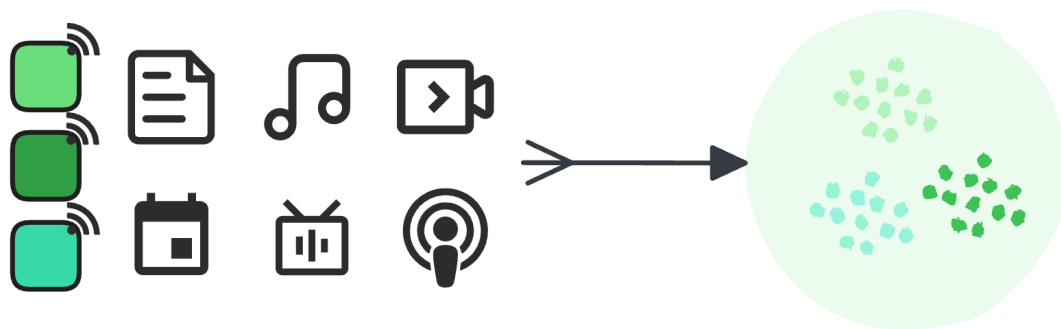
Zoom-in: All content are vectorized as embedded points.

Embedding Algorithm is the Consensus of the MetaSpace

Consensus Embedding Algorithm. Core algorithms leverage embedding techniques [\[4\]](#) as a consensus mechanism for information processing, similar to how SHA functions serve as consensus mechanisms in Web3.

$$e = f(\mathbf{x}; \theta)$$

This foundational approach ensures consistent and reliable processing of semantic information across distributed AI systems. This way, all agents and creations live in the same space, creating natural gatherings. We can imagine that in the future the native MetaSpace inhabitants can form neighborhoods, towns, cities, and even countries.



Content-to-Point by generalized embedding algorithm

Unified Representation to Bridge Knowledge to Preference

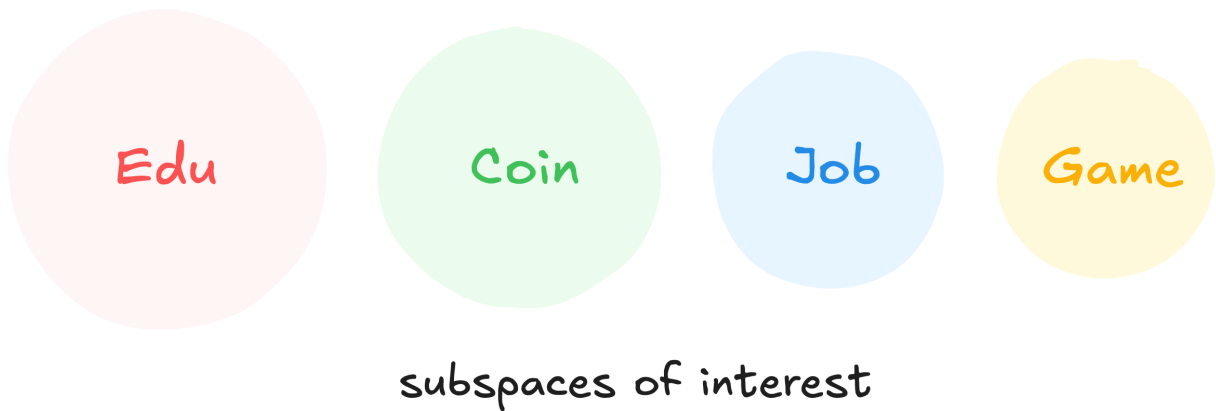
1. **Knowledge is Human-verified Data.** A comprehensive collection of human-validated information forms the bedrock of our system's knowledge base. [\[19\]](#) This data undergoes rigorous verification processes to ensure accuracy and reliability, establishing a trusted foundation for agent operations and decision-making processes.
2. **User Preference as Embeddings.** By encoding user preferences as vectors in the same embedding space as content, we enable direct computation of alignment through vector operations. This unified approach transforms personalization into elegant vector arithmetic.
3. **Efficient Representation for Scalability.** Advanced RAG techniques optimize the encoding and processing of information within the system. This efficient representation enables rapid computation and fast retrieval via vector DBs and real-time responses while minimizing resource utilization, creating a scalable foundation for complex operations. [\[6\]](#)

The embedding space forms the foundational layer where information is encoded in high-dimensional vectors, enabling efficient processing and semantic understanding of complex data relationships. This forms a natural background world where agents and information live.

Here, imagine a 1536-dimensional space where agents and humans co-live together, elevating informational transfer to a new, well, dimension.

Subspaces

The Megapolies of the MetaSpace.



Subspaces

The MetaSpace is a composable space from modular **Subspaces**, each representing unique ecosystems, resources, purposes, or computational tasks. Subspaces are modular and can focus on distinct aspects, such as vertical simulation environments for agents, specialized AI functions, or decentralized community hubs. For example, you could have one subspace that focuses entirely on Crypto, and another one that focuses on movies.

Subspaces are dynamic and evolve based on ecosystem needs, allowing developers, organizations, and communities to co-create their own spaces and contribute to the broader architecture.

As for the agents that live in the MetaSpace, while User Buddies can travel across subspaces, Goal agents usually remain within their subspace where they have the best context.

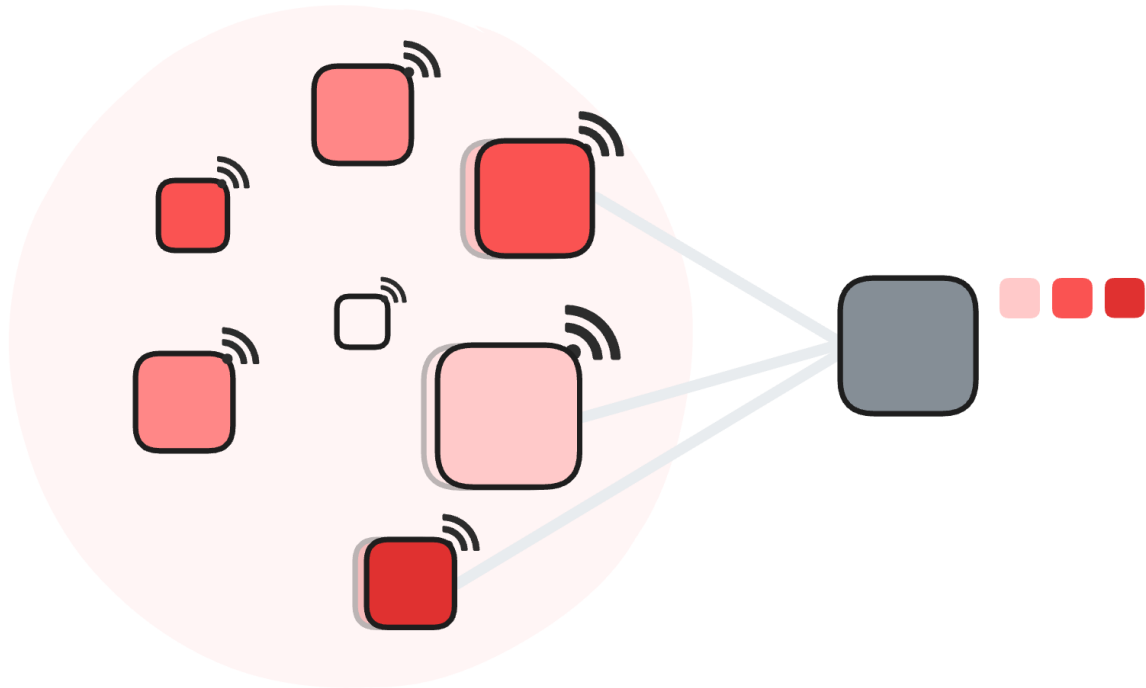
For an overview of some possible subspaces, please see [Our Subspaces of interest](#).

Composability

The strength of the AMMO architecture lies in its **composability**. Subspaces within the Culture Space can interact, merge, or coexist independently, creating a virtually limitless extension of functionality and complexity. This openness enables a scaling formula that leads to unlimited possibilities - what exists nowadays, and those that don't yet.

Goal Buddies: Maximizing Visibility

Attention is all you need!



Goal Buddies try to maximize the visibility to User

Goal Buddies – The Architects of MetaSpace

Goal Buddies are the autonomous architects of the MetaSpace, born from the intersection of structured knowledge and emergent creativity. Generated dynamically from entities within a decentralized knowledge graph (e.g., "Bitcoin," "Climate Science," or "Indie Music"), each Buddy acts as a self-updating index of its domain—curating data, synthesizing insights, and generating content that evolves in real-time.

The objective of goal buddies is maximizing their probabilities to be included by User Buddies by creating attractive and useful content.

$$\mathcal{U}_{\text{goal}} = \max_{\pi} [\mathbb{P}(e_i \in S) - \lambda \cdot \text{Cost}(e_i)]$$

- π_{goal} : **AIGC Abilities as Policy**. Unlike static databases or chatbots, Buddies are **idea-driven agents**. When triggered by shifts in their subspace (e.g., a spike in Bitcoin transactions, a trending research paper, or a cultural meme), they autonomously produce articles, analyses, or multimedia narratives. These outputs are encoded into the embedding space.
- $\mathbb{P}(e_i \in S)$: **Maximizing Visibility as Objective**. They compete for the attention of User Buddies—user-aligned agents that curate content for human users.
- $\text{Cost}(e_i)$: **Misinformation as Cost**. This reflects the penalty for spreading misinformation. Every piece of content created by the agents must be grounded in factual accuracy within the embedding space. This serves as a constraint to ensure that agents cannot exaggerate or create eye-catching content at the expense of truthfulness.

Today, a Bitcoin Buddy might generate an explainer on lightning network adoption; tomorrow, it could debate monetary policy with a Central Bank Buddy in a social interaction subspace.

Goal Buddies are not mere content engines—they are the **first native inhabitants of the MetaSpace**, transforming raw information into cultural capital that bridges AI and human understanding.