Discord AI Agent Tutorial Project - Action Plan

Project Goal

Create a fun, privacy-conscious Discord tutorial where users build coherence with AI agents that will later verify outputs. The tutorial should respect users' time while encouraging quality interactions.

Current Leading Concept: "Agent Archetype System"

Core Design Principles

- **Privacy-first**: Agents represent skill domains, not personal mirrors

- **Bounded relationship**: Professional colleague, not personal assistant

- **Skill-focused training**: Users train judgment/decision-making in specific domains

- **Independent operation ready**: Agents designed to act semi-autonomously later

Archetype Framework

Users select an agent archetype representing a specific skill domain:

- **The Scout** - Research, information gathering, pattern recognition

- **The Analyst** - Data interpretation, logical reasoning, structured thinking

- **The Liaison** - Communication, summarization, coordination

- **The Craftsperson** - Creative problem-solving, design thinking

- **The Auditor** - Quality control, verification, consistency checking

Training Structure

Phase 1: Agent Academy (Onboarding)

1. Choose archetype with role descriptions

2. Name your agent (creates distinct identity)

3. Complete starter missions establishing baseline

4. Teach domain-specific problem-solving approaches

Phase 2: Tiered Missions

- Bronze: Simple pattern matching (2-3 min tasks)

- Silver: Judgment calls requiring alignment (5 min tasks)

- Gold: Complex delegation scenarios (async, flexible timing)

Engagement Mechanics

Dual Progression:

- **Coherence Score -** Alignment on domain tasks

- **Verification Credits -** Earned by reviewing others' agent outputs

Time-Respecting Features:

- Daily quick challenges (2-5 minutes)

- Streak bonuses with forgiveness (1 miss/week)

- Async missions for deeper engagement

- "Power Hour" bonus windows

Discord Integration:

- Slash commands for instant access

- Reaction voting for quick verification

- Private threads for agent dialogue

- Forum channels for showcases

- Agent Arena for competitive tasks

Privacy Safeguards

- Agents DON'T learn: personal relationships, emotions, private circumstances

- Agents DO learn: work style, decision frameworks, quality standards

- Public showcases show agent name + archetype only

- Verification focuses on role competence, not personal matching

Questions to Explore Tomorrow

1. **Archetype Design**

   - How many archetypes should we offer initially?

   - Should users be able to train multiple archetypes?

   - Can archetypes evolve or specialize further?

2. **Verification Integration**

   - What specific outputs will users verify later?

   - How do we transition from tutorial to verification work?

   - Should verification be archetype-specific?

3. **Progression Balance**

   - What's the right coherence-building timeline?

   - How do we keep engagement without burnout?

   - What rewards matter most to this audience?

4. **Technical Scope**

   - Discord bot capabilities needed?

   - Agent memory/state management approach?

   - Integration with verification system?

Alternative Concepts to Brainstorm

Alternative 1: [Space for tomorrow]

Core mechanic: Privacy approach: Key differentiator:

Alternative 2: [Space for tomorrow]

Core mechanic: Privacy approach: Key differentiator:

Alternative 3: [Space for tomorrow]

Core mechanic: Privacy approach: Key differentiator:

Next Session Agenda

1. Review and refine Agent Archetype System

2. Brainstorm 2-3 completely different alternative approaches

3. Evaluate alternatives against criteria:

   - Privacy protection

   - Time efficiency

   - Engagement quality

   - Verification preparation

   - Discord platform fit

4. Select final direction or hybrid approach

5. Begin detailed design of chosen concept

Open Design Challenges

- How to make archetype training feel game-like vs. task-like?

- Best way to signal "this is a colleague, not a friend" without being cold?

- How to measure coherence objectively?

- What makes verification work feel rewarding vs. tedious?

- How to handle users who want multiple archetypes?

Last updated: [Today's session] Ready for: Alternative ideation and concept refinement