Syllabus Revision Document

Course

Course Title: CS 400: Architecting Autonomous AI Agents

Working Revision Goal: Adapt the existing course so that technical learning about agentic AI is taught through the 4-Stage Practice Model of Intergroup Dialogue and critical dialogic pedagogy.

Note: It is not taught yet, but the instructor hope to teach it soon in CEAS.

Revised Course Description

CS 400: Architecting Autonomous AI Agents introduces students to the design, implementation, evaluation, and governance of autonomous AI systems. Students learn core technical concepts such as ReAct-style reasoning loops, planning workflows, tool use, environment interaction, multi-agent orchestration, context engineering, and safety validation. At the same time, the course treats agentic AI as a social and ethical practice rather than a purely technical one. Students examine how design choices in agentic systems are shaped by power, access, labor, bias, institutional priorities, and assumptions about human oversight.

The course approaches learning through a combination of lecture, technical lab work, reflective writing, structured dialogue, collaborative case analysis, and project-based design. Using the 4-Stage Practice Model of Intergroup Dialogue, students begin by building relationships and shared agreements for inquiry, then explore differences and commonalities in experiences with AI systems, engage controversial questions about autonomy and responsibility, and finally develop collaborative action plans through the design of responsible agentic systems. Critical dialogic pedagogy tools such as positionality reflection, strategic questioning, storytelling and story-listening, dialogue protocols, and facilitated discussion are integrated into the technical curriculum.

By the end of the course, students will be able to build and test bounded agentic systems, explain and justify their design decisions, evaluate technical and social risks, and communicate how responsible AI development requires both engineering rigor and dialogic practice. The course prepares students to create agentic systems that are not only functional, but also transparent, accountable, and responsive to the communities they affect.

Course Learning Objectives

By the end of the course, students will be able to:

1. Explain the core architectures and workflows of agentic AI systems, including ReAct reasoning, tool use, planning, context management, and multi-agent orchestration.

2. Analyze how social identities, institutional power, and historical inequities shape the design, deployment, and impact of autonomous AI systems.
3. Apply agentic design patterns and development tools to build, test, and iteratively improve an autonomous AI workflow for a clearly defined task.
4. Evaluate tradeoffs among autonomy, transparency, usability, security, and human oversight by using technical evidence, case studies, and structured dialogue.
5. Facilitate and participate in critical dialogue about controversial AI topics by listening across difference, asking strategic questions, and responding constructively to disagreement.
6. Create a final project that proposes, prototypes, and justifies a responsible agentic AI system, including an implementation plan, risk analysis, and validation strategy.

Community Agreements

These agreements establish the learning conditions for technical rigor, dialogue across difference, and collaborative accountability. They should be introduced during the first week, revisited at each stage transition, and revised with student input as needed.

1. Speak from your own experience. Use "I" statements when sharing interpretations, concerns, and assumptions.
2. Listen to understand before responding. Practice story-listening, paraphrasing, and clarification before critique.
3. Critique ideas, systems, and design choices without demeaning people.
4. Ground claims in evidence. Bring together technical results, readings, examples, and lived experience where relevant.
5. Make room and take space. Students who speak often should create room for others; students who speak less should be supported in entering the conversation.
6. Expect discomfort and stay engaged. Discomfort is not the same as harm; we will examine why a topic feels difficult and what can be learned from that difficulty.
7. Name impact and invite repair. When something lands poorly, we pause, clarify impact, and work toward repair rather than defensiveness.
8. Respect confidentiality and context. Personal stories shared in class should not be circulated outside the class without permission.
9. Use tools responsibly. Do not expose secrets, private data, or unsafe outputs when demonstrating systems or sharing logs.
10. Take breaks seriously. Students may step back, pause, or request a reset when discussion or project work becomes overwhelming.

4-Stage Practice Model Integration

Stage	Course Focus	Dialogic Practices	Technical Focus	Sample Assessments
Stage I: Group Beginnings - Forming and	Build trust, establish classroom	Community agreements, positionality	Foundations of agentic reasoning,	Positionality reflection on AI, participation in

Stage	Course Focus	Dialogic Practices	Technical Focus	Sample Assessments
Building Relationships	norms, and introduce agentic AI as both a technical and social practice.	reflection, pair dialogues, active listening practice, low-stakes discussion protocols.	ReAct loops, basic tool use, introductory environment norms.	agreements process, introductory lab tracing a ReAct loop.
Stage II: Exploring Differences and Commonalities in Experiences	Examine how people experience AI systems differently across roles, identities, and institutions.	Storytelling and story-listening, strategic questioning, reflective writing, small-group dialogue, power/privilege analysis.	Planning workflows, context engineering, tool selection, user experience and access concerns.	Case analysis on differential impact of AI systems, reflective memo, tool-optimization lab.
Stage III: Dialoguing About Controversial "Hot" Topics	Engage disagreements about automation, surveillance, labor, bias, safety, and human authority in AI systems.	Fishbowl dialogue, structured controversy, facilitation moves for hot moments, evidence-based discussion, collaborative norm resets.	Safety, security, validation, model transparency, production mandates, risk mitigation.	Debate-to-dialogue exercise, risk assessment brief, safety review of an agent design.
Stage IV: Action Planning and Collaboration	Move from critique to collaborative design and responsible action.	Consensus building, collaborative design review, peer feedback, action planning, reflective synthesis.	Multi-agent systems, orchestration, capstone prototyping, verification, responsible deployment planning.	Final capstone project, implementation plan, peer review, validation report, public-facing design rationale.

Critical Dialogic Pedagogy Tools and Practices

The following tools and practices should be embedded across the revised syllabus:

- **Structured dialogue protocols:** Use opening rounds, paired listening, and timed reflection before whole-class discussion.
- **Strategic questioning:** Invite students to move from factual questions to analytical, reflexive, and transformative questions about agentic AI.

- **Storytelling and story-listening:** Ask students to connect technical systems to lived experiences with automation, decision support, surveillance, accessibility, and labor.
- **Positionality reflection:** Have students examine how their own disciplinary background, social location, and prior experiences shape their assumptions about AI.
- **Dialogue facilitation moves:** Normalize paraphrasing, summarizing tension, inviting quieter voices, and pausing for repair during conflict.
- **Critical reflection writing:** Use short reflections before and after difficult discussions to help students process learning and discomfort.
- **Collaborative knowledge building:** Treat peer dialogue as a site of inquiry where technical understanding is deepened through multiple perspectives.
- **Action orientation:** End units with design or policy responses that connect course learning to more just and accountable AI practices.

Revised Assessment Alignment

Assessment	Purpose	CLO Alignment
Reflection Journals	Support positionality work, dialogue preparation, and post-discussion synthesis.	2, 4, 5
Technical Labs	Build practical skills in reasoning loops, tool use, planning, and validation.	1, 3
Case Analysis Memos	Analyze the social and institutional implications of agentic AI systems.	2, 4
Facilitated Dialogue Participation	Practice listening, questioning, and constructive engagement across disagreement.	4, 5
Risk and Safety Review	Evaluate an agentic system for transparency, security, oversight, and failure modes.	3, 4, 6
Final Capstone Project	Design and justify a responsible agentic AI system with implementation and validation plans.	1, 3, 4, 5, 6

Sample Weekly Arc by Stage

Stage I: Weeks 1-3

- Introduce course purpose, agreements, and dialogic expectations.
- Teach foundational concepts: chatbot versus agent, ReAct loops, thought-action-observation cycles.
- Assign a positionality reflection: “What experiences shape how I interpret AI autonomy, intelligence, and trust?”

Stage II: Weeks 4-6

- Explore differential experiences with AI through case studies involving education, advising, hiring, accessibility, and public services.

- Connect planning, context engineering, and tool choice to user needs, institutional constraints, and inequitable outcomes.
- Use storytelling, listening pairs, and reflective memos to surface assumptions and differences in perspective.

Stage III: Weeks 7-10

- Organize dialogic sessions around contested topics such as explainability, surveillance, replacement of human labor, and acceptable failure risk.
- Pair controversial dialogue with technical work on safety, evaluation, and validation-driven development.
- Use facilitation protocols for hot moments, including pause, paraphrase, clarify impact, and re-enter.

Stage IV: Weeks 11-14

- Shift toward collaborative action through multi-agent architecture, capstone planning, peer review, and validation.
- Ask students to justify not only what their system does, but whom it serves, what risks it creates, and what accountability mechanisms it includes.
- Conclude with a final reflection on how dialogue changed each student's engineering practice.

Suggested Syllabus Language for Teaching Philosophy

This course treats learning as both technical and relational. Building agentic AI systems requires more than code proficiency; it requires the ability to reason with others, examine assumptions, respond to conflict, and design with accountability. For that reason, the class uses critical dialogic pedagogy alongside technical instruction. Students will be asked to test ideas, reflect on their own positions, listen seriously to others, and connect system design to broader social consequences. The goal is not consensus on every issue, but deeper understanding, stronger analysis, and more responsible action.

Suggested Closing Statement for the Revised Syllabus

Because agentic AI systems increasingly shape decision-making, labor, communication, and access, this course asks students to become both capable builders and accountable participants in public life. The revised syllabus frames technical mastery and dialogue across difference as mutually necessary practices. Students leave the course with stronger engineering skills, stronger habits of reflection, and a clearer sense of responsibility for how autonomous systems affect people and institutions.