What and How Should Urban Planners Learn in the AI Era? Exploring Urban AI Pedagogy from a Pilot Course in Urban Planning Education

Xiaofan Liang, University of Michigan, Ann Arbor, Ann Arbor, United States
xfliang@umich.edu

BIO:

Xiaofan Liang is an Assistant Professor of Urban and Regional Planning at University of Michigan Ann Arbor. Her research interests include Urban Networks and Urban AI.

Keywords:

Urban AI, Urban Informatics, Planning Pedagogy

Citation:

Liang, X. (2026). What and How Should Urban Planners Learn in the AI Era? Exploring Urban AI Pedagogy from a Pilot Course in Urban Planning Education. *Journal of Planning Education and Research*, *0*(0). https://doi.org/10.1177/0739456X261436057

ABSTRACT

Artificial Intelligence (AI) promises to transform urban planning research, practice, and education, yet few curricula address “Urban AI”. This paper presents the pedagogical design of a pilot Urban AI course and argues for three meta learning goals: applying AI effectively and appropriately in urban challenges, addressing its social, environmental, and governance impacts, and developing normative judgements and professional identities around AI. Pilot teaching produced a knowledge graph connecting essential skills to these goals and a critical framework for AI use and reflection, grounded in analysis of 235 student reflection journals, alongside course evaluations, syllabus materials, and student projects.

INTRODUCTION

Artificial Intelligence (AI) is increasingly used in urban planning research to review planning documents (Brinkley and Stahmer 2024; Fu, Li, and Zhai 2023), analyze zoning codes (Salazar-Miranda and Talen 2025), monitor traffic (Boukerche and Hou 2021), assess urban perception (Ito et al. 2024), classify land uses (Chaturvedi and Vries 2021), simulate urban patterns (Wu, Stouffs, and Biljecki 2022), and inform urban design (Quan 2022). As these applications expand, scholars have begun to examine AI’s conceptual foundations, including the flow of domain knowledge between AI and urban science (Ye et al. 2025; Zheng et al. 2025) and its ethical and societal implications (Sanchez, Brenman, and Ye 2025). Yet the field still lacks enough planners conversant in AI. A recent American Planning Association survey shows cautious adoption and mixed sentiments among practitioners despite broad acknowledgment of AI’s importance (Sanchez et al. 2023).

Few urban planning curricula have incorporated AI content, enabled students to meaningfully leverage GenAI tools in their learning, or even designed a course explicitly addressing the emerging field of Urban AI. This gap leaves educators facing critical uncertainties: What knowledge and skills define Urban AI? How should they be taught? And how can student learning be effectively evaluated?

This paper addresses these questions by presenting the pedagogical design and assessment of an Urban AI course offered at University of Michigan. The paper introduces a knowledge graph that characterizes the components and flows of knowledge essential for urban planning and urban technology students in the field of Urban AI, alongside a set of evaluation metrics to assess students' critical use and thinking with AI tools. These frameworks were refined through feedback from the pilot offering of the course and grounded in hundreds of examples drawn from student reflection journals.

The paper contributes to planning pedagogy in three ways: (1) by proposing an operational teaching framework for faculty seeking to introduce Urban AI into their classrooms; (2) by highlighting key pedagogical tensions and student growth observed through empirical evidence; and (3) by sharing a publicly available syllabus (LINK REDACTED), student final project portfolio (LINK REDACTED), and lessons learned from the pilot course as starting points for debate, refinement, and co-construction of Urban AI pedagogy. Together, these contributions illustrate how AI can be embedded in domain-specific pedagogy and provide adaptable assessment metrics for strengthening students' critical AI use and reflective practice, which holds values beyond urban planning.

LITERATURE REVIEW

Technology in Planning Pedagogy

Planning has long incorporated new technologies, from early computer use (French and Wiggins 1989) to GIS (Dueker 2007) and, more recently, Urban Science as a computational paradigm for analyzing cities (Batty 2013). Today, most planning programs require quantitative methods and GIS courses to satisfy analytical skills requirements set by the Planning Accreditation Board (PAB 2025). Programs that have embraced data-driven curricula (e.g., Urban Informatics, Urban Analytics, Spatial Data Science, etc.) tend to converge on two emphases: (1) teaching analytical and computational skills, and (2) addressing domain-specific applications and operational challenges (Kontokosta 2021; Clayton et al. 2024). Following this tradition, Artificial Intelligence (AI) represents the next major technological shift for planning education.

Prior efforts to integrate GIS and Urban Informatics highlight persistent barriers, including limited faculty expertise, uneven institutional capacity, and the rapid evolution of technologies outside planning (Godschalk and McMahon 1992; Drummond and French 2008; Kontokosta 2021). Unlike GIS, where Esri provides extensive teaching resources, Urban Informatics and AI rely on decentralized developer resources that require instructors to curate and contextualize materials. Another long-standing challenge is balancing conceptual, normative, and operational learning goals. Because GIS, data science, and AI span multiple disciplines, instructors often struggle to align technical tools with planning's pedagogical priorities (Montagu 2001; Clayton et al. 2024). Similar patterns appear in data science education, where ethics remains largely discussion-based while labs focus on technical execution (Lewis and Stoyanovich 2022). Suggested solutions emphasize more intentional integration of off-the-shelf technologies with planning needs, and greater reliance on problem- and practice-oriented teaching so that technologies become meaningful planning instruments rather than isolated technical skills (Montagu 2001; Kontokosta 2021; Lewis and Stoyanovich 2022).

Cross-disciplinary Integration of AI in Pedagogy

Across disciplines, educators have begun reflecting on how AI should be incorporated into teaching and learning. In data science, for example, scholars argue that the rise of large language models (LLMs) is reshaping the profession itself, shifting expectations from hands-on coding to LLM-guided reasoning, creativity, and critical evaluation (Tu et al. 2023). Geography has likewise embraced AI through the emerging subfield of GeoAI, demonstrating that AI can both automate traditional GIS tasks and enable entirely new forms of geospatial analysis (Huang et al. 2025). Evidence on teaching with AI remains mixed: while some studies report improvements in short-term learning and higher-order thinking (Deng et al. 2025), others caution that relying on AI to substitute rather than scaffold thinking may undermine learning (Lehmann, Cornelius, and Sting 2024).

Despite disciplinary differences, these debates converge on two core pedagogical questions: What forms of disciplinary knowledge matter in an AI-mediated environment, and how should learners strategically use AI? The AI-TPACK framework (Lee 2025), which is an extension of the original Technology, Pedagogical, and Content Knowledge framework (TPACK) that emphasizes effective integration of AI tools in teaching, provides one perspective to these questions. AI-TPACK outlines four core domains of knowledge for educators: 1) AI-related knowledge, knowing how AI works and how to use it, 2) AI-related pedagogy knowledge, knowing how to teach with AI, 3) AI-related content knowledge, knowing how AI is used in a field, and 4) AI-related pedagogical content knowledge, knowing how to teach a subject with AI. Similarly, learning goals for students also focus on metacognitive skills, such as how to navigate human-AI collaborative learning environment (Lee 2025) and making critical and normative judgments about when and why to use AI

(Lewis and Stoyanovich 2022). Thus, positioning AI within disciplinary knowledge and cultivating metacognitive skills are universal pedagogical challenges in the AI era.

Yet instructors frequently encounter inappropriate or shallow uses of AI and lack concrete pathways to guide improvement beyond prohibition or punitive measures. By proposing a structured knowledge-integration roadmap and a criticality-assessment framework, this paper contributes not only to urban planning pedagogy but also to broader conversations across fields seeking actionable strategies for responsible and meaningful AI use in education.

What Makes "Urban" and "AI" Unique?

We define Urban AI as a field of study and practice concerned with the application and governance of artificial intelligence in urban contexts. It encompasses both (1) the adaptation of AI methods to multimodal urban data, socio-technical complexity, and planning decision contexts, and (2) the social, economic, environmental, infrastructural, and everyday-life implications of AI for how urban systems operate, are governed, and are experienced by residents.

Different from Geography and Data Science, urban planning engages a broader range of multimodal data (texts, graphics, transactions) and domain-specific challenges (e.g., zoning code analysis, generative design, digital twins). Standard AI models are often ill-equipped to handle urban planning challenges because cities are complex, dynamic socio-technical systems and thus require domain-specific adaptations (Ye et al. 2025). For instance, using machine learning to automate zoning code extraction has failed to represent zoning's language heterogeneity, interdependent logic, and interpretation vagueness (Axelrod et al. 2023). Urban Planning also extends further into the governance, socio-politics, and the built environment consequences of AI in urban life. A prominent example is permitting the land, energy, and water uses of data centers, which has become a pressing operational task in urban planners' day jobs in recent years (Morley 2022, Nichols 2025) but typically neglected in AI education.

Unlike prior domain-specific technologies such as GIS, AI's impacts on society are both wide and deep, reshaping cognitive tasks and professional values across disciplines. Parallel to data-literacy discourse that emphasizes subjectification—Biesta's (2015) call for "the coming into presence of unique individual beings", AI literacy for planners must position AI relative to individual experience and interests so that planners can learn to work with AI, learn from AI, and make informed judgments about how professional identities should evolve, including strategies of resistance.

AI is also different from traditional statistical methods and mathematical modeling. AI learns patterns directly from data, enabling large-scale prediction, classification, and generation rather than executing predefined rules or testing researcher-specified hypotheses. Still, statistics remains foundational: AI methods rely on statistical evaluation, and simple models such as linear regression illustrate how algorithms learn from data.

The distinction between Urban AI and data-driven subfields such as Urban Informatics and Urban Science is less clear. Some argue that advanced AI simply strengthens the emerging "New Urban Science" by enabling more precise, dynamic understandings of urban systems (Ye et al., 2025), rather than constituting a separate subfield. Yet existing Urban Informatics courses emphasize wrangling diverse urban big data, APIs, data ethics, and governance, often with only brief attention to machine learning. These data competencies are widely used in practice and serve as a foundation for deeper exploration of Urban AI. Thus, they should be considered prerequisites rather than substitutes for Urban AI.

PEDAGOGICAL TENSION IN URBAN AI

Teaching Urban AI surfaces several key pedagogical tensions, many of which echo long-standing debates in teaching GIS and Urban Informatics (Montagu 2001; Drummond and French 2008; Kontokosta, 2021).

Breath versus Depth

How much should Urban AI courses expose students to the wide-ranging technical and societal dimensions of AI versus focus intensively on a narrower set of questions central to urban planning (e.g., policy and governance)? A breadth-oriented course might survey a range of topics without details, covering contents like AI mechanics and AI applications in transportation, housing, or civic process, while a depth-oriented course might focus on a few topics central to planning values (e.g., ethics and governance) and tasks (e.g., using AI to analyze public feedback) and rely on case studies to unpack methodological and governance complexity. Yet both approaches face limitations: a broad survey may leave students aware of AI's risks but without the understanding and skills needed to imagine solutions, while a narrowly focused course depends on detailed case studies that are not always publicly available.

Literacy versus Capacity Building

Should Urban AI education be broadly accessible, equipping all students with AI literacy, or selective and advanced, cultivating a smaller group with the capacity to design, build, or govern AI systems? A literacy approach emphasizes conceptual breadth, such as how AI supports planning, reshapes urban scenarios, or raises governance issues, without demanding advanced technical skills. A capacity-building approach develops technical expertise, preparing students to engage directly in system design and oversight. The former accommodates students from diverse technical backgrounds, while the latter requires sequenced courses, technical prerequisites, and a motivated audience seeking to innovate with AI in their careers.

Technique-Centered versus Context-Centered

Should Urban AI pedagogy be organized around *techniques* (methods, transferable mechanics, transversal literacy) or around *contexts* (urban domains, real problems, situated judgment)? For instance, a technique-centered module on natural language processing might introduce classic methods such as named entity recognition, sentiment analysis, and topic modeling. A context-centered module would begin with a planning problem, such as zoning code analysis, and examine how different methods succeed or fail in that setting. A challenge with the context-centered approach is that real-world applications of AI in planning remain limited; thus, educators may need to rely heavily on examples from their own research rather than standardized teaching materials.

Big Tech Conformity versus Alternative Practices

Do we train students in the dominant tools that structure current practice, or do we emphasize alternatives and critiques that align better with planning's values of equity, democracy, and public good? Popular tools like ChatGPT and OpenAI models are consistent, well-documented, and widely used, making them effective for pedagogy. Yet focusing solely on these tools risks reinforcing algorithmic injustice and the hegemony of Big Tech in shaping urban life. As Freire (2000) reminds us in *Pedagogy of the Oppressed*, education carries the obligation and possibility to support critical consciousness and dismantle structures of injustice. Urban AI pedagogy, framed in this way, must consider not only what tools we teach *but also* whose values and futures those tools sustain.

Learning About AI versus Learning with AI

Should Urban AI pedagogy focus on AI as an object of study or AI as a medium of learning? A learning *about* approach emphasizes AI as a topic, aiming to build conceptual and technical understanding of how AI works and how it affects cities, while a learning *with* approach treats AI tools as learning companions, with pedagogy centered on cultivating interaction skills and critical reflection. These approaches imply different measures of success: the first evaluates students on their knowledge of AI systems, which is easier to operate, while the second assesses their ability to work with AI tools productively and critically in planning contexts, which has few existing guidelines.

PILOT COURSE CONTEXT AND PEDAGOGICAL APPROACH

The *Urban AI* pilot course was first offered at the Taubman College of Architecture and Urban Planning, University of Michigan – Ann Arbor, between January and May 2025. Students were expected to have prior knowledge of statistics and proficiency in a programming language, which can be learned through other courses.

Twenty students enrolled in the pilot course: roughly half from the Master of Urban and Regional Planning program, half from the Bachelor of Science in Urban Technology program, and one student from the Master of Public Health program. In the pre-course survey, all students reported prior interaction with AI tools such as ChatGPT and some awareness of AI's applications and implications for urban planning. However, few had any understanding of AI mechanics or knew how to address misalignments between AI tools, domain needs, and human values.

To address the pedagogical tensions discussed above, the course adopted an ambitious design that balanced breadth and depth, exposing students to a variety of perspectives while providing enough technical training to build capacity among a technically prepared audience (see Table 1). Most weeks consisted of a lecture paired with a hands-on Python lab. Lectures introduced AI concepts, methods, and theoretical perspectives, while labs focused on applying those methods to planning-relevant problem contexts. Beyond OpenAI's models (including ChatGPT), the course introduced students to the Hugging Face open-source model ecosystem and taught them how to evaluate and use models from this platform, encouraging exposure to alternative AI practices.

Students were encouraged to use AI tools freely throughout the course but were required to submit weekly reflection journals (12 out of 16 weeks) on what AI tools they use, how they use the tools, and their reflections on the use. Because lectures and labs emphasized technical skills, assignments encouraged domain-specific applications: students created code-based tutorials to teach their peers on selected topics and completed a final project on a topic of their choice. Additional details are available in the open-access course syllabus [link redacted for peer review].

RESEARCH METHODS

This study adopts an experiential learning and reflective practice approach (Kolb 1984; Schön, 1983) to refine a pedagogical framework for teaching Urban AI, addressing what to teach, how to teach, and how to evaluate success.

The research design itself follows an experiential cycle (Kolb 1984). During the course design phase, a three-session Urban AI pedagogy workshop was hosted with three academics, a practitioner, and a local government staff to provide the initial concrete experience, gathering diverse perspectives on essential knowledge and skills for planners in the era of AI. These insights from the workshop were distilled into a 16-week course schedule through deliberation about trade-offs and what best constitutes the body of knowledge for Urban AI, representing abstract conceptualization. These insights served as the foundation for developing a knowledge graph of Urban AI, mapping interconnections between knowledge components and their directions of growth. Course delivery and continuous adjustments during the semester served as active experimentation, allowing the framework to evolve in real time.

Schön's reflective practice was embedded through systematic analysis of students' weekly reflection journals (Schön 1983). These journals were iteratively coded and analyzed using ATLAS.ti, a qualitative analysis software. The software allows researchers to highlight text segments ("quotes") and assign one or multiple user-defined codes, facilitating the development of emerging themes and enabling easy review, merging, or splitting of codes to ensure internal consistency. All except for one student granted permission for their journals to be analyzed for research purposes.

Each reflection journal was coded for the following elements (not all discussed in depth in this paper): 1) AI tool(s) used, 2) Application behaviors and tasks, 3) Application critical level, 4) Whether the application was for this course or external tasks, 5) Reflection topics, and 6) Reflection critical level. Not every journal contained all six elements; some included only application descriptions or reflections.

Coding for reflection journals was conducted in three rounds. The first round identified the unit of analysis (text chunks describing distinct behaviors or reflections) and captured emerging themes. The second round refined and defined codes, ensuring internal consistency of quotes assigned to each category. The third round merged or split codes as necessary, aligning them with finalized application task categories, reflection topics, and five levels of critical engagement for both application and reflection. In the second and third round, ChatGPT-4o was used as a human-in-the-loop assistant to clarify borderline cases, suggest category names that captured the essence of each group of quotes, and generate preliminary code definitions. While using AI for qualitative coding carries the risk of embedding systemic biases and of relying on the very technology this study critiques, AI also provides consistent coding at scale which speed up qualitative analysis. To mitigate the risks, the author only treated AI strictly as a thought partner rather than a decision maker (see Appendix A for detailed methods and discussion).

A total of 655 quotes from 235 reflection journals were coded. Among them, 328 quotes were categorized into 22 AI Tool Application categories (roughly half and half for tasks conducted inside or outside the course), and 327 quotes were categorized into 18 AI

Reflection Topic categories. Both *Application Critical Level* and *Reflection Critical Level* were classified into five levels.

LEARNING GOALS, ESSENTIAL KNOWLEDGE, AND SKILLS FOR URBAN AI

Defining the essential knowledge and skills for Urban AI begins with clarifying the learning goals for urban planners and technologists. For the pilot Urban AI course, three learning goals are: (1) to apply AI effectively and appropriately in urban contexts, (2) to address AI's social, environmental, and governance implications in cities, and (3) to develop normative judgements and professional identities in relation to AI. The first two are knowledge-driven, while the third is practice-driven, helping students become critical AI users and thinkers, moving from passive users to active innovators, and from describing isolated interactions toward developing broader perspectives and normative judgments about AI's role in their work, values, and society. These learning goals are grounded in the unique positioning of Urban AI outlined in the literature review. Applying AI effectively and appropriately requires planning domain knowledge and professional discretion, while addressing AI's implications is inseparable from core planning practice. Identity development becomes particularly important in the AI era, where metacognitive skills must be prioritized over cumulative technical skills, especially as the latter quickly become outdated.

Table 1 summarizes five modules of essential knowledge and skills that was condensed in the Urban AI pilot course, which are 1) AI Mechanics, 2) Urban Application, 3) Governance and Society, 4) AI's Systematic Impacts, and 5) AI Proficiency and Reflection. These five modules were designed not merely as content areas but as a scaffolded pathway for developing Urban AI competence.

Modules 1–4 progressively move students from understanding *how AI works* (mechanics), to *how it applies to planning* (applications), to *how it should be governed* (society and ethics), and finally to *how it might transform urban systems* (systemic impacts). Module 5 functions as a cross-cutting practice layer: students repeatedly interact with AI tools, reflect on successes and failures, and progressively gain agency and critical thinking.

Table 1: Five Modules of Essential Knowledge and Skills for Urban AI

Modules	Knowledge Outcomes	Skills Outcome	Sample Topics / Activities

#1 AI Mechanics	Explain what Machine Learning (ML)/AI is; Distinguish model types; Understand how models learn from structured and unstructured data; Describe how large language models (LLMs) generate original contents	Train and evaluate a basic ML model; Apply Natural Language Processing (NLP) and Computer Vision (CV) pre-trained models from Hugging Face; Interact with LLMs via API.	Machine Learning and GenAI Sandbox (e.g., Teachable Machine; ChatGPT; Adobe Firefly); Common ML models (e.g., supervised versus unsupervised); Machine learning process and evaluation metrics; Model tradeoffs (e.g., bias versus variance)
#2 Urban Application	Understand how AI can support planning tasks and how planning might change in response.	Customize and prompt-tune LLMs; Apply AI to domain-specific problems; Develop transferable workflows.	Case Studies; Guest Lectures; peer tutorials (e.g., GeoAI; Video Detection; NLP for Public Comments; Chatbot Auditing); Final Project.
#3 Governance and Society	Identify social risks, costs, and mitigation strategies; Articulate the roles of human oversight and value alignment.	Gain exposure to major AI ethics frameworks, governance approaches, and human-in-the-loop concepts; Build vocabulary for discussing value alignment.	Operational definition of human values as in EU AI Act; Various AI ethical concerns and social costs (e.g., inequality, racial discrimination, loss of local representation, privacy, security, manipulation, etc.); AI Policy and Governance; Interpretable ML, Human-in-the-loop ML, Participatory AI; Guest Lectures
#4 AI's Systematic Impacts	Anticipate AI's systematic impacts on the environment, labor, economy, and human dynamics	Recognize key domains where AI affects urban systems (e.g., environment, labor, economy); Gain familiarity with current debates and planning considerations.	Environmental Impact of AI; Data Center; Geography of GenAI labor; Guest Lectures
#5 AI Proficiency and Reflection (Cross-cutting)	Understand the underlying concepts (e.g., tokens, context windows) and mental models (e.g., open source versus proprietary models); Understand Five-level of critical AI use and AI thinking as proposed by this study	Interact with AI tools to understand AI behaviors; Evaluate AI outputs for quality, relevance, and bias; Customize AI models through web interface or code; Decide when to accept, refine, or discard AI results; Reflect on how professional identities shape AI use.	Lab Exercises (many are identified above); Weekly Reflection Journals; Peer Tutorials, Final Projects

HOW TO DELIVER ESSENTIAL KNOWLEDGE AND SKILLS

Having defined the knowledge and skill pillars for Urban AI, the next step is to consider how to deliver them coherently and pedagogically effectively. Figure 1 presents a knowledge graph that illustrates how the course contents, originated from domain knowledge of both Computer Science and Urban Planning, converges toward the two of the three meta-level learning goals: *Apply AI Effectively and Appropriately in Urban Contexts*, and *Address AI's Social, Environmental, and Governance Implications in Cities*.

Each block in the diagram represents a cognitive milestone, while the connecting lines show how one milestone supports the next. For example, understanding AI mechanics helps students grasp the technical capacities and limitations of AI models, including their strengths and weaknesses (pink arrows). On the other side, urban planning and urban technology domain knowledge highlights key contextual challenges that influence AI's effectiveness, such as the complexity of zoning language or the sensitivity of certain urban systems to algorithmic decisions (cyan arrows). These two lines of knowledge converge to help students build intuition about where and how AI can add value in urban contexts, ultimately enabling more effective use of AI in practice.

Figure 1 also acknowledges the diversity of students' technical backgrounds and career aspirations by shading the center of the graph as a reflective space for the third learning goal: *Develop normative judgements and professional identities in relation to AI*. In the pilot course, this learning is deliberately practiced through weekly reflection journals and a required final-project reflection on how students' planning expertise led them to ask different questions and notice implications that engineers might overlook. Cultivating this reflective space is conductive for planners to find multiple pathways to work with digital technologies and for technologists to engage with urban planning (Clayton et al. 2024).

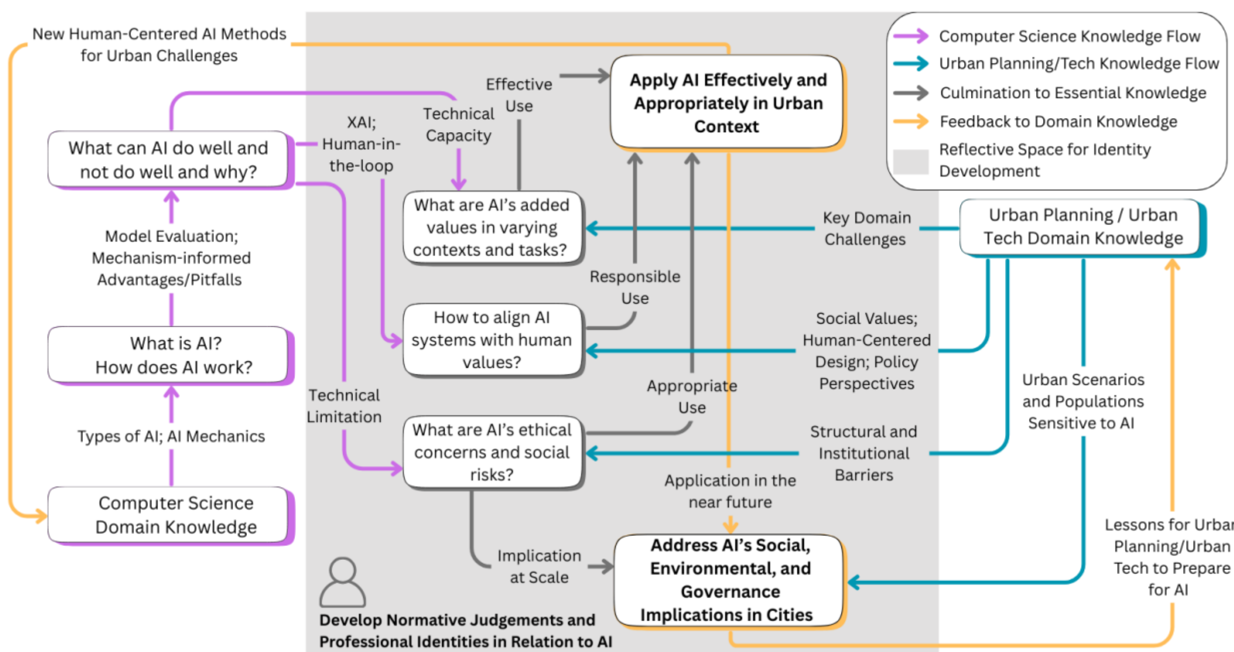


Figure 1: Knowledge Graph of Urban AI. Left side draws knowledge flows from computer science domain and the right side draws knowledge flows from urban planning and urban technology domain.

HOW TO ASSESS CRITICAL AI APPLICATION AND REFLECTION

The critical levels for AI application and reflection correspond to two dimensions of the identity development learning goal (Table 2). The first is human agency with AI, which goes from *consume and observe* AI products (Level 1), *iterate and explore* AI tools/models (Level 2), *analyze and control* AI use (Level 3), *strategize and contextualize* AI application (Level 4), *design and build* new AI application (Level 5).

The second is critical thinking about AI, which moves from simply *describe* an interaction experience (Level 1), *reflect* on personal takeaway from a single instance of use (Level 2), *analyze* why AI performance vary in context (Level 3), *evaluate and ethicize* AI use (Level 4), and *synthesize and innovate* perspectives of AI and its role in personal career, disciplines, and society at large (Level 5). Table 2 provides a definition and an example for each level.

Table 2: AI tool / model application and AI reflection defined by critical levels.

	AI Tool / Model Application	**AI Reflection**
Level 1	*Consume and Observe:* Consume AI as designed or observe AI outputs passively with little analysis. Complete exercises as instructed in class. Report minimal details in application. *Example:* "I used AI to summarize readings into key learning points."	*Describe:* Report what the tool did or what the user observed and felt, such as features tried, outputs produced, and surface feelings without explaining why it worked/failed or what that means. *Example:* "AI summarized the document in 2s, and the outline looks great!"
Level 2	*Iterate and Explore:* Adapt, refine, or repurpose AI outputs using iteration, targeted exploration, or basic feature adjustments. Work often builds on prior results, showing some agency beyond default use yet without full technical control *Example:* "I used AI to generate study guides for readings and Q&A packet to help with understanding."	*Reflect:* Reflect on personal takeaways about usefulness or friction, such as feelings, quick pros/cons, small tweaks to prompts, without digging into mechanisms. *Example:* "I like how the AI-generated study guides help me see the big picture and focus my attention in doing readings."
Level 3	*Analyze and Control:* Uses AI with clear goals and hypotheses, shaping outputs through mechanism-aware choices, edge-case testing, and validation against human reasoning or benchmarks. Prioritizes quality, accuracy, and learning over automation. *Example:* "I used AI to summarize these documents with six focus areas and cross-checked the outcomes with my own understanding."	*Analyze:* Explain why performance varies. Name causal factors (e.g., data, chunking, prompts, context, etc.) and compare settings to link AI behaviors to underlying mechanisms. *Example:* "The final summary missed important details. When I went back and check, these details were presented in a table. I suspect that AI cannot process the tabular contents well because AI model processes texts line by line."
Level 4	*Strategize and Contextualize* Integrate AI into broader strategies and contextual goals by adapting methods, testing systematically, and aligning outputs with domain-specific needs or personal objectives.	*Evaluate & Ethicize* Make judgments about suitability and risk, weigh trade-offs, compare against professional and contextual workflow, recommend when to use/avoid—sometimes with ethical or societal

	Example: "To make it easy for human validation, I tested a few strategies to summarize a planning document with AI, such as returning the page number for the original sources, showing reasoning for the summary, and return details around keywords that can be easily checked through a document keyword search."	stance, and considering broader implications for society. *Example:* "There is a risk if I had just used the summary AI gave me, as state-published materials must be approved and verified. This raised concerns for me about how bureaucracies might use AI in the future—and how their verification processes would need to adapt."
Level 5	*Design and Build* Designs, implements, or meaningfully customizes AI models, pipelines, or applications to address domain-specific challenges. *Example:* "When building the civic pothole reporter, I used AI to summarize and structure user submitted reports into variables that can support downstream categorization and task assignment."	*Synthesis & Innovate* Turn lessons into new designs, workflows, or policy/professional moves. Identify missing opportunities. Proposes concrete improvements. Conceptualize new roles of human and AI. *Example:* "Building the pothole report made me see AI not just as a tool for efficiency, but as a way for cities to respond faster and align more closely with residents' needs. It felt less like a coding exercise and more like creating a new channel that connects people with the systems shaping their daily lives.

Figure 2 visualized how students used and reflected on AI across semesters, showing 21 AI application behaviors and 17 reflection topics (See Appendix Table 1-6 for category definitions and concrete reflection-journal examples). Cell color indicates how frequently each behavior or reflection appeared in student journals. The figure can be read in two ways. Reading across the x-axis highlights the most common ways students applied AI and the topics they reflected on. Reading up the y-axis shows which types of use and reflection were more likely to reach higher levels of criticality. Together, these patterns provide empirical evidence of how students engaged with AI and where deeper critical thinking emerged.

Although the dataset comes from a small pilot course, many of the AI uses and reflections emerged from students' everyday interactions with AI outside the classroom or from their prior experiences. As such, the behavioral and reflection categories captured here reflect a broader pool of generalizable behaviors, which can help instructors anticipate how students may use AI in their coursework and what kinds of reflections are likely to arise during the learning process.

Reading across the x-axis for AI application, students use AI most frequently for tasks under theme *Understanding, Reasoning, and Expression (59 uses)*, followed by *Code Development and Technical Assistance (55 uses)*, *Data and Text Analysis (54 uses), Meta-AI (44 uses), Life Planning and Task Productivity (38 uses),* and *Domain Application (28 uses)*. The top theme includes core learning tasks such as writing, learning new concepts, and distilling content, suggesting that AI played a significant role in supporting students' cognitive learning. While *Domain Application* appears less often, reading across the y-axis shows that this category is associated with higher critical levels, reflecting deeper integration of AI with urban planning knowledge in final projects where students exercised greater discretion.

A similar pattern emerges in reflections when frequency is considered alongside criticality. Themes related to *AI Strengths and Weaknesses* (133 uses) appear most often but cluster at lower critical levels, suggesting that students commonly begin with descriptive evaluations of what AI can and cannot do. In contrast, *Effective and Appropriate Use in Context* (92 uses) and *Human–AI Co-evolution* (82 uses) show strong concentration at levels 3 and 4, indicating a shift from description to judgement and synthesis as students connect AI use to professional responsibilities and identity formation. Categories tied to governance and social implications (i.e., User, Design, and Governance Considerations (21 uses) and Social Costs, Risks, Ethics, and Implications (42 uses)) appear less frequently and show occasional spikes at higher critical levels, but many entries remain at mid or low-level criticality rather than sustained, in-depth critique. This uneven distribution suggests that while students do reach moments of advanced critical engagement, such reasoning is not yet routine. Taken together, this distribution suggests a developmental trajectory: frequent, lower-level reflections anchor early learning, while less frequent but higher-critical reflections mark deeper intellectual and professional growth. More interpretations for Figure 2 are provided in Appendix B.

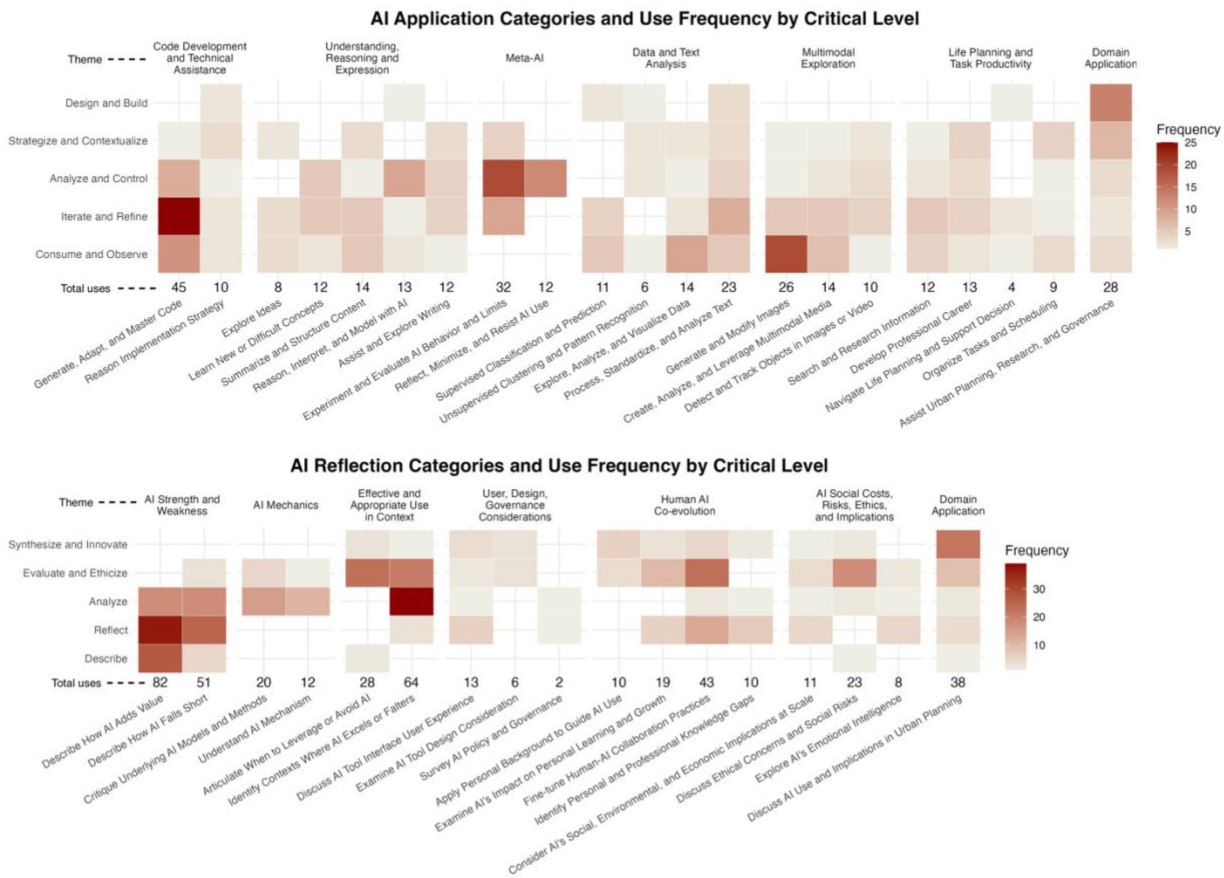


Figure 2: AI Application and reflection categories and use frequency by critical levels. X-axis represents application or reflection categories. Y-axis represents critical levels. Deeper cell color represents high frequency of use and vice versa. Numbers above the x-axis labels are total uses in each category, and the texts on top of the blocks are themes.

DISCUSSIONS AND LESSONS LEARNED

This Urban AI pilot course illustrates a pedagogy design that balances breadth of topics with depth of technical training, leaning more toward capacity building for a technical audience than general literacy. This study identifies three learning goals positioned to highlight the uniqueness of Urban AI, presents a knowledge graph that shows the road map to weave cognitive milestones toward learning goals, and provides a criticality assessment framework to help enhance students' AI use and thinking.

Reflecting on the "Pedagogical Tensions in Urban AI" section, student feedback shows strong appetites for both breadth and depth of the course, appreciating the wide range of topics covered, while being able to build on concrete technical skills in labs and dive into topics of their interest at group and final projects.

In contrast to prior concerns that Urban AI may be "too technical" for planning students, students in this class (a small, self-selected group who had completed at least a programming and a statistics course) with varied levels of technical confidence still enjoyed learning AI fundamentals and the agency to build AI for their purposes. For example, one student reflected that "In a world where AI is suddenly in everything and you're not sure what's real value and what's just hype, learning how these tools fundamentally work was huge… This whole (final) project made me feel way more empowered compared to when I've just used some black-box AI online." Empirical observations from Figure 2 reinforces this point that high criticality emerges not from the most technically intensive uses, but from context-sensitive applications and reflections that connect AI to planning judgement and responsibility. This finding suggests that a course emphasizing AI fundamentals, while focusing evaluation on the criticality of AI use and thinking rather than technical sophistication, can foster diverse yet equally valid pathways to success. This strategy mirrors responsible data science pedagogy, which emphasizes model interpretability as a core learning goal for both technical and non-technical students (Lewis and Stoyanovich 2022).

Student reflections and final projects also demonstrate that integration between technical AI methods and planning context knowledge can emerge organically. For example, multiple students attempted to build AI models to extract zoning codes in the final projects and creatively adapted existing AI models to effectively interpret zoning codes. One student reflected that Natural Language Processing (NLP) tools "were helpful, but only if you already know what you're looking for," noting that they had to draw on their urban planning background to decide which zoning rules mattered, design a planning-relevant dataset, and correct extraction errors. This reflection is also consistent with general patterns observed in Figure 2 where *Domain Application* appears less frequently but consistently aligns with higher critical levels, indicating that meaningful integration of planning knowledge and AI tends to produce deeper learning.

The strategy of encouraging students to *learn with AI* also proved effective to motivate *learning about AI*. Introducing ChatGPT in Week 2 and prompting students to experiment with AI tools through weekly reflection journals created a feedback loop between interaction and inquiry. Students reported that hands-on experimentation raised questions about AI limitations (e.g., hallucinations) that motivated them to seek deeper understanding of AI mechanics. Likewise, learning about model processes prompted further tool experimentation and reflection the following week, suggesting a synergistic cycle between knowledge acquisition and tool use.

Despite these successes, final projects revealed a strong preference for commercial AI tools (especially the ChatGPT API) over open-source alternatives, and for application-focused projects over those examining policy or governance. Of all AI tool instances reported in the journals, 62% is ChatGPT. These preferences may reflect both student interests and the relative ease and abundance of resources for commercial, application-oriented approaches. This mirrors a broader research bias toward novel AI applications over critical examinations of AI's implications. This preference is also reflected in Figure 2, where governance- and ethics-related reflections appear less frequently and spread across different levels, indicating that without deliberate instructional prompting, students gravitate toward application rather than in-depth critique.

Beyond reflection journals, classroom interactions created teachable moments in which technical uncertainty and ethical risk became opportunities for developing professional judgement. In one lab, a widely used Hugging Face open-source model was removed days before class, requiring an improvised replacement. Rather than treating this as a logistical failure alone, the incident surfaced a critical lesson: AI systems are contingent, unstable, and shaped by platform governance beyond instructors' control. Discussing this with students reframed the event as part of professional preparation where planners must evaluate reliability, anticipate breakdowns, and make pragmatic decisions about when to rely on open, proprietary, or hybrid tools. In another incident, an AI auditing assignment prompted students to confront the real-world consequences of technical experimentation. One group initially designed a chatbot audit that would send hundreds of automated queries to a city-service system. After reflection, several students chose not to pursue this method in the final projects, recognizing the potential to disrupt public services. This decision illustrates advanced professional judgement: choosing restraint based on ethical and civic considerations rather than technical feasibility. Such moments demonstrate how Urban AI education extends beyond skill-building toward cultivating responsibility, risk awareness, and accountability in practice.

The framework and lessons derived from the pilot course have some limitations. The small class size (n = 20) increases variability in findings, and the course design assumes a minimum level of student technical preparation. The technically prepared student population may explain the inclination toward application and outstanding performance in integrating domain knowledge and technical skills. Programs without a sufficient pool of technical students, or without a supporting technical course sequence, may want to adapt and use existing AI tools to help future planners raise awareness on how AI might bring new urban challenges and planning implications (see Appendix C for a non-technical version of the course). For these reasons, the proposed framework should be understood as a flexible learning scaffold rather than a rigid grading system. Future research can gather more teaching examples across institutions for a comparative study and use the proposed criticality framework to assess student performance across modules.

While this paper offers one model for teaching Urban AI, it does not imply every planning program should open a course. Urban AI is best taught as the last piece of the technical sequence, not only because it builds on technical foundations, but because effective planners must first be capable of evidence-based planning and data governance before engaging with AI. For many local governments eager to adopt AI, the most responsible first step is strengthening their data infrastructure and governance practices, an area where planners can play a pivotal role in shaping policy and direction.

Thus, this paper calls for more educators to cultivate materials that support bold experimentation and critical consciousness around AI. Effective and appropriate AI governance is still rare in planning, partly because planning demands strong alignment with values which are areas of significant technical, human, and systemic challenges. As a result, there is a pressing need for pedagogical resources focused on auditing and governing AI to build student interest and capacity in policy-oriented work. This need aligns with broader movements toward critical and creative AI pedagogy (Atenas, Havemann, and Nerantzi 2025), including playful tools such as *A Planner's AI Dilemma Cards*, which assemble concrete examples, case studies, and questions that mirror the AI ethics and governance challenges planners face daily (Boyco and Robinson 2025). This pilot course leads by example through sharing its syllabus (LINK REDACTED), encouraging students to develop peer tutorials, assigning a group project on auditing AI chatbots, and inviting students to publish final projects, underscoring the importance of Urban AI education in shaping future planners.

ACKNOWLEDGEMENT

I would like to thank University of Michigan Center for Research on Learning and Teaching, Taubman College, and Robert Goodspeed (Chair of the Urban and Regional Planning Program) for providing funding to support the curriculum development; Qifan Wu for drafting course lab materials; Gene Leynes, Varun Adibhatla, Heyu Huang, Scott Campbell, Sarah Williams, and Anthony Vanky for offering their valuable perspectives in the Urban AI Pedagogy Workshop which sets the tone for the pilot course. I am particularly grateful for the twenty students in my Urban AI course who did their best to engage with the course and share their reflection journals with me for research purposes. Their initials are listed below: E.A., S.B., T.B., M.B., P.C., S.C., T.C., N.C., M.D., A.H., Z.H., I.K., S.S., T.W., X.T., J.W., X.W., H.X., M.Y., R.Z.

REFERENCE

Atenas, Javiera, Leo Havemann, and Chrissi Nerantzi. "Critical and creative pedagogies for artificial intelligence and data literacy: an epistemic data justice approach for academic practice." *Research in Learning Technology* 32 (2025).

Axelrod, Judah, Lydia Lo, and Sara C. Bronin. *Automating Zoning Data Collection: Results from a Pilot Effort to Automate National Zoning Atlas Methodologies*. Urban Institute, 2023.

Biesta, Gert JJ. *Good education in an age of measurement: Ethics, politics, democracy*. Routledge, 2015.

Brinkley, Catherine, and Carl Stahmer. "What is in a plan? Using natural language processing to read 461 California city general plans." *Journal of Planning Education and Research* 44, no. 2 (2024): 632-648.

Boukerche, Azzedine, and Zhijun Hou. "Object detection using deep learning methods in traffic scenarios." *ACM Computing Surveys (CSUR)* 54, no. 2 (2021): 1-35.

Clayton, Paige, Robert Goodspeed, Jamaal Green, Allison Lassiter, William Riggs, and Bev Wilson. "More than analytics: Five approaches to educating professionals to shape today's digital cities." *Journal of Planning Education and Research* (2024): 0739456X241261372.

Deng, Ruiqi, Maoli Jiang, Xinlu Yu, Yuyan Lu, and Shasha Liu. "Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies." *Computers & Education* (2024): 105224.

Drummond, William J., and Steven P. French. "The future of GIS in planning: Converging technologies and diverging interests." *Journal of the American Planning Association* 74, no. 2 (2008): 161-174.

Dueker, Kenneth J. "Geographic information systems and computer-aided mapping." *Journal of the American Planning Association* 53, no. 3 (1987): 383-390.

Fu, Xinyu, Chaosu Li, and Wei Zhai. "Using natural Language Processing to read plans: a study of 78 resilience plans from the 100 resilient cities network." *Journal of the American Planning Association* 89, no. 1 (2023): 107-119.

Freire, Paulo. Pedagogy of the Oppressed. Translated by Myra Bergman Ramos. 50th Anniversary Edition. London: Bloomsbury Academic. 2000.

French, Steven P., and Lyna L. Wiggins. "Computer adoption and use in California planning agencies: Implications for education." *Journal of Planning Education and Research* 8, no. 2 (1989): 97-108.

Godschalk, David R., and Gerard McMahon. "Staffing the revolution: GIS education for planners." *Journal of Planning Education and Research* 11, no. 3 (1992): 216-226.

Lee, Jongwon. "Generative AI and geography education." In *GeoAI and human geography: The dawn of a new spatial intelligence era*, pp. 297-312. Cham: Springer Nature Switzerland, 2025.

Lehmann, Matthias, Philipp B. Cornelius, and Fabian J. Sting. "AI meets the classroom: When does ChatGPT harm learning?." *Available at SSRN 4941259* (2024).

Lewis, Armanda, and Julia Stoyanovich. "Teaching responsible data science: Charting new pedagogical territory." *International Journal of Artificial Intelligence in Education* 32, no. 3 (2022): 783-807.

Huang, Xiao, Siqin Wang, John Wilson, and Peter Kedron. *GeoAI and Human Geography*. Springer Nature, 2025.

Ito, Koichi, Yuhao Kang, Ye Zhang, Fan Zhang, and Filip Biljecki. "Understanding urban perception with visual data: A systematic review." *Cities* 152 (2024): 105169.

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development* (Vol. 1). Englewood Cliffs, NJ: Prentice-Hall.

Kontokosta, Constantine E. "Urban informatics in the science and practice of planning." *Journal of Planning Education and Research* 41, no. 4 (2021): 382-395.

Montagu, A. Simon. "Repackaging the revolution: Making GIS instruction relevant to planners." *Journal of Planning Education and Research* 21, no. 2 (2001): 184-195.

Morley, David. (2022). "Zoning for Data Centers and Crytocurrency Mining". American Planning Association – Zoning Practice". https://www.planning.org/publications/document/9249180/

Nichols, Carlie. (2025). "The Physical Footprint of Artificial Intelligence". American Planning Association – Zoning Practice. https://www.planning.org/publications/document/9318086/

Planning Accreditation Board. (2025). "PAB Accreditation Standards and Criteria". https://www.planningaccreditationboard.org/wp-content/uploads/2025/02/Amended2022Stds.pdf

Quan, Steven Jige. "Urban-GAN: An artificial intelligence-aided computation system for plural urban design." *Environment and Planning B: Urban Analytics and City Science* 49, no. 9 (2022): 2500-2515.

Boyco, Morgon, and Pamela Robinson. “A Planner’s AI Dilemmas Cards”. Canadian Institute of Planners and Ontario Professional Planners Institute National Planning Conference. (2025).
https://www.researchgate.net/publication/393479227_A_Planner's_AI_Dilemmas_Cards

Salazar-Miranda, Arianna, and Emily Talen. "An AI-based analysis of zoning reforms in US cities." *Nature Cities* (2025): 1-12.

Sanchez, Thomas W., Hannah Shumway, Trey Gordner, and Theo Lim. "The prospects of artificial intelligence in urban planning." *International journal of urban sciences* 27, no. 2 (2023): 179-194.

Sanchez, Thomas W., Marc Brenman, and Xinyue Ye. "The ethical concerns of artificial intelligence in urban planning." *Journal of the American Planning Association* 91, no. 2 (2025): 294-307.

Schon, Donald A. *The Reflective Practitioner: How Professionals Think in Action*. New York: Basic Books, 1984.

Tu, Xinming, James Zou, Weijie J. Su, and Linjun Zhang. "What should data science education do with large language models?." *arXiv preprint arXiv:2307.02792* (2023).

Wu, Abraham Noah, Rudi Stouffs, and Filip Biljecki. "Generative Adversarial Networks in the built environment: A comprehensive review of the application of GANs across data types and scales." *Building and Environment* 223 (2022): 109477.

Ye, Xinyue, Tan Yigitcanlar, Michael Goodchild, Xiao Huang, Wenwen Li, Shih-Lung Shaw, Yanjie Fu, Wenjing Gong, and Galen Newman. "Artificial intelligence in urban science: why does it matter?." *Annals of GIS* 31, no. 2 (2025): 181-189.

Zheng, Yu, Fengli Xu, Yuming Lin, Paolo Santi, Carlo Ratti, Qi R. Wang, and Yong Li. "Urban planning in the era of large language models." *Nature computational science* (2025): 1-10.

APPENDIX

Appendix A: ChatGPT Use in Qualitative Coding

Appendix B: More Interpretations for Figure 2

Appendix C: A Non-Technical Roadmap for Urban AI

Appendix Table 1: AI Tool and Model Application by Theme and Definition.

Appendix Table 2: AI Tool and Model Application by Critical Level.

Appendix Table 3: Simplified Version - AI Tool and Model Application by Critical Level.

Appendix Table 4: AI Tool and Model Reflection by Theme and Definition.

Appendix Table 5: AI Reflection by Critical Level.

Appendix Table 6: Simplified Version - AI Reflection by Critical Level.

Appendix A: ChatGPT Use in Qualitative Coding

ChatGPT-4o was used as a human-in-the-loop assistant to accelerate qualitative coding of student reflection journals. Specifically, it suggested category names, refined category definitions based on sets of quotes, and helped the author deliberate edge cases when determining the most appropriate category. The goal was to assign each quote an AI application or reflection category and a critical level.

The motivation for using ChatGPT-4o was to speed up analysis and maintain consistency across the coding process. A total of 655 quotes from 235 journals were coded, many reviewed multiple times during the iterative development of categories. This amounted to substantial cognitive work that, in a traditional approach, might require multiple trained research assistants whose consistency can also vary. The human-in-the-loop approach with AI ensured all decisions ultimately went through the author as the final judge.

The author carefully considered where AI could be used to maximize benefits and minimize risks. Coding proceeded in three rounds.

Round 1 identified unit of analysis (distinct text chunks) and captured emerging themes. This stage was conducted manually because humans are better than AI at identifying boundaries of unique instances of applications or reflections, and the process exposed the author to the raw data without bias from AI-generated labels.

Round 2 refined and defined the codes. The author reviewed all emerging categories, adjusted the schema, and then reassigned quotes using AI for initial suggestions, with all assignments confirmed through human judgment.

Round 3 used AI to generate finalized definitions for each category–criticality combination (e.g., Level 1 for *Generate, Adapt, and Master Code*) based on the assigned quotes. The author then revisited all quotes to ensure alignment with the final definitions and conducted a sanity check to ensure each quote had no more than one application and one reflection category.

Example prompt templates used in ChatGPT 4o are attached below:

------------------For Classification-----------------------

Task: Classify the quotes below under the category "{{CATEGORY}}".

Constraints:

- Use the unique numbers in front of each quote as ID.

- Use this rubric:

Descriptive = Plain what/was done

Reflective = Personal pros/cons, light tweaks

Analytical = Why/how patterns, mechanisms, context

Evaluative & Ethicize = Judgments, trade-offs, ethical stance

Synthesis & Innovate = New workflows/policies, concrete improvements

Output as a table with columns:

[ID | Category | Assigned Critical Level | Rationale (1–2 sentences)]

Quotes:

{{QUOTES_BLOCK}}

------------------For Generating Definitions-----------------------

Task: For category "{{CATEGORY}}", write category-specific definitions for each critical level that appears in the quotes below.

Constraints:

- Filter by the quotes provided; do NOT include levels with zero quotes.

- For each level present, provide:

(a) "Definition (Category-Specific)"

(b) "How the Quotes Show It" → list 3–5 distinct patterns observed.

- Keep language specific, not generic. Tie patterns to what people actually did/said.

Quotes:

{{QUOTES_BLOCK}}

Return as a compact table:

[Critical Level | Definition (Category-Specific) | How the Quotes Show It (bullet list)]

The author recognizes that AI is prone to hallucination. To mitigate this, every assignment (category and critical level) was manually verified and tagged with its corresponding quote in ATLAS.ti. As such, all coded results remain grounded in actual text and can be traced back to specific quotes.

The author also acknowledges that AI performance can deteriorate when the context window becomes long. To address this, classification and summarization were conducted at modular levels—by category and by critical level—limiting both the number of quotes and the amount of reference information included in each prompt.

Finally, the author notes that large-scale reliability of AI-generated results remains under-tested. When the author checked for sample quotes, AI was generally consistent for unambiguous cases but varied on edge cases. The author argues that such ambiguity also affects human coders, and with the author serving as the final decision-maker, edge cases were still placed into acceptable categories.

For instance, here is a quote that AI struggles to classify for application in prior iterations. It does not fit into prior categories such as (code) **debugging** or **idea generation**, because the student isn't just fixing code or passively brainstorming. They're actively engaging in a dialogue with AI to challenge its assumptions and reinterpret incomplete outputs. This prompts the author to develop a new application category called **Reason Implementation Strategy** to better capture this type of nuanced usage.

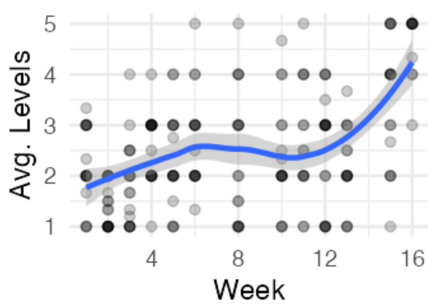
"This week, I used AI tools in a few different ways while working on my final project. One of the most helpful aspects was being able to brainstorm with ChatGPT when my original approach wasn't working as expected. I had planned to use the GNews API to automatically gather articles about opposition to agrivoltaics in Wisconsin, but the results weren't producing the kind of localized or specific content I was looking for. ChatGPT tried to tell me that the lack of results around opposition to agrivoltaics in the GNews searches was meaningful, but I didn't agree. It used the logic that no results meant no articles or no opposition, but based on my personal experience, some emails with [faculty] I knew there was NIMBY related opposition out there, and my model simply wasn't catching it. So, I explained my thoughts to ChatGPT which also helped me think through alternative strategies, like expanding the geographic focus to the entire Midwest and refining my search terms to target zoning and land use disputes. It was a good reminder of the role of human processing in using AI tools, as ChatGPT was ready to accept the lack of opposition results as a passive acceptance of renewable energy by the Midwest, which obviously is not accurate."

Overall, the study focuses on the broader frameworks and patterns observed across student reflections rather than the exact frequency counts, which are sensitive to the small class size and specific student population. Thus, marginal AI errors in exact classification have minimal impacts on the study's results.

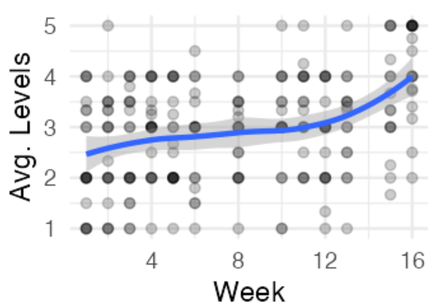
Appendix B: More Interpretations for Figure 2

Low or absent use in certain categories can have several explanations. Some application categories, such as those in *Life Planning and Task Productivity*, were not explicitly introduced in the course, which may indicate limited instructional prompting rather than lack of relevance. Some categories are also narrowly defined; for example, *Reflect, Minimize, and Resist AI Use* captures moments when students intentionally reduced or resisted AI use or reflected on their overall usage patterns. By definition, this category aligns closely with the *Analyze and Control* critical level.

Similarly, some reflection categories were deliberately designed to scaffold learning progression. For instance, *Describe How AI Adds Value/Falls Short* represents a basic account of AI's strengths and weaknesses and typically appears at lower critical levels. More advanced categories, such as *Articulate When to Leverage or Avoid AI* or *Identify Contexts Where AI Excels or Falters*, require integrating domain knowledge and making normative judgments and thus tend to land on high criticality. This scaffolding reflects the knowledge flow shown in the Urban AI Knowledge Graph (Figure 1), where students begin with basic domain knowledge and gradually build a more nuanced understanding of AI's capacities and constraints, eventually forming personal stances. The following subgraphs also show a steady but varied progression of such growth throughout the semester.



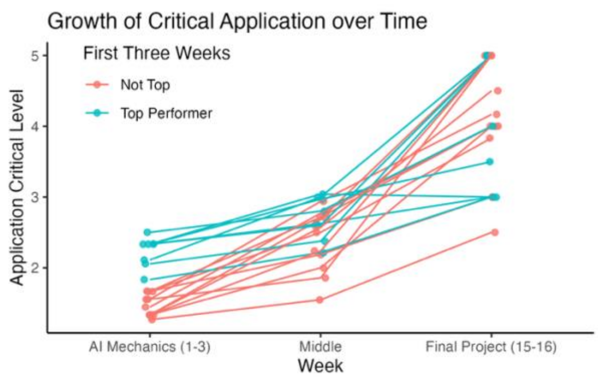
(Average Application Criticality Over Weeks for the Class)

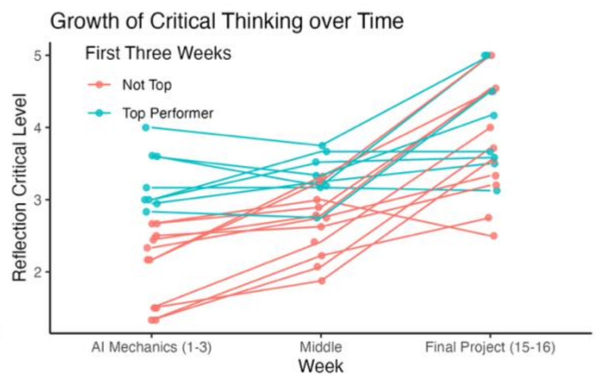


(Average Reflection Criticality Over Weeks for the Class)

The following subfigures are generated to test the hypothesis descriptively that whether students who are top performers in learning AI mechanics module (i.e., denoted as higher than medium average critical levels in reflection journals) will consistently outperform others in final projects. The reasoning behind is that students who are top performers (with higher criticality) in the AI mechanics module are likely to be more technical in class and thus perform better in the final projects. If the hypothesis stands, the graphic will show the cyan lines (Top Performers) to be consistently on top of the pink lines (Not Top).

However, the results reject the hypothesis and reveal that being top performers (and potentially more technical) do not necessarily lead to higher criticality in the final project. Many less technical students, despite having slightly lower average criticality in technical learning, can still achieve success in final projects and demonstrate high criticality. This evidence is used to support the following statement in the manuscript: "a course emphasizing AI fundamentals, while focusing evaluation on the criticality of AI use and thinking rather than technical sophistication, can foster diverse yet equally valid pathways to success."





Appendix C: A Non-Technical Roadmap for Urban AI

This Urban AI course module table is designed for a less technical audience. This design presumes that the learning outcomes are the same, yet the emphasis is on establishing a cognitive coordinate system to position technical knowledge (or technologies) in relation to planning problems and appropriate contexts. The skill outcomes also shift from "building models" to "use, interpret, and adapt AI tools" for students with little coding background.

Modules	**Knowledge Outcomes**	**Skills Outcome**	**Sample Topics / Activities**
#1 AI Mechanics	Explain what AI and machine learning are in everyday terms; Develop a "mental map" of different AI tools and what tasks they are suited for (e.g., language, images, prediction); Understand the basic idea of how AI learns (training, testing, feedback); Recognize AI's capabilities and limitations	Use no-code AI tools; Experiment with pre-built models to see how data, prompts, and settings shape outputs; Interpret results and identify strengths, weaknesses, and risks	Machine Learning and GenAI Sandbox (e.g., Teachable Machine; ChatGPT; Adobe Firefly).

#2 Urban Application	Understand how AI can support planning tasks and how planning work may change; Recognize opportunities and limits of AI in real planning contexts.	Explain how AI is applied across planning domains; Identify fragile claims and implementation risks; Design simple prompt-based workflows for planning tasks; Use AI to support tasks such as brainstorming, organizing data, literature search (with verification), and design exploration	Case Studies and Guest Lectures (e.g., GeoAI; Traffic Video Detection; NLP for Public Comments; Image Recognition for Blighted Property)
#3 Governance and Society	Understand how technical limitations can produce social consequences (bias, surveillance, loss of local knowledge); Recognize alternative AI development models (e.g., open-source, smaller, distributed systems); Understand how cities procure and deploy AI systems through vendors and contracts	Discuss ethical and governance challenges in practical planning language; Articulate what contexts and domain challenge makes AI easy to fail; Evaluate tradeoffs in real deployment decisions	Case Studies and Guest Lectures (e.g., NYC chatbot; civic organization that used AI to solve community challenges, etc.); Play with "A Planner's AI Dilemma Cards" to gain concrete views of how ethical challenges show in the practice work.
#4 AI's Systematic Impacts	Anticipate AI's systematic impacts on the environment, labor, economy, and human dynamics	Recognize key domains where AI affects urban systems (e.g., environment, labor, economy); Gain familiarity with current debates and planning considerations; Practice translating technology advancement (e.g., data center infrastructure or AI models) into implications for planning jobs.	Environmental Impact of AI; Data Center; Geography of GenAI labor; Guest Lectures Student design appropriate causal links and use AI as a conversational, scenario planning assistant to expand immediate constraints into future, long-term implications.
#5 AI Practice and Reflection (cross-cutting as background)	Practice five-level of critical AI use and AI thinking as proposed by this study; Understand how professional identity shapes AI use	Interact with AI tools to understand AI behaviors; Evaluate AI outputs for quality, relevance, and bias; Customize AI models through web interface or prompt engineering (no-code); Find communities (e.g., APA organizations, thought leaders in planning) for continuous self-learning.	Reflection journals; Case study review and assessment; Student & instructor examples of appropriate AI use

Appendix Table 1: AI Tool and Model Application by Theme and Definition.

The definitions are grounded in examples from AI reflection journals in the pilot course.

AI Tool/Model Application	Theme	Definition
Generate, Adapt, and Master Code	Code Development and Technical Assistance	Use of AI tool(s) to generate, adapt, debug, and explain code, refresh coding syntax, and generate edge and test cases
Reason Implementation Strategy	Code Development and Technical Assistance	Use of AI tool(s) to design and refine a project's implementation approach, such as overall strategies and information architectures and key decisions about which models and methods to use for analysis or prototyping.
Explore Ideas	Understanding, Reasoning and Expression	Use of AI tools to brainstorm, expand, or refine early-stage ideas—often for creative, conceptual, or design-oriented tasks.
Learn New or Difficult Concepts	Understanding, Reasoning and Expression	Use of AI tools to support learning, memorization, and comprehension of unfamiliar, abstract, or complex concepts across technical, academic, or real-world domains; to actively review, recall, and rehearse knowledge in preparation for assessments or public speaking; to reframe explanations given specific audience
Summarize and Structure Content	Understanding, Reasoning and Expression	Use of AI to distill, restructure, or visualize key information from longer or complex texts (e.g., PDFs, readings, datasets) for improved comprehension or learning, including generating summaries, study guides, timelines, mind maps, or restructuring user-generated content.
Reason, Interpret, and Model with AI	Understanding, Reasoning and Expression	Use of AI tools to verify the correctness of assignment answers, validate reasoning, and build confidence in one's approach to academic tasks; to analyze, explain, or contextualize results from data analysis, model outputs, or visualizations; to support the construction, testing, or refinement of logical, mathematical, or system-based models; to explore and evaluate the causes of unexpected or incomplete results, to rethink problem framing, or to co-develop alternative strategies when original approaches fail
Assist and Explore Writing	Understanding, Reasoning and Expression	Use of AI tools to improve the clarity, fluency, tone, grammar, or stylistic quality of written language; to compose original written content—such as paragraphs, essays, or sections of text—based on user-provided inputs like notes, outlines, or prompts; to generate, edit, or verify citations in specific academic or professional styles and assist with formatting tasks related to document layout; to produce a structured outline or suggested framework for a piece of writing, presentation, or project based on user-provided prompts, drafts, notes, or assignment instructions

AI Tool/Model Application	Theme	Definition
Experiment and Evaluate AI Behavior and Limits	Meta-AI	Use of AI tools to systematically test, manipulate, or observe model behavior, limitations, and internal mechanics through experimentation.
Reflect, Minimize, and Resist AI Use	Meta-AI	Use of AI tools was intentionally avoided, limited, or deemed unnecessary due to contextual, personal, ethical, or task-specific reasons; was used to summarize past AI use patterns for reflection.
Supervised Classification and Prediction	Data and Text Analysis	Use of AI tools to train, evaluate, and apply machine learning models on labeled data to classify outcomes or predict target variables, such as selecting appropriate models on labeled datasets, applying models to detect or classify patterns in new data, and using tools like Teachable Machine for real-time classification (e.g., poses, objects, or voices)
Unsupervised Clustering and Pattern Recognition	Data and Text Analysis	Use of AI or machine learning tools to detect hidden structures, groupings, or patterns in datasets without predefined labels, such as categorizing qualitative or quantitative data into meaningful clusters (e.g., archetypes or themes), recognizing recurring patterns across diverse inputs (e.g., interview transcripts, behavioral data, or tabular datasets), and applying advanced models like XGBoost to uncover complex, non-linear relationships.
Explore, Analyze, and Visualize Data	Data and Text Analysis	Use of AI tools to explore, analyze, and visually represent structured or semi-structured datasets; to generate mock data for simulation
Process, Standardize, and Analyze Text	Data and Text Analysis	Use of AI tools to explore, understand, or navigate textual content through question answering, summarization, or explanation; to extract, structure, and standardize information from unstructured or semi-structured text for further analysis, comparison, or integration; to conduct NLP tasks such as sentiment analysis.
Generate and Modify Images	Multimodal Exploration	Use of AI tools to create, alter, or refine static images or visual scenes, often based on textual prompts or selected areas within images.
Create, Analyze, and Leverage Multimodal Media	Multimodal Exploration	Use of AI tools to generate and analyze original artistic, musical, narrative, or multimedia content where the primary goal is creative expression, storytelling, entertainment, or conceptual exploration; to convert non-text media—such as audio or video—into written transcripts.
Detect and Track Objects	Multimodal Exploration	Use AI to identify and follow objects in visual data — from simply reporting detection results, to experimenting with pre-

AI Tool/Model Application	Theme	Definition
in Images or Videos		trained models and parameters, systematically testing hypotheses about model performance, and iteratively improving detection and tracking for context-specific tasks such as satellite imagery analysis or pedestrian behavior studies.
Search and Research Information	Life Planning and Task Productivity	Use of AI tools as a digital research assistant to retrieve, explore, or synthesize information across academic, professional, and everyday domains.
Develop Professional Career	Life Planning and Task Productivity	Use of AI tools to support professional development activities such as exploring career interests, improving job application materials, preparing for interviews, navigating salary negotiations, and understanding industry tools and expectations.
Navigate Life Planning and Support Decision	Life Planning and Task Productivity	Use of AI tools to support personal decision-making and future planning by generating structured options, evaluating trade-offs, and clarifying complex information
Organize Tasks and Scheduling	Life Planning and Task Productivity	Use of AI tools to plan, prioritize, and manage time-bound responsibilities across academic, personal, and professional domains.
Assist Urban Planning, Research, and Governance	Domain-specific Application	Use of AI to support planning, research, and governance by observing or interacting with planning-oriented tools, summarizing or analyzing planning documents and data, evaluating domain-specific AI applications, and adapting or building new models and systems tailored to urban challenges.

Appendix Table 2: AI Tool and Model Application by Critical Level.

The definitions are grounded in examples from AI reflection journals in the course. Empty cells can mean one of the two things: 1) category definition determines that critical level is unlikely to be observed, 2) no examples are observed in the AI flection journals.

AI Tool/Model Application	Level 1	Level 2	Level 3	Level 4	Level 5
Theme: Code Development and Technical Assistance					
Generate, Adapt, and Master Code	generate code from scratch; explain codes from documentation or others; add comments to	adapt/debug existing code to work or generate / explain code as part of an iterative code development	learn and understand why code does not work just problem-solving; refresh memories of	develop edge and test cases to contextualize function uses and improve code robustness	

AI Tool/Model Application	Level 1	Level 2	Level 3	Level 4	Level 5
	code; show minimal iterations or description of the interaction	process that builds on prior efforts	coding syntax rather than generate solutions		
Reason Implementation Strategy	develop a strategy or information architecture to approach a big project; understand possible implementation paths to build prototypes and conceptualize projects	consult how to implement a model building on previous efforts; iteratively debug and improve implementation strategy throughout the project	propose specific implementation steps and ask AI to piece them together	steer strategic decisions on which models to use, why current implementation does not work well and places for improvement	implement the strategy for analysis and prototypes
Theme: Understanding, Reasoning and Expression					
Explore Ideas	brainstorm ideas and examples with limited user inputs	iterate on prompts, model setting, and model inputs to improve exploration results		support creative, imaginative work (e.g., naming characters in speculative fiction, world construction in novel writing) under a user-defined ideation world	
Learn New or Difficult Concepts	explain a new or difficult concept	explore multimodal functions (e.g., create a PPT) to explain a new concept; develop quiz packets or prepare presentations	understand difficult concepts in an iterative problem-solving process; double-check on conceptual understanding; aid learning by generating examples		
Summarize and Structure Content	summarize key points from long texts (e.g., readings, lecture notes, etc.);	summarize long texts into multimodal formats (e.g., mind maps);	summarize and structure long texts with user-specified goals and human	develop a summarization strategy (e.g., how and what to summarize) that	

AI Tool/Model Application	Level 1	Level 2	Level 3	Level 4	Level 5
	explore an AI-assisted notetaking app that summarize other people's notes for a given textbook; review user experience of text processing apps such as NotebookLM	interact with contents from summarized texts, such as generating Q&A and study guides (e.g., using Notebook LM) from readings.	validate the quality of summary	makes it easy for human validation for a specific context; use summarization as a part of the AI tool building to support strategic goals (e.g., summarize user inputs for pothole reporting; summarize news article for sentiment analysis)	
Reason, Interpret, and Model with AI	provide reasoning for a problem (e.g., solve a math problem); interpret statistical results	debug reasoning or logical faults (e.g., incorrect mathematical proof) from prior efforts	validate writing /coding/model selection /problem-solving reasoning against human thoughts; double-check on statistical interpretations		
Assist and Explore Writing	generate writing with minimal inputs; lack of details in interactions	generate writing based on a user-provided outline; generate a writing outline based on user-provided, prior contents; rewrite paragraphs with specific goals / tones; play with writing tones and styles; derive personal writing styles given user-provided writing samples	generate an outline to double-check on existing writing to ensure argument comprehensiveness; check grammar; format citations	adapt writing to meet specific audience (e.g., simplify concept explanations for five years old); experiment with an AI app to convert AI-generated texts to human tones or apply tone-checker for professional emails	
Theme: Meta-AI					
Experiment and Evaluate AI		experiment with prompts to	experiment with AI model	conduct scientifically	

AI Tool/Model Application	Level 1	Level 2	Level 3	Level 4	Level 5
Behavior and Limits		observe changing AI behaviors and outcomes; play with AI persona that changes tone in generated texts; compare AI tools and models based on user experience (without explicit criteria); use AI platforms to evaluate different AI models side-by-side; mention interactions with model parameters but lack details	parameters and observe changes; test and iterate model performance with hypotheses and specific goals; compare AI tools and models with explicit criteria (e.g., accuracy at a specific task)	designed, systematic prompt experiments to evaluate AI performance for a contextual goal (e.g., assess response biases by feeding carefully varied prompts to the NYC Chatbot); experiment with AI model parameters to enhance contextual use (e.g., detect trees from satellite imagery)	
Reflect, Minimize, and Resist AI Use			minimize or resist AI use; summarize past AI use history to reflect use patterns		
Theme: Data and Text Analysis					
Supervised Classification and Prediction	generate code for training a supervised model as instructed in the lab exercise; explore an AI application in a museum that uses supervised classification; mention the use without interaction details	interact with teachable machine as directed in class to explore how supervised classification works			build or apply new, sophisticated AI models and code pipeline for classification and prediction on domain-specific topics (e.g., recognize different objects from traffic videos; predict bike sharing patterns based on prior data)
Unsupervised Clustering and	mention the use without		identify categories or derive themes	classify people, text, and data as a strategic step	build or apply new, sophisticated AI models and code

AI Tool/Model Application	Level 1	Level 2	Level 3	Level 4	Level 5
Pattern Recognition	interaction details		from user provided inputs (e.g., interview transcripts, crosswalk behaviors in videos)	to reach the outcome (e.g., generate a hierarchical affinity diagram)	pipeline to derive cluster categories and facilitate knowledge discovery from data (e.g., categorize and assess US dams; understand skill ecosystems from urban tech job posts)
Explore, Analyze, and Visualize Data		iteratively explore, analyze, or visualize data with user provided instructions	refine instructions based on failure hypotheses to improve outcomes; validate AI's outcomes with human-led / existing analyses	generate mock data to enrich training set in a data-scares planning domain	
Process, Standardize, and Analyze Text	interact with text-based AI software or chatbot and describe user experience; complete text extraction lab exercise as instructed without further details; vague description of what models are used and how	interact with hugging face pre-trained models for text Q&A or sentiment analysis; Interact with text-processing AI (e.g., Notebook LM) to interactively summarize and QA for specific tasks; Improve prompts to improve text-processing, text extraction, and theme inference outcomes	improve text extraction and standardization outcomes via model parameters (e.g., chunk size) and human-in-the-loop validation; analyze performance of hugging face pre-trained text-based and sentiment analysis models given different types of sentences with specific benchmarks	strategize use of text extraction in building an AI system to reach a specific goal (e.g., extract key information from pothole reports); read research paper to explore how to compare different types of semantic data and text-based emotional expression in the context of public space analysis; try different strategies to analyze sentiments from Airbnb posts	develop code pipeline to extract specific, structured information from zoning document; develop AI process to identify topical themes from news article for energy planning context
Theme: Multimodal Exploration					

AI Tool/Model Application	Level 1	Level 2	Level 3	Level 4	Level 5
Generate and Modify Images	generate and modify images based on prompts; create and fuse new image given image inputs (e.g., create baby faces from adult pictures); perform image activities as requested in class	experiment with prompts and input images to refine outputs; explore a model again after functions update	analyze examples of well-curated generative art examples and their associated prompts from a museum	use AI-generated images as scene inputs in the process of video creation and editing	
Create, Analyze, and Leverage Multimodal Media	create multimodal media (other than images) with simple prompts; play AI games; transcrip t texts from video and summarize	iteratively refine multimodal media such as video, songs, mind maps, etc.; use AI voice mode to simulate conversations to prepare for class presentation	analyze AI-created multimedia (e.g., music) with experiments		
Detect and Track Objects in Images or Videos	simply report detection results with no interaction details or analyses.	explore object detection models (built by others or a pre-trained model) and its parameters to observe changing outcomes	systematically change user inputs and targeted outputs with clear hypothesis of why performance falls short and how to improve	iteratively improve object detection performance for contextual use (e.g., detect trees in satellite imagery); conduct literature review to find AI methods suitable for contextual use (e.g., effectively track pedestrian behaviors in public space videos)	
Theme: Life Planning and Task Productivity					

AI Tool/Model Application	Level 1	Level 2	Level 3	Level 4	Level 5
Search and Research Information	search/research to understand a specific concept, code, or topic with limited human interactions; use deep research function in ChatGPT	refine search queries in AI to reach desirable outcomes; support technical tool, software learning, life planning, or decision deliberation in an iterative Q&A process	use AI search when Google search is exhausted and compared performance	strategic search to support contextual goals, such as interview preparation	
Develop Professional Career	review a resume reviewing tool from HR perspectives; upload resume to AI to refine texts and match job description; align cover letters to different jobs	refine with personal take, such as skills to emphasize, ATS-friendly, specific bullet points to improve, and tones appropriate for reach out emails	analyze and critique underlying models behind career recommendation software	mock interviews with specific inputs	
Navigate Life Planning and Support Decision	consult a life plan in a hypothetical scenario (e.g., got a job; move to a city); interpret and explain life documents (e.g., insurance, tax)	iterate on the life plan, learn new concepts, and deliberate pro and cons under different conditions (e.g., whether to buy a EV car)			Build an AI application to support life planning (e.g., an AI-assisted itinerary planning app)
Organize Tasks and Scheduling	prioritize and schedule tasks without a clear optimization goal or descriptions of human interactions; misidentify or unclearly identify AI in the scheduling function of an app	explore AI calendar and task scheduling apps with a clear goal, providing detailed descriptions of user interactions to illustrate iterations	assist in small parts of task planning (e.g., come up with task names) while human do the main planning	organize tasks, schedule time, arrange courses, and design a financial plan while accounting for specific personal needs, constraints, and defined goals; reflect on and evaluate performance	

AI Tool/Model Application	Level 1	Level 2	Level 3	Level 4	Level 5
				against human expectations	
Theme: Domain-specific Application					
Assist Urban Planning, Research, and Governance	observe an AI tool for planning audience (e.g., a web app teaching planners how to write better prompts); brainstorm planning suggestions with limited inputs and contexts; summarize planning documents into tables without further iterations	interact with pre-trained and/or domain-specific AI models others have built (e.g., AI models built by peers through mandatory class tutorials) and explore performance	analyze performance of a planning-specific AI application developed by other researchers with data and explicit metrics	conduct literature review to understand how to adapt existing AI methods to planning-specific problems and data (e.g., pedestrian behaviors tracking in traffic videos); use AI to analyze data/text to understand urban phenomenon (e.g., analyze tweets, Airbnb post, news articles for sentiments; analyze which factor can best predict housing affordability)	build new AI models, systems, and applications tailored to domain-specific challenges and needs, demonstrating careful handling of contextual nuances (e.g., pothole reporter, bike flow prediction, purpose-built zoning data extraction pipeline, itinerary planner app, user dashboard for civic reports)

Appendix Table 3: Simplified Version - AI Tool and Model Application by Critical Level.

This table highlights which critical levels have been observed in the AI reflection journals across all AI tool/model application categories, marked with a check mark. A widespread distribution of check marks means that many instances are observed in a wide range of application behaviors and that students' application behaviors may have evolved throughout the. A category with less check mark could either mean that the definition of the category is narrow by design (e.g., Reflect, Minimize, and Resist AI Use) and thus are less likely to include other levels of use, or the course or students did not explore this type of application in-depth.

AI Tool/Model Application	Level 1	Level 2	Level 3	Level 4	Level 5
Theme: Code Development and Technical Assistance					
Generate, Adapt, and Master Code	✓	✓	✓	✓	
Reason Implementation Strategy	✓	✓	✓	✓	✓
Theme: Understanding, Reasoning and Expression					

AI Tool/Model Application	Level 1	Level 2	Level 3	Level 4	Level 5
Explore Ideas	✓	✓		✓	
Learn New or Difficult Concepts	✓	✓	✓		
Summarize and Structure Content	✓	✓	✓	✓	
Reason, Interpret, and Model with AI	✓	✓	✓		
Assist and Explore Writing	✓	✓	✓	✓	
Theme: Meta-AI					
Experiment and Evaluate AI Behavior and Limits		✓	✓	✓	
Reflect, Minimize, and Resist AI Use			✓		
Theme: Data and Text Analysis					
Supervised Classification and Prediction	✓	✓			✓
Unsupervised Clustering and Pattern Recognition	✓		✓	✓	✓
Explore, Analyze, and Visualize Data		✓	✓	✓	
Process, Standardize, and Analyze Text	✓	✓	✓	✓	✓
Theme: Multimodal Exploration					
Generate and Modify Images	✓	✓	✓	✓	
Create, Analyze, and Leverage Multimodal Media	✓	✓	✓		
Detect and Track Objects in Images or Videos	✓	✓	✓	✓	
Theme: Life Planning and Task Productivity					
Search and Research Information	✓	✓	✓	✓	
Develop Professional Career	✓	✓	✓	✓	
Navigate Life Planning and Support Decision	✓	✓			✓
Organize Tasks and Scheduling	✓	✓	✓	✓	
Theme: Domain-specific Application					
Assist Urban Planning, Research, and Governance	✓	✓	✓	✓	✓

Appendix Table 4: AI Tool and Model Reflection by Theme and Definition.

The definitions are grounded in examples from AI reflection journals in the pilot course.

AI Tool/Model Reflection	Theme	Definition
Describe How AI Adds Value	AI Strength and Weakness	Celebrate clear instances where AI reliably accomplishes specific tasks, such as being correct, timely, or creative; Explain why and where AI contributions matter beyond mere task completion, such as speeding workflows, sparking ideas, guiding debugging, uncovering insights, or improving accessibility.

AI Tool/Model Reflection	Theme	Definition
Describe How AI Falls Short	AI Strength and Weakness	Calls out where AI attempts a task but delivers incomplete, imprecise, or unreliable results; Highlights why AI underperforms, citing lack of depth, relevance, or reliability, and explain the concrete impact on the user's work or decisions (e.g. hallucinations, bias, unpredictability, domain gaps).
Understand AI Mechanism	AI Mechanics	Explain or speculate with curiosity about how the AI tool actually works, describing its core algorithms, data representations, model architectures or pipelines.
Critique Underlying AI Models and Methods	AI Mechanics	Probe the model's design choices, such as its training data, feature extraction, label schemas or decision logic, to explain why it behaves as it does, identify biases or failure modes tied to those design decisions, and imagine technical improvements or alternatives.
Identify Contexts Where AI Excels or Falters	AI Performance and Appropriateness in Context	Pinpoint the situational when and where that make AI tools effective or problematic, considering data quality, prompt clarity, task complexity, UI adaptivity, or domain specificity.
Articulate When to Leverage or Avoid AI	AI Performance and Appropriateness in Context	Guide decisions about whether and when to use AI, balancing benefits against risks, accuracy stakes, domain expertise needs, personalization, or simpler non-AI alternatives.
Examine AI Tool Design Consideration	AI's User, Design, Governance Considerations	Reflect on the architectural and business trade-offs behind a specific AI tool, such as who built it, why, and how choices around APIs, platforms, cost, or accessibility shape its core functionality.
Discuss AI Tool Interface User Experience	AI's User, Design, Governance Considerations	Focus on hands-on interactions with the tool's UI/UX, such as what feels intuitive or confusing, and how input modes (chat, menus, mobile) affect adoption and satisfaction.
Survey AI Policy and Governance	AI's User, Design, Governance Considerations	Examine the institutional frameworks, partnerships, regulations, and transparency gaps that govern AI at scale, such as identifying what formal agreements exist or are lacking and how oversight is documented.
Fine-tune Human-AI Collaboration Practices	Human AI Co-evolution	Describe how users themselves must evolve, such as their prompt engineering, verification routines, task decomposition, and interactive workflows, to partner effectively with AI rather than simply judging its raw performance.
Examine AI's Impact on Personal	Human AI Co-evolution	Explore how sustained AI use reshapes users over time, such as their skills, habits, confidence, and self-concept, affecting

AI Tool/Model Reflection	Theme	Definition
Learning & Growth		memorization, critical-thinking routines, domain expertise, and identity as a learner or professional.
Apply Personal Background to Guide AI Use	Human AI Co-evolution	Show how who you are, such as your discipline, training, or lived experience, shapes what you ask, trust, and build with AI.
Identify Personal and Professional Knowledge Gaps	Human AI Co-evolution	Reveal specific areas of expertise or skill where users realize they still need to learn, often surfaced when AI hits its limits, highlighting new learning needs for effective collaboration.
Explore AI's Emotional Intelligence	AI Social Costs, Risks, Ethics, and Implications	Probe whether AI can genuinely understand or convey human emotions, such as examining how it interprets emotional content or offers support, and the comfort or confusion its simulated empathy provokes.
Discuss Ethical Concerns and Social Risks	AI Social Costs, Risks, Ethics, and Implications	Raise broader moral, social, and governance questions (normative/moral concerns), such as fairness, bias, privacy, labor impacts, accountability, or misuse, focusing on should we? and what if? rather than technical capability.
Consider Social, Environmental, and Economic Implications of AI at Scale	AI Social Costs, Risks, Ethics, and Implications	Explore how widescale AI deployment reshapes societies, economies, and the environment (structural/systemic impacts), covering democratization of work, labor-market shifts, energy/resource demands, equity/bias across communities, and regulatory or infrastructural changes.
Discuss AI Use and Implications in Urban Planning	Domain Application	Focus on AI applications in urban planning, such as GIS segmentation, policy extraction, incident mapping, highlighting domain-specific use cases, limitations, human-in-loop needs, and social or environmental benefits within civic systems.

Appendix Table 5: AI Reflection by Critical Level.

The definitions are grounded in examples from AI reflection journals in the course. Empty cells can mean one of the two things: 1) category definition determines that critical level is unlikely to be observed, 2) no examples are observed in the AI flection journals.

AI Reflection	Level 1	Level 2	Level 3	Level 4	Level 5
Theme: AI Strength and Weakness					

AI Reflection	Level 1	Level 2	Level 3	Level 4	Level 5
Describe How AI Adds Value	Report surface success, including what the tool did, how it performed, and the immediate benefits observed, such as saving time, simplifying tasks for beginners, offering smooth interfaces, generating accurate outputs	Describe personal takeaways about why AI felt useful in practice, such as guiding technical problem-solving, accelerating drafting and brainstorming, boosting productivity in work and personal tasks, enhancing creativity with visual or narrative options, structuring information into summaries or visualizations, and offering reassurance that builds user confidence in decisions.	Explain why AI adds value in a specific context, such as leveraging broad training (e.g., coding, GIS, music), personalizing with context-aware inference, reducing effort in categorization/summarization, accelerating analysis, search, and design, and scaffolding explanations for learning, tutoring, and workflow sequencing.		
Describe How AI Falls Short	Misidentify AI; Describe incidences of failures without further analysis	Report AI's shortcomings, including failures in accuracy (hallucinations, misclassifications, poor citations), weak clarity or style (generic, confusing, awkward phrasing), and task execution (scale, realism, nuance, conventions), and technical/contextual mismatches	Explain AI failures by tracing them to mechanisms, such as training gaps, domain mismatches, structural biases (sequential logic, TF-IDF sparsity), parameter sensitivity (temperature, thresholds), prompt design issues, and representation/terminology limits—that cause hallucinations, oversimplified categories, unrealistic images, mislabels, or missed nuances.		
Theme: AI Mechanics					
Understand AI Mechanism			Explain *how AI works under the hood* by breaking outputs down into their algorithms, math, or model design, such as describing GPT as working like a		

AI Reflection	Level 1	Level 2	Level 3	Level 4	Level 5
			decision tree that narrows options, explaining how embeddings turn text into numbers, comparing how different models process information, pointing out that outputs depend on input format (resolution, features, or text chunks), and asking if AI can handle different modalities like manga or visuals		
Critique Underlying AI Models & Methods			Explain *why AI sometimes fails* by connecting errors to flaws in training data, design, or processing methods, such as pointing out bias (e.g., QA system excluding women, cats mislabeled as people), explaining domain mismatch (street-trained model failing on animals, houses misclassified as buildings), Linking errors to data gaps (English-dominant sources skewing answers about Chinese culture), highlighting structural flaws (TF-IDF missing rare words, chunk sizes cutting summaries short), Noting input variability (color, style, resolution) leading to false positives	Take a clear stance on AI's limits by calling out how human values and biases (political, gender, cultural) are built into training data. Question what counts as creativity versus pattern-copying, use outside evidence (like UNESCO/IRCAI reports) to show systemic bias, and emphasize why critical awareness and education are needed.	
Theme: AI Performance and Appropriateness in Context					
Identify Contexts Where AI		Identify use scenarios that AI excels or falters from personal,	Show *conditions and contexts* that shape when AI succeeds or fails such as data	*Evaluate usefulness across contexts* and takes a stance on when AI is better as	

AI Reflection	Level 1	Level 2	Level 3	Level 4	Level 5
Excels or Falters		simple interactions without further articulation, such as AI failed at searching information about a job as compared with Google	quality and structure, prompting quality and contextual fit, algorithmic and parameter setting, and task and domain-specific contexts. Examples include noticing AI detects objects well in high-resolution video but struggles in low-resolution footage, explaining how zoning rules hidden in tables or "special use" clauses confuse the model, pointing out cultural mismatches, like restaurant classification depending on word associations, describing parameter effects, such as confidence thresholds leading to over- or under-detection, highlighting task/domain differences, like sentiment models mislabeling neutral statements or policy summaries missing equity details.	support, inspiration, or efficiency tool, rather than a replacement, such as arguing AI is helpful for scanning large zoning documents but can't replace planners' judgment, comparing AI's predictions with real outcomes (e.g., bike flow demand, trip itineraries), judging visualization outputs as incomplete or unrealistic compared to human-made versions, concluding image tools are better for inspiration than realism, or one design tool suits concept work while another suits art, emphasizing AI speeds up data processing but requires human interpretation and oversight	
Articulate When to Leverage or Avoid AI				Makes *value judgments about when to use or not use AI* and sets clear boundaries for responsible application, such as leveraging AI for brainstorming, structure, or filtering messy data, but not as the final "truth", avoiding AI for high-stakes or detail-heavy work (e.g.,	Propose new ways of deciding when to use AI by rethinking existing practices and suggesting better alternatives, such as suggesting that AI should be evaluated through side-by-side comparisons (e.g., Chatbot Arena) to align with human values, not just abstract metrics, critiquing an AI-

AI Reflection	Level 1	Level 2	Level 3	Level 4	Level 5
				zoning laws, precise image generation), using AI when fact-checking and validation are possible, avoiding it when results can't be verified, choosing AI only when it improves workflows; skipping it when simpler tools (like search or SMS) are more efficient, trusting AI to support explanation or empathy (AI personas), but keeping humans at the center for accuracy and judgment.	powered pothole-reporting chatbot for offering little improvement over existing online forms, and instead proposing an SMS system to expand accessibility.
Theme: AI's User, Design, Governance Considerations					
Discuss AI Tool Interface User Experience		Share personal impressions of how easy or confusing an AI tool is to use, focusing on its design, accessibility, or first-time experience, such as noting whether a tool requires an API key, subscription, or mobile app preference, reacting to features (e.g., AI that adjusts tone from writing samples or uses natural conversation), describing confusion with unclear UI on first try, reflecting on how AI is marketed as a "magic button" while hiding functions—especially for elderly or less tech-savvy users.	Speculate AI system design from the tool interaction, such as guessing that the model design of a campus housing bot is a rule-based system rather than an LLM-driven system.	Evaluate how well an AI tool's interface supports real use by weighing trade-offs, comparing its recommendations to human knowledge, and judging when the design helps or hinders effective work, such as comparing campus AI tool's recommendations against personal expertise or campus resources, weighing trade-offs (e.g., advanced models give better answers but force longer wait times), deciding when to split tasks across tools/models for efficiency, Assessing whether the interface helps allocate effort wisely or slows progress.	Suggest new applications to improve user experience, such as adding ML and NLP to O*NET Interest Profiler to optimize career interest analysis and personalize career advice.

AI Reflection	Level 1	Level 2	Level 3	Level 4	Level 5
Examine AI Tool Design Consideration				Evaluate how design choices shape AI tools by comparing different approaches and reflecting on trade-offs between flexibility, accessibility, cost, and usability, such as contrasting two mind-mapping tools: a startup tool that breaks problems into modular steps vs. an established platform that enhances usability without losing core functions, weighing chatbot design trade-offs: conversational flexibility with ChatGPT vs. low-cost accessibility of SMS forms, comparing web-based vs. mobile-centered designs, noting how technical barriers can limit adoption even if the design is more user-friendly	Proposes concrete redesigns or scaling strategies that re-center AI tools around accessibility, context, and human use, such as questioning whether a student-built prototype could scale for cities and proposing advanced features (e.g., NLP to filter high-volume reports), reframing an AI project from a coding exercise into a human-centered design opportunity, suggesting integration with everyday tools like phones or cars, envisioning a civic chatbot evolving into a mobile app for safer, more accessible infrastructure reporting, shifting focus from technical fixes to people-centered design.
Survey AI Policy and Governance			Reflect on why AI policy documents are difficult to find and tend to be outdated		
Theme: Human-AI Co-evolution					
Fine-tune Human-AI Collaboration Practices		Describes surface-level tweaks to improve collaboration, without going deep into why it worked, such as mentioning that rewording a prompt, adding detail, or using a clear tone made outputs better, saying that constraints or context (like word	Reflect on how to change model design to achieve better AI outputs, such as experimenting with system messages or text-splitting strategy	Evaluate how to balance human judgment with AI outputs by showing when and how people must step in to guide, correct, and validate results, such as double-checking summaries or transcripts to confirm key details aren't missing, simplifying AI's overly complex	Strategic redesign of the workflow and actively shape how AI fits into process while ensuring humans remain essential for accuracy and context, such as suggesting a better image-generation workflow (e.g., sketch first, then let AI fill in details instead of text-only prompts), noticing when AI searches

AI Reflection	Level 1	Level 2	Level 3	Level 4	Level 5
		count or style) improved results, noting that AI worked best when paired with your own judgment or prior knowledge.		answers or correcting errors in context, using AI as a starting point for searches, then validating with trusted sources, comparing outputs from different models or devices to cross-check accuracy, reflecting on time saved versus time spent re-judging AI's results, describing how clear goals and specific prompts make collaboration more effective, noting when AI brainstorming boosted creativity (e.g., for zine ideas) but needed human direction, choosing the right model for the right task and ensuring results meet design or accessibility standards.	return nothing and inventing new strategies (e.g., expanding geography, refining keywords) to guide the process, iterating text extraction on zoning codes while checking results against human judgment to spot gaps or errors, describing how changes in collaboration method—like sequencing steps differently—lead to better outputs.
Examine AI's Impact on Personal Learning & Growth		Show a growing awareness of how AI fits into one's learning process, balancing excitement with caution about overuse, such as: Emphasizing that humans should double-check AI outputs to ensure accuracy and alignment with personal learning goals; Expressing worry that depending too much on AI could weaken memory, critical thinking, or trial-and-error learning; Noticing		Highlight deliberate choices about when AI supports growth and when personal effort matters more, showing a thoughtful balance between efficiency and independence, such as: Acknowledging the trade-off between speed and long-term retention when relying on AI for studying; Choosing to troubleshoot code personally or review lab materials without AI to strengthen problem-solving, practicing with AI as a mock	Redefine what learning itself means with AI, reframing tasks, perspectives, and practices through AI use, such as: Appreciating the agency to use AI as a method, not just consume its outputs—solving real challenges with it rather than passively accepting answers; Shifting from initial skepticism about AI's costs to recognizing its value as a planning and problem-solving partner through structured coursework; Realizing that learning is now

AI Reflection	Level 1	Level 2	Level 3	Level 4	Level 5
		how practice with prompting or trying diverse use cases has improved personal confidence in working with AI; Noting patterns or gaps in personal use of AI, such as how to cook or how to take medicine.		interviewer, gaining confidence while noting it can't fully replace human feedback; Expressing healthy skepticism about AI's accuracy in tasks like citations, even while using it for convenience, recognizing that AI planning leaves out human factors, like learning style or workload preferences; Building confidence in prompt writing and coding skills through repeated practice with AI; Wondering whether multimodal AI tools might reduce the need for careful prompt engineering in the future.	partly about learning how to use AI effectively to reach goals, not only doing tasks by hand; Describing how using AI/data analysis methods transformed perspectives—such as seeing Airbnb house rules not as minor details, but as a system shaping tone, authority, and social expectations.
Apply Personal Background to Guide AI Use				Bring one's own expertise into the process, ensuring AI's use is grounded in disciplinary knowledge and lived experience, such as: Applying an urban planning/tech perspective to balance AI's technical accuracy with broader social or economic concerns; Framing policy conflicts (like governance opposition) in ways that reflect community voices and local autonomy, shaping which themes AI should detect in text analysis; Anticipating everyday resident	Redesign AI tools and workflows so they align with professional knowledge, community needs, and broader social values, such as: Designing civic reporting tools (e.g., SMS-based systems) that reduce barriers between residents and government, grounded in knowledge of bureaucracy and accessibility; Guiding NLP models to focus on meaningful planning rules, correcting extraction errors by applying professional expertise in regulations; Interpreting technical outputs (like object-detection data) not

AI Reflection	Level 1	Level 2	Level 3	Level 4	Level 5
				questions (e.g., zoning issues) so AI outputs are relevant to real-world contexts.	just as numbers, but in terms of the built environment and human behavior; Evaluating who benefits or is excluded, ensuring AI serves communities responsibly rather than just optimizing efficiency; Reframing AI tools as human-centered systems, such as turning a coding prototype into a mobile app that improves safety and accessibility for residents; Using disciplinary expertise to ask bigger questions about fairness, trust, and social impact (e.g., "Who is this tool really for?").
Identify Personal & Professional Knowledge Gaps		Spot and clarify knowledge gaps, motivating further learning and skill-building, such as: Asking AI to explain quantitative analysis concepts in urban planning and realizing one does not fully understand the logic behind them; Using AI to debug code, but seeing that deeper technical knowledge is still needed to solve complex problems; Critiquing dense documentation as overwhelming, while leaning on AI to simplify it for beginners; Struggling to interpret outputs from a Hugging	Use AI responses to figure out what key skills, concepts, or metrics one is missing, and connect those gaps to what's essential for learning in one's field, such as: Deriving essential metrics for supervised ML models (e.g., accuracy, precision, recall, F1) by observing patterns in AI's explanations		

AI Reflection	Level 1	Level 2	Level 3	Level 4	Level 5
		Face model, showing the need for stronger foundational skills; Experimenting with AI website design, which sparks motivation to learn more about web development; Recognizing that AI can't fill in gaps you don't know yourself—it can guide, but not replace your own learning			
Theme: AI Social Costs, Risks, Ethics, and Implications					
Explore AI's Emotional Intelligence		Share personal takeaways about AI's ability (or inability) to handle emotion, such as: Pointing out that AI-generated videos struggled with realistic facial expressions; Appreciating AI's role as emotional support during debugging or interview prep; Suggesting that a customized chatbot could serve as a virtual partner for encouragement or motivation.	Explain why AI responses feel human by linking it to model design, training, or persona choices, such as: Comparing ChatGPT and DeepSeek for comfort requests, and noticing that ChatGPT felt more "human" because of its explicit training for supportive dialogue; Recognizing that AI personas (tone, character, or role-play settings) strongly shape emotional impact and user acceptance.	Evaluate how AI's lack of empathy or contextual awareness can distort sensitive situations, especially when others rely too heavily on its outputs. Judge whether AI is trustworthy in handling human emotions or interpersonal dynamics, such as: Describing how an AI meeting summary oversimplified classmates' concerns, stripping away nuance and sensitivity and that AI's tone-deaf summaries could misrepresent group dynamics and damage trust if used uncritically.	
Discuss Ethical Concerns & Social Risks	Surface observation of others' discussions on ethical concerns and social risks, such as through a professional planning conference			Evaluate the ethical risks of AI and judge whether current safeguards are enough and highlight where accountability is missing, such as: Raising concerns	Develop application and analysis that address AI's ethical concerns and social risks, such as building a fake news detector and suggesting ways to improve AI

AI Reflection	Level 1	Level 2	Level 3	Level 4	Level 5
				about bias (e.g., face-fusion apps, hiring systems, or cultural/gender bias in LLMs across countries); Questioning regulatory gaps, such as weaker U.S. protections compared to the EU, or accountability in AI-assisted government work; Highlighting risks to autonomy and trust, like AI personas manipulating emotions, or people over-trusting AI-generated notes; Discussing cultural and creative risks, such as copyright, consent, or fears of AI replacing artists; Noting social harms, from beauty filters harming teens to risks in political decision-making; Pointing to responsible practices, such as frameworks that mitigate bias, as examples of positive steps.	recognition of fake news
Consider Social, Environmental & Economic Implications		Share personal takeaways about large-scale implications without further analysis, such as Imagining daily life impacts, like voice-enabled AI making things easier for elderly users; Noting job concerns, such as AI automating entry-level work and threatening artists' livelihoods;		Examine the bigger trade-offs of AI adoption, weighing social, environmental, and economic risks against its benefits, such as: Questioning how AI hiring systems create stress for students navigating black-box filters; Reflecting on the paradox of progress, where advanced tech and productivity still	Identify missing links between AI policy and energy policy as knowledge gaps and consider AI's infrastructural and energy requirements

AI Reflection	Level 1	Level 2	Level 3	Level 4	Level 5
		Sharing curiosity about tech integration, like AI features in new hardware; Observing societal adoption, such as municipalities starting to expect AI use; Recognizing that coding barriers are lowering, making programming accessible beyond software engineers.		coexist with job insecurity; Raising concerns that AI-driven feeds may trap people in narrow information bubbles; Highlighting sustainability issues, such as the high energy costs of large AI systems.	
Theme: Domain-specific AI Applications and Extensions					
Discuss AI Use & Implications in Urban Planning	Report AI applications without interpretation, such as commenting that the state now uses AI-driven wetland monitoring through image segmentation of raster data.	Share personal takeaways on how AI changed the experience of doing planning tasks, such as: Noting how AI makes policy comparisons across cities more consistent; Highlighting when AI helps spot hidden urban patterns; Using AI-generated images to explain or advocate for planning issues, like bike-friendly streets; Finding value in AI for classifying urban data (e.g., waste types) more effectively.	Connect AI's success or failure in planning to data quality, context, or design limits, such as Pointing out that zoning documents hide solar farms under "special use", so AI misses key rules; Explaining how object detection fails when video has heavy traffic, poor weather, or low resolution; Noting that AI summaries of planning policies can overlook equity details because the training data skews toward technical language; Observing that classification results depend on cultural word associations, such as restaurant names tied to cuisine type.	Evaluate trade-offs and risks of using AI in planning, balancing efficiency gains with fairness, accuracy, and the need for human expertise, such as: Questioning AI's role in political decision-making, asking whether it should ever guide zoning or housing policies; Warning bias in training data—especially the lack of socio-cultural and planning texts—can distort outputs; Weighing AI object detection against manual observation, debating which better captures public space vitality; Comparing AI-driven civic reporters with SMS systems, considering inclusivity, government moderation, and city scale; Stressing safeguards like planner oversight, accuracy checks, and bias mitigation	Propose new ways AI could reshape planning practice that make human–AI collaboration more meaningful and people-centered, such as: Proposing integrations, like combining ChatGPT with ArcGIS to automate GIS tasks or adding NLP to detect place names for civic reporting; Designing new features, such as a confidence score for image segmentation or a more accessible chatbot interface; Spotting gaps in current research, noticing when methods focus only on technical optimization and suggesting ways to address human behavior or planning concerns; Identifying limits and offering fixes, like explaining why zoning text extraction fails and proposing advanced NLP/ML methods instead; Envisioning future scenarios,

AI Reflection	Level 1	Level 2	Level 3	Level 4	Level 5
				before trusting zoning extractions, meeting summaries, or rapid-response systems; Speculating on broader impacts, like how AI could reshape remote work, housing choices, or enable shared urban data platforms.	where AI civic reporters move beyond automation to support richer, two-way reporting between government and residents; Re-centering on people, stressing that AI should act as a collaborator that enhances planning expertise rather than replacing it.

Appendix Table 6: Simplified Version - AI Reflection by Critical Level.

This table highlights which critical levels have been observed in the AI reflection journals across all AI reflection application categories, marked with a check mark. A widespread distribution of check marks means that many instances are observed in a wide range of reflections. A category with less check mark could either mean that the definition of the category is narrow by design (e.g., Understand AI Mechanism) and thus are less likely to include other levels of use, or the course or students did not explore this type of application in-depth (e.g., Survey AI Policy and Governance).

AI Reflection	Level 1	Level 2	Level 3	Level 4	Level 5
Theme: AI Strength and Weakness					
Describe How AI Adds Value	✓	✓	✓		
Describe How AI Falls Short	✓	✓	✓		
Theme: AI Mechanics					
Understand AI Mechanism			✓		
Critique Underlying AI Models and Methods			✓	✓	
Theme: AI Performance and Appropriateness in Context					
Identify Contexts Where AI Excels or Falters		✓	✓	✓	✓
Articulate When to Leverage or Avoid AI				✓	✓
Theme: AI's User, Design, Governance Considerations					
Discuss AI Tool Interface User Experience		✓	✓	✓	✓
Examine AI Tool Design Consideration				✓	✓
Survey AI Policy and Governance			✓		
Theme: Human AI Co-evolution					
Fine-tune Human-AI Collaboration Practices		✓	✓	✓	✓
Identify Personal and Professional Knowledge Gaps		✓	✓		
Examine AI's Impact on Personal Learning and Growth		✓		✓	✓
Apply Personal Background to Guide AI Use				✓	✓

AI Reflection	Level 1	Level 2	Level 3	Level 4	Level 5
Theme: AI Social Costs, Risks, Ethics, and Implications					
Explore AI's Emotional Intelligence		✓	✓	✓	
Discuss Ethical Concerns and Social Risks	✓			✓	✓
Consider AI’s Social, Environmental, and Economic Implications at Scale		✓		✓	✓
Theme: Domain-specific AI Applications and Extensions					
Discuss AI Use and Implications in Urban Planning	✓	✓	✓	✓	✓