

# Systematic AI-Human Collaboration for Worldbuilding in Cybersecurity Education: A Practice-Based Framework

Ryan Straight  
College of Information Science  
University of Arizona  
Tucson, AZ 85721 USA

**Abstract:** This paper documents a systematic framework for AI-human collaboration in educational worldbuilding, the practice of constructing coherent fictional universes that provide persistent narrative environments for pedagogical case studies. Developed through implementing seven interconnected cases in undergraduate cybersecurity policy education, our methodology comprises four integrated components: structured prompt engineering for scenario generation, YAML-based entity management for cross-case consistency, iterative human validation for pedagogical alignment, and reproducible technical infrastructure. Whereas traditional case studies typically present bounded, isolated scenarios, our approach treats AI as constrained collaborative partner operating within explicit pedagogical boundaries to generate persistent narrative environments. Implementation across three semesters with undergraduate and graduate students (n=110) suggested distinctive engagement patterns: spontaneous recognition of recurring entities, cross-scenario analytical synthesis, and more nuanced ethical reasoning than observed with traditional isolated cases. Iterative co-creation with AI, supported by systematic quality control, maintained fictional universe integrity and representational appropriateness throughout development. We discuss transferability beyond cybersecurity contexts and offer a reproducible methodology educators may adapt to their own educational settings.

## Introduction

Cybersecurity ethics and policy education operates within a fundamental tension: the field demands that students develop nuanced understanding of complex socio-technical systems, understanding that emerges through sustained engagement with the messy particularity of actual contexts rather than from abstract principles alone. Yet the very particularity that makes case studies pedagogically valuable often fragments student experience. A data breach scenario illuminates privacy considerations; a social engineering case reveals human vulnerability; a content moderation dilemma exposes platform governance challenges. Each case stands in isolation, a discrete pedagogical moment disconnected from broader narrative continuity.

Such fragmentation carries pedagogical consequences beyond mere aesthetics. When students encounter organizations, individuals, and technological artifacts that appear once and vanish, they miss opportunities to observe how ethical challenges evolve across contexts, how organizations respond to cumulative pressures, and how individual actors navigate shifting terrain. Each case study becomes a snapshot rather than a chapter: pedagogically useful but lacking the narrative momentum that would allow students to trace consequences across time and situation.

As a response to this fragmentation, worldbuilding offers the systematic construction of coherent fictional universes that provide persistent narrative environments for pedagogical scenarios (Breuleux et al., 2019; Wolf, 2012b). Borrowed from speculative fiction and game design, this practice provides a framework for creating interconnected cases that share characters, organizations, technologies, and histories. A journalist who reports on a data breach in one scenario might later investigate algorithmic bias; a tech company facing privacy violations in one case confronts content moderation challenges in another; advocacy organizations persist across multiple scenarios, developing positions and strategic responses.

Yet worldbuilding at scale demands resources that may exceed typical educator bandwidth. Constructing a coherent fictional universe with hundreds of entities, including people, organizations, technologies, events, and documents, while maintaining consistency across multiple pedagogical scenarios represents time-intensive authorial labor. The emergence of capable large language models (LLMs) offers potential relief, but raises immediate

questions: How might educators collaborate systematically with AI to construct educationally appropriate narrative environments? What mechanisms ensure consistency when AI generates content across multiple scenarios? How can human validation preserve pedagogical alignment while leveraging AI's generative capacity?

This paper documents a practical framework developed through constructing seven interconnected worldbuilding cases for undergraduate cybersecurity law, ethics, and policy education at a US research university. We offer methodological documentation rather than theoretical claims about AI in education or narrative pedagogy: a reproducible approach to AI-human collaboration for worldbuilding within the post-authorship condition that other educators might adapt to their contexts.

Our framework comprises four integrated components: structured prompt engineering that guides AI scenario generation toward pedagogical goals; YAML-based entity management that tracks people, organizations, technologies, and events across cases; iterative human validation that ensures pedagogical alignment and ethical appropriateness; and dynamic content integration through reproducible technical infrastructure that maintains consistency at scale. We position AI not as autonomous content generator but as collaborative partner. This relationship reflects what Veletsianos et al. (2024) characterize as treating AI simultaneously as object (tool serving pedagogical ends) and subject (agentic participant in creative process).

Having established the pedagogical problem and our methodological response, the sections that follow document this framework's development and classroom implementation in sufficient detail for adaptation. We begin by situating worldbuilding within broader conversations about narrative pedagogy, immersive learning, and AI-human collaboration in educational content creation. We then describe our methodological approach so that others might implement similar systems. Concrete examples in the implementation section demonstrate how narrative elements persist and evolve across scenarios. We present findings from classroom integration, acknowledging the limitations of practice-based documentation while identifying patterns that warrant further investigation. Finally, we discuss transferability beyond cybersecurity education and implications for educators navigating the opportunities and constraints of AI-assisted content creation.

We make no claim that worldbuilding represents a pedagogical panacea or that AI collaboration resolves all challenges in case study development. Rather, we document what became possible when systematic methodology met emergent technology, and we invite others to adapt, critique, and extend these practices in their own educational contexts.

## Literature Review

Our approach intersects worldbuilding scholarship, case-based pedagogy, and emerging AI-human collaboration practices. We position our methodological contribution within these conversations rather than claiming theoretical innovation.

### Worldbuilding and Case-Based Learning

Worldbuilding, understood as the systematic construction of coherent fictional universes, has received limited attention in educational technology despite widespread application in adjacent fields. Wolf (2012a) provides foundational analysis of worldbuilding mechanics in speculative fiction, while Breuleux et al. (2019) propose a worldbuilding framework for design education, and Friedrichsen (2021) extends these principles to game-based learning environments. What distinguishes AI-assisted worldbuilding from conventional case development is its operation within what Abdul-Jabbar & Bhatt (2025) term the *post-authorship condition*: a collaborative space where human educators and AI systems jointly originate content, with neither serving as exclusive author. Traditional educational cases present bounded scenarios, whereas worldbuilding establishes persistent contexts from which educators generate numerous related scenarios.

Case studies effectively develop critical thinking (Mahdi et al., 2020), and case-based approaches have proven particularly valuable for ethics education in technical fields including cybersecurity (Blanken-Webb et al., 2018). Yet traditional approaches risk what we term *pedagogical atomism*: treating each case as an isolated learning event disconnected from narrative continuity. Students miss opportunities to observe ethical challenges as ongoing, to see organizations navigate cumulative pressures, and to understand policy debates as historically situated.

From a learning theory perspective, worldbuilding’s pedagogical value aligns with situated learning frameworks (Lave & Wenger, 1991), which emphasize that knowledge develops through sustained engagement with authentic communities and contexts rather than through decontextualized instruction. Continuous narrative settings create conditions for situated learning by embedding ethical reasoning within ongoing relationships, institutional histories, and accumulated consequences, enabling students to develop professional judgment through immersion in complex, evolving contexts. Worldbuilding addresses traditional case study limitations by creating persistent narrative environments where the same organizations, characters, and regulatory bodies appear across multiple scenarios.

## **AI Collaboration in Educational Content**

Kasneci et al. (2023) and Farrokhnia et al. (2023) identify that effective AI integration requires clear pedagogical strategies and systematic quality control. Williamson (2024) further argues that AI in education operates within “political and economic contexts” rather than as neutral technology—a perspective that informs our positioning of AI as constrained collaborative partner rather than autonomous generator. Central to our approach is Veletsianos et al. (2024)’s framing of AI relationships: people relate to AI sometimes as object (tool serving predetermined ends), sometimes as subject (agentic partner), and often both simultaneously. Such dual positioning characterizes our methodology, wherein structured prompts direct AI toward pedagogical goals while we remain responsive to unexpected narrative possibilities. Recent prompt engineering scholarship (Giray, 2023; White et al., 2023) situates our systematic approach within emerging best practices.

## **Gap and Contribution**

Existing scholarship addresses components of our approach in isolation: worldbuilding in creative contexts, immersive learning environments (Kampe, 2022), and AI’s educational potential. Researchers have explored AI-generated educational content, automated assessment, and chatbot tutoring systems. However, we lack documented methodologies for *systematic AI-human collaboration in educational worldbuilding* that integrate these elements into a coherent framework. Our contribution addresses this gap: a reproducible framework for leveraging LLM capabilities while maintaining pedagogical control, developed through actual classroom deployment and building upon these prior explorations of AI in education.

## **Methodological Framework**

Our framework for AI-assisted worldbuilding emerged iteratively during textbook development, not designed comprehensively at the outset but rather discovered through practice. We present it here as systematic methodology others might adapt. The framework comprises four integrated components designed to support scalable, consistent worldbuilding across multiple scenarios and semesters: structured prompt engineering, YAML-based entity management, iterative human validation, and reproducible technical infrastructure.

Throughout development, we employed multiple large language models rather than relying on a single platform, using Google’s Gemini for initial scenario ideation and narrative variation and Anthropic’s Claude for refining pedagogical framing and maintaining consistency with established entities. This multi-model approach reflected practical considerations, as different models exhibited varying capabilities for specific tasks, rather than systematic comparison. We deliberately avoided exclusive reliance on any single model, recognizing that the AI landscape evolves rapidly and that methodology transferability requires platform flexibility.

## **Worldbuilding Architecture: Entity Management Through YAML**

At the foundation lies a structured data representation of the constructed world. We employ YAML (YAML Ain’t Markup Language), a human-readable data serialization format widely used in software development, to document entities across six categories: people (individuals who appear across cases, including journalists, executives, advocates, regulators, and victims); organizations (companies, advocacy groups, regulatory bodies, and research institutions); locations (cities, facilities, and neighborhoods that host narrative events); events (conferences, policy debates, and incidents that create shared history); documents (reports, legislation, and manifestos that fictional actors reference); and technology (platforms, systems, and algorithms that drive ethical dilemmas).

Each entity includes structured metadata. A person entity, for instance, specifies name, role/occupation, organizational affiliation, biographical details, and narrative purpose. An organization entity documents mission, history, key personnel, and previous case appearances. Multiple functions emerge from this structure: it creates a searchable reference during new case development; it ensures consistency when entities reappear; it reveals narrative opportunities through relationship mapping.

Consider a simplified example:

```
people:
- name: "Jordan Pham"
  name_first: "Jordan"
  name_last: "Pham"
  salutation: "Mr."
  role: protagonist
  position: "Investigative Technology Journalist"
  affiliation: "The Silicon Valley Observer"
  bio: "Known for data privacy reporting, particularly corporate surveillance"
  appears_in: ["case_03_data_breach", "case_05_algorithmic_bias"]
  narrative_arc: "Evolves from privacy advocate to algorithmic accountability critic"

organizations:
- name: "Digital Rights Now"
  type: "Advocacy Organization"
  mission: "Privacy protection and algorithmic transparency"
  founded: "2019"
  key_personnel: ["Elena Rodriguez - Executive Director"]
  appears_in: ["case_02_privacy", "case_04_content_moderation", "case_07_ai_ethics"]
  narrative_trajectory: "Expands scope from privacy to broader AI ethics concerns"
```

Structured thus, YAML renders worldbuilding infrastructure visible and maintainable. As cases accumulate, the file grows, documenting not just individual entities but the evolving narrative ecosystem they inhabit. Relationships emerge through use: Jordan Pham quotes Elena Rodriguez in multiple cases, Digital Rights Now submits testimony citing Pham's reporting, and fictional events create history that subsequent cases reference.

Why YAML rather than a database or spreadsheet? Unlike proprietary database systems, YAML files are plain text, making them readable, versionable through Git, editable in any text editor, and processable through standard programming libraries. Such accessibility reduces technical barriers for educators adapting the approach.

## Structured Prompt Engineering: Guiding AI Generation

Creating instructionally aligned scenarios through AI collaboration requires systematic prompting strategies. We distinguish three prompting phases, each serving specific methodological functions. The first phase, scenario scaffolding, establishes pedagogical parameters before requesting narrative content. A typical scaffolding prompt specifies the learning objectives the scenario must address, the ethical and legal concepts to illuminate, the complexity level appropriate to the student audience, relevant constraints such as avoiding certain stereotypes or maintaining cultural sensitivity, and the desired scenario structure including case facts, decision points, and stakeholder perspectives.

Consider the following example scaffolding prompt:

*“Generate a cybersecurity case study for undergraduate students that explores tensions between user privacy and platform content moderation. The scenario should involve a social media platform facing pressure from multiple stakeholders (users, advertisers, regulators). Include realistic technical details about content recommendation algorithms without requiring advanced CS knowledge. The case should raise questions about privacy in behavioral data collection, transparency in automated decision-making, and competing definitions of harm. Avoid framing any stakeholder group as simple villain or hero.”*

This prompt positions AI as collaborator working toward pedagogical specifications rather than autonomous content generator, reflecting prompt engineering best practices’ emphasis on clear task decomposition and constraint specification (White et al., 2023).

The second phase, worldbuilding integration, ensures consistency with the existing narrative ecosystem once AI generates initial scenario content. We provide AI with relevant YAML excerpts (entities that might appear in the new case) and request narrative incorporation:

*“Revise this scenario to include Jordan Pham (investigative journalist, Silicon Valley Observer) discovering and reporting on the content moderation controversy. Reference Digital Rights Now as advocacy organization filing formal complaints. Ensure Jordan’s reporting style matches previous appearances (data-driven, critical of corporate tech power, seeks technical expert sources).”*

This phase treats the YAML file as canonical reference—the “ground truth” about the constructed world—that AI-generated content must respect. It positions human curation as authoritative while leveraging AI’s generative capacity for narrative elaboration.

The third phase, pedagogical refinement, adjusts tone, complexity, and pedagogical framing. These final prompts might request simplification of technical jargon without sacrificing accuracy, addition of specific decision points for student analysis, inclusion of diverse stakeholder perspectives, or adjustment of narrative pacing to maintain engagement.

Multiple iterations often prove necessary. AI might generate compelling narrative that fails pedagogical objectives; subsequent prompts redirect toward learning goals. This iterative process of generation, evaluation, and refinement reflects the AI-as-collaborative-partner relationship Veletsianos et al. (2024) describe.

## **Human Validation: Ensuring Pedagogical and Ethical Alignment**

We address concerns about AI-generated educational content through systematic validation processes, employing four criteria applied to all AI-generated scenarios before classroom deployment. Pedagogical validity demands that each scenario illuminate targeted concepts with complexity and technical detail appropriate to the intended audience; embedded decision points must support desired learning objectives in ways that advance student understanding. Scenarios failing to serve clear pedagogical purposes undergo revision or rejection regardless of narrative quality. Fictional universe integrity presents a distinct challenge: while scenarios reference real-world technologies, policies, and regulatory frameworks as backdrop, they must avoid attributing actions or positions to actual people or organizations, since large language models trained on real-world data sometimes insert references to real companies, executives, or public figures when generating narrative content.

Our validation process identifies these boundary violations before students encounter them. Ethical appropriateness requires that educational scenarios avoid harmful stereotypes while representing diverse perspectives authentically, responding to An et al. (2024)’s finding that AI ethics frameworks in education frequently neglect user perspectives. We deliberately constructed the scenario setting to include diverse representation across gender, ethnicity, and professional background, rejecting content that reduces complex stakeholders to caricatures or presents simplistic moral binaries. Narrative consistency maintains the integrity of the worldbuilding environment across the full span of case materials, with the YAML file serving as the canonical consistency standard; when AI-generated content contradicts documented entities, we pursue revision through additional prompts or manual editing rather than accepting narrative inconsistencies.

In practice, validation involved systematic instructor review against documented criteria before any scenario reached students. Pedagogical validity assessment used a rubric checking whether scenarios addressed stated learning objectives, included appropriate decision points, and matched target complexity levels. Fictional universe integrity verification checked for inadvertent references to real companies, public figures, or actual events that should remain outside the fictional universe, while confirming accurate representation of real-world regulatory frameworks and technical concepts serving as backdrop. Ethical appropriateness review examined stakeholder representation for stereotyping, power dynamic framing, and implicit assumptions about technology's social role. Narrative consistency checking compared new scenarios against the YAML documentation, flagging contradictions in character behavior, organizational positions, or timeline sequences.

Through iterative implementation, we observed that treating AI as collaborative partner from the outset produced better results than sequential generate-then-review workflows. Our process began with brainstorming scenarios alongside AI based on learning objectives, selecting the most pedagogically promising directions, building outlines collaboratively, fleshing out narratives together, and polishing through successive refinement. This iterative co-creation meant that validation occurred continuously throughout development rather than as a discrete post-generation phase; the quality of final scenarios reflected accumulated human judgment across multiple collaborative exchanges. We developed a custom AI agent specifically tasked with worldbuilding consistency checking, parsing YAML entity files to identify continuity violations: characters appearing with inconsistent biographical details, organizations referenced before their documented founding dates, or events described with contradictory timelines. This computational approach to consistency verification enabled systematic quality control at scale, flagging potential contradictions for human review before they propagated across interconnected cases.

This validation framework recognizes AI's limitations while leveraging its strengths. AI excels at generating narrative variety within constraints; humans excel at evaluating pedagogical appropriateness and ethical nuance. The division of labor proves productive.

### **Technical Infrastructure: Reproducible Integration**

Worldbuilding infrastructure must support content creation without creating maintenance burden. We employ Quarto, an open-source scientific publishing system, integrated with R for data processing. This technical stack enables several critical capabilities that make large-scale worldbuilding tractable.

Dynamic content rendering transforms entity management from manual duplication to automated consistency. Case studies incorporate entity data directly from YAML files, with character descriptions, organizational details, and location information rendering dynamically throughout the narrative. When entity information updates in the YAML documentation (perhaps a character receives a promotion or an organization shifts strategic focus) all cases referencing that entity automatically reflect these changes. This automation prevents the inconsistencies that plague manual content management while reducing maintenance burden as the narrative ecosystem expands.

Consistency checking through automated scripts scales more effectively than manual cross-referencing as worldbuilding complexity grows. R scripts parse YAML files to flag potential inconsistencies, including characters appearing in undocumented cases, organizations with conflicting descriptions across scenarios, and timeline contradictions where events reference impossible sequences. This computational approach to narrative consistency transforms an overwhelming cognitive challenge into a manageable technical process. Educators receive alerts about potential problems before they reach students, maintaining the narrative coherence essential to worldbuilding's pedagogical value.

Version control through Git provides essential infrastructure for tracking worldbuilding evolution and enabling collaborative development. All changes to YAML files and case content create audit trails documenting how the constructed world develops over time. When experimental narrative changes prove problematic, version control enables rollback to previous states. Multiple educators can contribute to the same worldbuilding environment without coordination challenges, as version control manages concurrent development and highlights conflicts requiring resolution. This infrastructure proves particularly valuable when worldbuilding spans multiple courses or institutional collaborations.

Cross-platform accessibility ensures that worldbuilding-enriched content deploys consistently across diverse educational contexts. Quarto renders content to multiple formats from a single source, including PDF for printed materials, HTML for web delivery, ePub for e-readers, and Word for institutional requirements. This format flexibility means worldbuilding cases work equally well in traditional textbooks, learning management systems, assignment handouts, or interactive web resources. Educators need not choose between narrative richness and technical compatibility.

The technical infrastructure choice reflects methodological commitments: open-source tools reduce financial barriers, plain-text formats ensure longevity, and scripted processes enable reproducibility. Educators adapting this framework need not adopt identical tools, as the principles transfer to alternative technical implementations.

### **Framework Synthesis: Workflow Integration**

Figure 1 illustrates the complete workflow, showing how the four components integrate into a coherent process for AI-assisted worldbuilding.

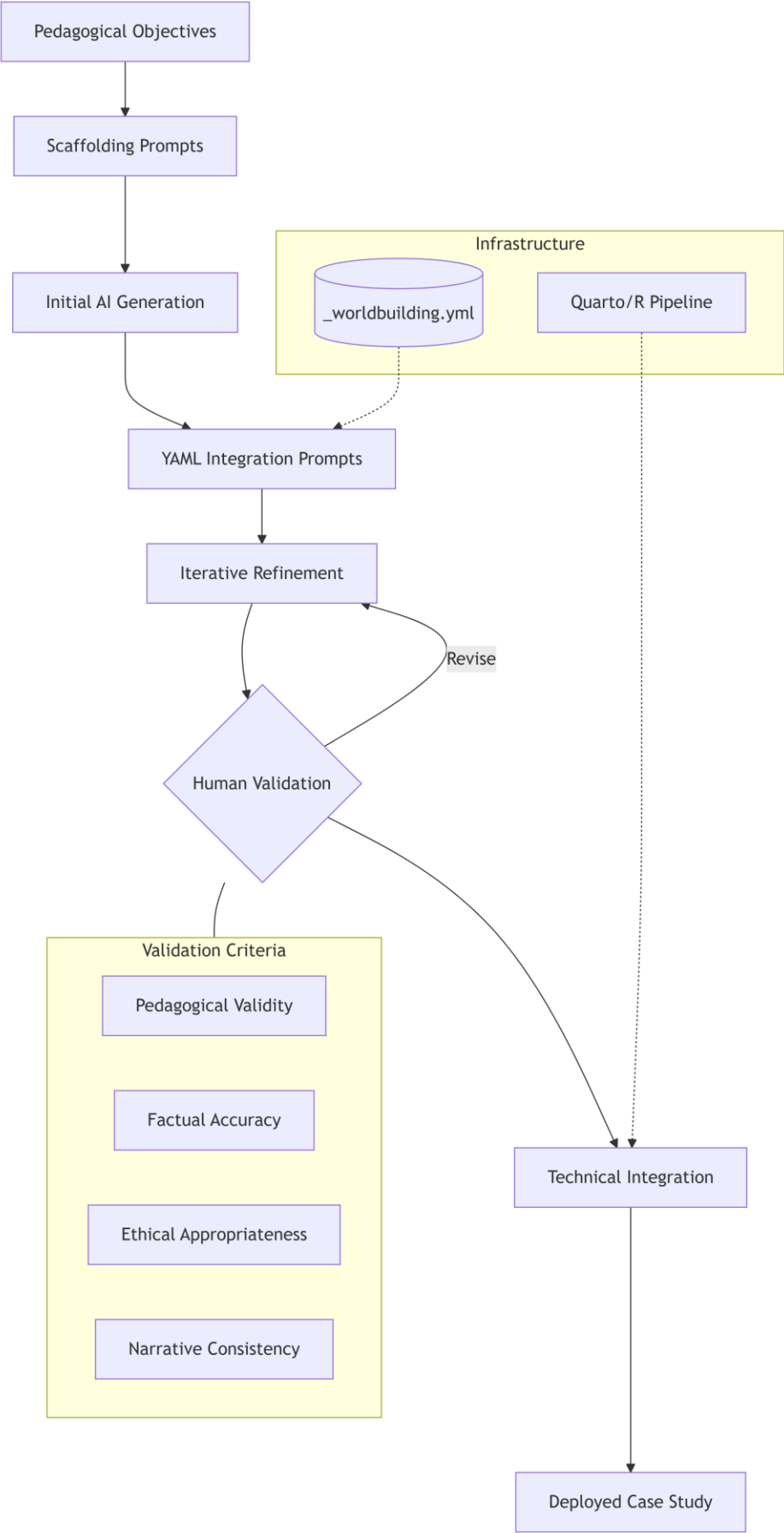


Figure 1: AI-Human Collaboration Framework for Educational Worldbuilding

Interdependently, these four components strengthen one another. Structured prompting without validation risks pedagogically inappropriate content; validation without YAML consistency checking permits narrative contradictions; technical infrastructure without systematic prompting merely automates poor practice. Together, these components create conditions for scalable worldbuilding that maintains both pedagogical rigor and narrative coherence.

Aiming for reproducibility without prescriptiveness, this framework documentation provides sufficient detail for adaptation while remaining open to alternative implementations. Educators might adopt different YAML schemas, employ different prompting strategies, implement different validation criteria, or use different technical tools. What transfers is the systematic approach: treating worldbuilding as documented practice, positioning AI as constrained collaborator, validating rigorously, and maintaining narrative consistency through structured data management.

## **Implementation: Worldbuilding in Practice**

Guided by this framework, we developed seven interconnected case studies deployed across multiple semesters in undergraduate cybersecurity education. We document here both the worldbuilding implementation and its pedagogical integration, offering concrete examples that illustrate the methodology in practice.

### **Course Context**

The worldbuilding cases serve an undergraduate course in cybersecurity law, ethics, and policy, a core requirement for cybersecurity majors at a research university. The course enrolls 25-35 students per semester, primarily juniors and seniors with technical backgrounds but limited social science training. Course objectives emphasize critical analysis of socio-technical systems, ethical reasoning about technology-mediated harms, and understanding regulatory frameworks governing cyberspace.

Traditional case study approaches confronted persistent challenges in this context. Students with strong technical skills often struggled with ethical ambiguity, seeking “correct answers” to dilemmas that require situated judgment. Isolated cases failed to convey how technological controversies evolve, how organizations navigate cumulative ethical failures, and how advocacy campaigns develop strategic sophistication. Worldbuilding addressed these limitations by creating narrative continuity across an entire semester’s worth of cases.

Beyond direct classroom application, this framework carries implications for faculty professional development. The systematic methodology provides a replicable model that teacher educators might incorporate into instructional design training, demonstrating how educators can leverage AI collaboration without sacrificing pedagogical control. This transfer potential extends the framework’s relevance beyond single-course implementation to broader educational technology adoption conversations.

### **The Fictional Universe Implementation**

Seven cases share a common fictional universe, each illuminating distinct ethical and legal concepts while maintaining narrative coherence across the semester’s arc. Spanning the breadth of contemporary cybersecurity policy concerns, these cases address distinct but interconnected challenges. SPECTER explores right-to-be-forgotten legislation and digital privacy policy, while TechGiant examines AI companion ethics and emotional manipulation in consumer technology. AVALANCHE addresses government surveillance, whistleblowing, and national security tensions, whereas Reverb investigates platform content moderation and online speech governance. EXODUS navigates trade secrets, employee mobility, and intellectual property disputes. Co(g)nect probes brain-computer interfaces and cognitive liberty, and Bellator confronts autonomous weapons and AI in economic warfare. These cases connect through shared entities: Jordan Pham, the investigative journalist, first appears in SPECTER investigating a major data breach, then resurfaces in Reverb analyzing the same company’s content moderation failures, allowing students to trace how an organization’s privacy missteps create conditions for later platform governance crises.

Populating these cases are persistent entities documented in a YAML file containing over 170 named individuals, 40 organizations, and dozens of technologies, locations, and documented events. These recurring elements create the narrative continuity that distinguishes worldbuilding from isolated case studies.

Among the persistent characters, Jordan Pham exemplifies how individual trajectories model professional development. This investigative technology journalist for The Silicon Valley Observer appears across multiple cases, initially reporting on data breaches with straightforward privacy concerns and later investigating algorithmic bias and content moderation controversies that demand more sophisticated analytical frameworks. Through Pham's evolving methodology, students observe how journalistic practice adapts to technological complexity. The character's growing sophistication in understanding technical systems, navigating corporate communications strategies, and balancing advocacy with objectivity provides an accessible model for students' own intellectual development.

Organizational trajectories demonstrate institutional evolution in response to technological change. Digital Rights Now, a fictional advocacy organization appearing in five cases, illustrates how social movements adapt strategies and expand scope as new challenges emerge. Students first encounter the organization focused narrowly on privacy protection, deploying traditional advocacy tactics such as public campaigns and legislative lobbying. Across subsequent cases, they observe Digital Rights Now's strategic expansion to address algorithmic transparency, platform governance, and AI ethics, mirroring how actual advocacy organizations evolved from single-issue privacy groups to comprehensive digital rights advocates. This trajectory helps students understand that institutional responses to technology evolve through learning, adaptation, and strategic recalibration rather than remaining static.

Beyond human actors, worldbuilding's technological dimension operates through interconnected platforms and systems rather than isolated technologies. A social media platform facing content moderation challenges in one case later confronts algorithmic bias concerns; the recommendation algorithm implicated in privacy violations reappears driving different ethical dilemmas. Through these recurring technological actors, students come to recognize technologies as evolving systems embedded in broader socio-technical contexts rather than static artifacts operating in isolation.

Equally significant, the worldbuilding establishes a shared regulatory environment through persistent institutional actors. Fictional regulatory bodies (the Federal Privacy Commission and the Digital Platform Oversight Board) persist across cases, creating consistent institutional context that shapes stakeholder options and constraints. Students observe how regulatory responses in early cases establish precedents that subsequently constrain or enable actions in later scenarios, developing understanding of policy path dependencies and institutional evolution.

## **Pedagogical Integration**

Worldbuilding serves specific pedagogical functions beyond narrative engagement, creating structured opportunities for increasingly sophisticated engagement with cybersecurity ethics concepts. Scaffolded complexity emerges through deliberate conceptual progression across the semester. Early-semester cases introduce fundamental concepts, including privacy, consent, and harm, through relatively straightforward scenarios where ethical boundaries appear clear and stakeholder positions remain distinct. Later cases revisit these same concepts in more complex contexts where students must recognize familiar principles operating under different constraints. The persistent narrative ecosystem facilitates this scaffolding without requiring entirely new contextual setup for each case, allowing pedagogical focus to remain on conceptual sophistication rather than context acquisition.

Transfer and application emerge naturally when students encounter familiar organizations, characters, or technologies across multiple scenarios. This recognition prompts them to bring knowledge from previous cases to bear on new situations, mirroring authentic professional expertise. Practitioners rarely analyze situations in isolation but instead draw on accumulated understanding of actors, institutions, and precedents. Worldbuilding creates pedagogical affordances for practicing this transferability, enabling students to recognize patterns, apply precedents, and anticipate stakeholder responses based on established behavioral histories.

Historical contextualization develops through accumulated narrative events that create shared reference points across the scenario setting. A data breach that occurred in Case 3 shapes stakeholder positions in Case 5's

content moderation debate; regulatory responses to early privacy violations constrain available options in subsequent algorithmic bias scenarios. Through these connections, students develop habits of historical reasoning, understanding current controversies as emerging from prior decisions and institutional trajectories rather than appearing spontaneously. This temporal dimension proves particularly valuable in technology ethics, where rapid change often obscures historical continuities.

Perhaps most significantly, persistent characters model ethical reasoning evolution across the semester's narrative arc. Jordan Pham's reporting shifts from straightforward privacy advocacy to more nuanced engagement with competing values as the journalist encounters increasingly complex scenarios. Digital Rights Now confronts internal debates about strategic priorities, balancing ideological purity against coalition-building pragmatism. These narrative arcs provide accessible models for students' own ethical development, demonstrating that principled positions can evolve through encounter with complexity without abandoning core commitments. The characters' struggles with ambiguity normalize the difficulty of ethical reasoning while modeling productive approaches to navigating uncertainty.

### **Assessment Through Worldbuilding**

Worldbuilding enables assessment approaches that traditional isolated cases cannot support, creating opportunities to evaluate higher-order cognitive skills through narrative continuity. Cumulative analysis assignments leverage the persistent fictional universe to demand sophisticated synthesis across multiple scenarios, moving beyond single-case analysis toward longitudinal reasoning. Students might trace an organization's evolution across three cases, analyzing how institutional responses to early challenges shape subsequent options and constraints. This analytical work requires synthesizing multiple scenarios, recognizing organizational patterns, and evaluating long-term consequences of strategic decisions. Such higher-order cognitive skills (pattern recognition, temporal reasoning, and strategic analysis) prove difficult to assess through isolated case analysis where each scenario exists independently.

Persistent entities enable substantive perspective-taking exercises that move beyond superficial role-play. When students adopt stakeholder roles, whether Jordan Pham writing a follow-up investigation, Digital Rights Now revising advocacy strategy, or corporate ethics officers responding to cumulative controversies, they can research character history within the accumulated case materials. This depth transforms perspective-taking from imaginative exercise into evidence-based analysis. Students must maintain consistency with established character positions while responding to new challenges, producing more sophisticated engagement with stakeholder constraints and motivations.

Counterfactual reasoning assignments become pedagogically tractable through worldbuilding's narrative coherence. Students might revise earlier case outcomes and trace implications through subsequent scenarios: if the company in Case 2 had adopted different privacy practices, how would Case 4's algorithmic bias controversy unfold differently? This counterfactual reasoning, central to policy analysis and ethical judgment, requires understanding causal relationships and institutional dynamics. The shared narrative logic makes these relationships visible and traceable in ways disconnected cases cannot achieve.

Most significantly, worldbuilding enables assessment of authentic stakeholder complexity rather than simplified ethical decision-making. Rather than asking students to identify "the ethical course of action" from an omniscient perspective, assignments position them within specific stakeholder roles navigating competing pressures and constraints. Digital Rights Now must balance ideological purity against coalition-building pragmatism, corporate ethics officers navigate between fiduciary duties and public accountability, and journalists manage the tension between maintaining access and pursuing adversarial scrutiny. The established fictional universe makes these role constraints feel consequential: decisions have implications for future scenarios, relationships affect subsequent options, and reputations accumulate across cases. This consequentiality transforms ethical reasoning from abstract exercise into situated practice.

### **Implementation Challenges**

We encountered four significant challenges across implementation. Cognitive load from boundary confusion emerged when fictional scenarios incorporated actual legislation alongside invented organizations;

students struggled to distinguish between real-world legal frameworks and fictional institutional actors, a problem we addressed through explicit framing mechanisms and consistent naming conventions. Continuity tracking proved unexpectedly difficult as well, since not all students recognized recurring entities during initial encounters with cases. This challenge required explicit continuity markers within case materials and cumulative assignments that rewarded attention to persistent characters and organizations.

Two additional challenges emerged later in deployment. Temporal coherence created productive tension between narrative chronology and pedagogical flexibility, which we resolved by maintaining separate fictional timeline documentation that enabled instructors to reorder cases for pedagogical purposes while preserving narrative coherence for students tracking chronological developments. By late semester, fictional fatigue emerged for some students. We addressed this by grounding fictional scenarios in recognizable real-world contexts: cases reference actual regulatory frameworks, real historical events, and authentic technological systems as backdrop while maintaining entirely fictional characters and organizations. Explicitly positioning these scenarios as analytical training grounds for engaging with real-world complexity, rather than as replacements for authentic contexts, helped students understand the pedagogical purpose of the fictional frame.

## Findings: Observational Evidence from Classroom Implementation

We present findings from three semesters of implementing worldbuilding cases in undergraduate and graduate cybersecurity policy education. These observations reflect practitioner experience rather than controlled research, representing patterns that warrant systematic investigation while documenting what became visible through sustained classroom deployment. Table 1 summarizes the implementation scope and context.

| Component                 | Details                                            |
|---------------------------|----------------------------------------------------|
| Cases deployed            | 7 interconnected scenarios                         |
| Implementation period     | Fall 2024–Fall 2025 (3 semesters)                  |
| Undergraduate enrollments | F24: 24; S25: 26; F25: 36 (n=86 total)             |
| Graduate enrollment       | F25: 24 students                                   |
| Course level              | Upper-division undergraduate and graduate          |
| Delivery format           | Hybrid (synchronous in-person and online sections) |
| YAML entities documented  | 170+ named individuals, 40+ organizations          |

Table 1: Implementation summary. Note: n=86 reflects undergraduate students; total n=110 includes graduate students.

### Student Engagement with Narrative Elements

Observable student engagement manifested in several consistent patterns. We observed spontaneous entity recognition without explicit prompting: students regularly demonstrated unprompted awareness of recurring characters and organizations, noting connections such as “Jordan Pham reported on this in the earlier privacy case” or “Digital Rights Now has been fighting these issues since the data breach scenario.” Most tellingly, students sometimes corrected perceived continuity errors in draft materials, suggesting worldbuilding created expectations of narrative coherence worth preserving.

Narrative synthesis further distinguished worldbuilding-based analysis from traditional isolated cases. Students drew connections across cases with notably greater frequency. They referenced earlier advocacy efforts, discussed how prior controversies shaped current positioning, and traced technological evolution across scenarios. We did not systematically collect formal comparison data, but instructor observation across four semesters suggested meaningful differences in analytical approach. The worldbuilding framework appeared to activate different analytical habits, prompting students to seek patterns and trajectories rather than treating each case as discrete.

Character-based reasoning in role-playing assignments proved particularly striking. Students writing as Jordan Pham adopted investigative journalism conventions naturally; they cited “previous reporting” and maintained critical distance consistent with the character’s established skepticism. These performances were substantive.

Worldbuilding appeared to transform stakeholder roles from arbitrary perspectives into positions with histories, constraints, and established behavioral patterns that students felt obligated to honor.

### **Learning Outcomes and Conceptual Transfer**

We cannot claim causality without controlled comparison. Several learning outcome patterns appeared more pronounced with worldbuilding implementation nonetheless, and they merit documentation even as we acknowledge the limitations of practitioner observation.

Development of ethical nuance emerged most strikingly. Rather than seeking definitive “correct answers,” students increasingly framed analyses around competing values and contextual constraints. Jordan Pham’s narrative arc, evolving from straightforward privacy advocate to nuanced critic engaging multiple value frameworks simultaneously, appeared to model this sophistication. Late-semester assignments demonstrated comfort with ambiguity that earlier submissions lacked; students had learned to dwell productively in uncertainty.

Historical contextualization capabilities developed through worldbuilding’s temporal dimension. Students demonstrated stronger grasp of technological controversies as historically situated, routinely referencing earlier cases when analyzing later scenarios. This matters. Historical reasoning requires understanding present dilemmas as emerging from accumulated prior decisions, a cognitive move that proved significantly more challenging with isolated cases presenting acontextual problems.

Recognition of regulatory complexity evolved as well. Early-semester responses often proposed simplistic solutions, typically “pass a law” framings that treated regulation as straightforward. By semester’s end, student analyses acknowledged competing objectives, lobbying influences, and implementation challenges. The persistent regulatory bodies across cases supported this development. Students observed regulators navigating complex terrain rather than wielding unconstrained authority, and this observation appeared to reshape their understanding of policy processes. Transfer across cases produced the most sophisticated analytical work; cumulative assignments revealed students identifying patterns, tracing causality, and anticipating stakeholder responses based on established behavioral precedents.

### **Qualitative Feedback and Student Reflections**

Informal student feedback emerged primarily through unsolicited comments in assignment reflections and course discussions rather than systematic course evaluation analysis. Students frequently characterized the fictional world as “rich and nuanced,” noting that realistic materials provided multiple entry points for exploring content and guiding analysis. Representative comments highlighted narrative continuity as enhancing motivation: “It felt like following a story rather than just analyzing disconnected problems.” The affective dimension—caring about fictional entities—appeared to create investment in understanding complex issues.

Initial concerns about authenticity often evolved across the semester. Students who initially questioned fictional scenarios’ value later noted that “the fictional cases actually helped me understand the real issues better because I could see how things connected over time.” Graduate students, who engaged with the same cases during Fall 2025, offered particularly sophisticated observations about how the persistent fictional universe supported deeper policy analysis than isolated real-world cases typically permit. Preference variation remained consistent across both levels, suggesting worldbuilding aligns better with certain learning styles, with students favoring narrative approaches appearing more receptive than those preferring concrete factual analysis.

### **Implementation Insights**

The strategies that proved most effective shared a common principle: making implicit narrative connections explicit for students encountering the fictional universe for the first time. Continuity markers, implemented as textbox reminders of previous entity appearances, helped students recognize recurring characters and organizations they might otherwise overlook. These explicit prompts proved essential early in the semester but became less necessary as students developed expectations of narrative coherence and began seeking connections independently. Cumulative assignments reinforced this development by requiring cross-case analysis that rewarded attention to persistent entities.

Character design choices similarly shaped pedagogical outcomes. Complex, flawed characters proved more effective than archetypal heroes and villains for supporting nuanced ethical reasoning; students engaged more substantively when stakeholder positions admitted legitimate disagreement rather than clear moral judgment. Scaffolded introduction, beginning with relatively standalone cases before adding narrative complexity, eased students into the worldbuilding framework without overwhelming them with connections they lacked context to appreciate.

Iterative refinement produced additional adaptations. Visual distinctions between fictional scenarios and real-world legal frameworks helped students navigate boundary confusion. Supplementary timeline documentation enabled pedagogical flexibility while preserving narrative coherence for students tracking chronological developments. Automatically generated character reference guides supported student memory without requiring case rereading, addressing the cognitive load that accumulated narrative complexity imposed.

### **Limitations of Practice-Based Evidence**

These findings emerge from practitioner observation rather than systematic research design. We lack controlled comparison groups, standardized learning outcome measures, or longitudinal tracking beyond individual semesters. Student feedback reflects self-reported perceptions rather than objective assessment. The patterns we identify suggest directions for future investigation but do not constitute empirical validation of worldbuilding's pedagogical effectiveness.

What we can claim with confidence is more modest but still meaningful: worldbuilding cases proved pedagogically viable at scale, students engaged with narrative elements in ways that traditional cases did not produce, and implementation challenges proved manageable with iterative adaptation. Whether worldbuilding produces superior learning outcomes compared to traditional approaches remains an empirical question requiring more rigorous investigation.

## **Discussion**

Our implementation of AI-assisted worldbuilding for cybersecurity ethics education reveals both possibilities and constraints worthy of broader consideration. We organize this discussion around three themes: methodological contributions to educational AI collaboration, transferability beyond cybersecurity contexts, and limitations that bound our claims.

### **Methodological Contributions: Systematic AI Collaboration**

By advancing educational AI integration beyond ad hoc experimentation toward systematic methodology, this framework positions AI as constrained collaborator operating within explicit pedagogical boundaries. Structured prompts specify learning objectives, ethical boundaries, complexity levels, and narrative consistency requirements. Human judgment remains authoritative while leveraging AI's generative capacity. This directly addresses concerns about quality control in educational AI deployment (Kasneci et al., 2023).

YAML-based entity management transforms worldbuilding maintenance from memory challenge to data management practice, creating an external memory system that both humans and AI can reference. The validation framework transforms quality control into productive pedagogical practice. Our four-criterion system—pedagogical validity, fictional universe integrity, ethical appropriateness, and narrative consistency—deepens educator engagement with content while ensuring AI-generated materials meet pedagogical standards before deployment.

### **Transferability: Beyond Cybersecurity Education**

While developed for cybersecurity ethics, this framework transfers to educational contexts sharing similar pedagogical challenges. Fields requiring complex systems thinking benefit particularly from durable fictional contexts, as students must understand how decisions ripple across interconnected systems, how institutional histories constrain present options, and how stakeholders navigate competing pressures. Disciplines where case-based pedagogy dominates (such as law, business, medicine, and social work) confront the same fragmentation challenges

worldbuilding addresses. In such contexts, isolated cases prevent students from recognizing how professional situations connect and accumulate over time. Fields where ethical reasoning involves navigating competing values benefit from worldbuilding's capacity to model sophisticated moral development through persistent characters whose thinking evolves across scenarios.

Adaptation requires domain-specific customization, as different fields would document different entity types and employ different validation criteria. AI collaboration reduces initial development time, and once established, YAML infrastructure makes maintaining the fictional universe more tractable than writing entirely new cases. The systematic approach—structured data management, constrained AI collaboration, and rigorous validation—transfers across contexts even as specific implementations vary.

## **Limitations and Boundary Conditions**

Our fundamental limitation involves documenting practitioner experience rather than testing hypotheses through controlled research. We lack randomized comparison groups, standardized assessments, or longitudinal tracking. The patterns observed suggest worldbuilding's pedagogical value but do not constitute empirical validation, and systematic research remains necessary before claiming superior learning outcomes.

Context-specific conditions constrain generalizability. Our framework emerged from a research university with technically proficient students and adequate infrastructure. Different institutional contexts, such as those with varied student preparedness, limited technical resources, or international settings, might require fundamental adaptation. The framework's dependency on AI platforms introduces sustainability risks: changes to platform policies, escalating costs, or institutional restrictions could undermine the methodology. Educators must consider whether materials dependent on commercial AI services align with their sustainability requirements.

Narrative continuity imposes structural constraints alongside pedagogical benefits, as established entities constrain subsequent scenario development in ways that may not suit courses needing rapid response to current events. Cultural and linguistic limitations reflect our US higher education context, and transferring to different linguistic or cultural contexts would require substantial reconceptualization beyond translation.

## **Implications for Educational Practice**

Our experience suggests that educators considering worldbuilding adoption benefit from beginning modestly rather than ambitiously. Starting with simple recurring elements, perhaps one character or organization appearing across two cases, allows instructors to develop entity management practices before complexity overwhelms capacity. Early infrastructure investment proves essential precisely because it prevents the cascading inconsistencies that emerge when documentation standards and version control arrive only after extensive content already exists.

Validation demands particular attention. AI generates fluent text that can mask pedagogical misalignment, factual inaccuracies, or ethical problems; this fluency requires systematic checking rather than superficial review. Transparent communication with students about why fictional scenarios serve pedagogical purposes shapes appropriate engagement while simultaneously supporting AI literacy. Throughout, pedagogy must prevail when narrative interests conflict with learning objectives: the constructed world exists to support student development, not as an end in itself.

## **Conclusion**

Through constructing and deploying seven interconnected case studies for cybersecurity ethics education, we have developed a systematic framework for AI-human collaboration in educational worldbuilding. The framework that emerged differs from what we initially envisioned: we began expecting AI to reduce development time, but discovered that the value lay less in efficiency gains than in the consistency mechanisms, iterative refinement processes, and validation frameworks that AI collaboration made both necessary and tractable. We offer this methodological documentation as a reproducible approach that others might adapt, critique, and extend.

Four integrated components comprise the framework: structured prompt engineering that constrains AI generation toward pedagogical objectives; YAML-based entity management that maintains narrative consistency at scale; systematic validation ensuring pedagogical appropriateness, fictional universe integrity, and ethical soundness; and reproducible technical infrastructure enabling dynamic content integration. Together, these components enable what individual educators would find exceedingly difficult to achieve through either unaided authorship or unstructured AI use: coherent fictional universes populated with persistent characters, organizations, and technologies that create narrative continuity across multiple pedagogical scenarios.

Classroom implementation revealed both affordances and constraints. Students engaged with recurring narrative elements in ways traditional isolated cases did not produce, recognizing characters and organizations, synthesizing across scenarios, and developing more sophisticated ethical reasoning about evolving situations. Implementation challenges proved manageable through iterative adaptation: explicit continuity markers, fiction-reality boundary clarification, balanced deployment with real-world cases, and accessible entity reference guides.

We demonstrate that systematic worldbuilding proves pedagogically viable at classroom scale, that AI collaboration can support rather than undermine pedagogical rigor when properly constrained and validated, and that narrative coherence creates learning affordances isolated cases cannot achieve. Whether these affordances translate to superior learning outcomes compared to traditional approaches remains an empirical question requiring more rigorous investigation than our practitioner observations provide.

For the broader educational technology community, this work illustrates productive middle ground between AI skepticism and uncritical adoption. Lang (2024) argues that generative AI integration requires redesigning courses to promote authentic learning within AI-driven environments, and our worldbuilding framework offers one model for such redesign. We position AI as collaborator whose capabilities prove valuable when systematic methodology constrains generation, rigorous validation ensures appropriateness, and human judgment remains authoritative. This positioning avoids both technophobic dismissal and uncritical adoption, offering a template for other educational AI integration efforts through the model of AI as constrained collaborator within a systematic pedagogical framework.

## Implementation Essentials

For educators seeking to implement worldbuilding, we recommend beginning with modest scope: two interconnected cases, one recurring character, and one organization. Educators should establish entity documentation immediately and use scaffolded prompts specifying learning objectives before requesting content generation. Explicit continuity markers prove essential, as narrative connections are not self-evident to students encountering cases for the first time. Finally, balancing worldbuilding with real-world cases addresses varied learning preferences. These elements constitute minimum viable implementation.

Open for adaptation, this framework accommodates different entity schemas, prompting strategies, or validation criteria while preserving the systematic approach. For those confronting fragmented student experience, desire for narrative coherence, or resource constraints, our framework offers a tested starting point; for educational technology scholars, empirical grounding for future research on worldbuilding's learning impacts and AI collaboration's pedagogical consequences.

## References

- Abdul-Jabbar, W. Kh., & Bhatt, I. (2025). Post-authorship and AI-Assisted Writing: Singularity, Communication, and Pedagogy in Higher Education. *Postdigital Science and Education*. <https://doi.org/10.1007/s42438-025-00586-5>
- An, Q., Yang, J., Xu, X., Zhang, Y., & Zhang, H. (2024). Decoding AI ethics from Users' lens in education: A systematic review. *Heliyon*, 10(20). <https://doi.org/10.1016/j.heliyon.2024.e39357>
- Blanken-Webb, J., Palmer, I., Deshaies, S.-E., Burbules, N. C., Campbell, R. H., & Bashir, M. (2018). A Case Study-based Cybersecurity Ethics Curriculum. *ASE @ USENIX Security Symposium*.

- Breuleux, Y., De Coninck, B., & Therrien, S. (2019). The World Building Framework for Immersive Storytelling Projects. *SHS Web of Conferences*, 64, 00003. <https://doi.org/10.1051/shsconf/20196400003>
- Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2023). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International*, 1–15. <https://doi.org/10.1080/14703297.2023.2195846>
- Friedrichsen, D. (2021). *Building Worlds, Creating Atmospheres: Postcritique, Worldbuilding, and the Properties of Speculative Fiction* [PhD thesis, Justus Liebig University Giessen]. <https://doi.org/10.22029/JLUPUB-18229>
- Giray, L. (2023). Prompt Engineering with ChatGPT: A Guide for Academic Writers. *Annals of Biomedical Engineering*, 51(12), 2629–2633. <https://doi.org/10.1007/s10439-023-03272-4>
- Kampe, C. (2022). *Complex Games and Game Design as a Vehicle for Technical & Critical Learning* [PhD thesis, North Carolina State University]. <https://repository.lib.ncsu.edu/items/dd40e4b6-5550-4d82-a814-2185ca34864b>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Lang, J. C. (2024). Embracing Generative AI for Authentic Learning. *Creative Education*, 15(1), 1–20. <https://doi.org/10.4236/ce.2024.151001>
- Lave, J., & Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press.
- Mahdi, O. R., Nassar, I. A., & Almuslamani, H. A. (2020). The Role of Using Case Studies Method in Improving Students' Critical Thinking Skills in Higher Education. *International Journal of Higher Education*, 9(2), 297. <https://doi.org/10.5430/ijhe.v9n2p297>
- Veletsianos, G., Houlden, S., & Johnson, N. (2024). Is Artificial Intelligence in Education an Object or a Subject? Evidence from a Story Completion Exercise on Learner-AI Interactions. *TechTrends*. <https://doi.org/10.1007/s11528-024-00942-5>
- White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J., & Schmidt, D. C. (2023). *A Prompt Pattern Catalog to Enhance Prompt Engineering with ChatGPT*. arXiv. <https://doi.org/10.48550/ARXIV.2302.11382>
- Williamson, B. (2024). The Social life of AI in Education. *International Journal of Artificial Intelligence in Education*, 34(1), 97–104. <https://doi.org/10.1007/s40593-023-00342-5>
- Wolf, M. J. P. (2012a). Building Imaginary Worlds: The Theory and History of Subcreation. Taylor & Francis Group.
- Wolf, M. J. P. (2012b). Introduction. In *Building Imaginary Worlds: The Theory and History of Subcreation*. Taylor & Francis Group.