A Five-Level Classification Framework for Intelligent Textbooks: Lessons from Autonomous Vehicle Standards

Dan McCreary
https://www.linkedin.com/in/danmccreary/

Version 0.01
November 29, 2025

Abstract

The rapid advancement of artificial intelligence is transforming educational content, yet the field lacks a standardized framework for classifying intelligent textbook capabilities. Inspired by automotive industry standards for autonomous vehicles, we propose a five-level classification: (1) Static textbooks; (2) Interactive textbooks with multimedia and assessments; (3) Adaptive textbooks that adjust to individual performance; (4) Chatbot-integrated textbooks using large language models; and (5) Autonomous AI textbooks with comprehensive real-time personalization.

Evidence from METR research shows AI task capabilities doubling every seven months. Extrapolating to 2030, the cost of producing Level 2 content approaches zero—pennies per student per day. This commoditization challenges publishers whose value rests on content production alone.

A critical finding is the identification of Level 3 as a privacy inflection point. Below this threshold, textbooks require minimal student data. At Level 3 and above, systems require detailed learning histories and behavioral patterns, raising concerns under FERPA, COPPA, and GDPR. We argue this threshold demands differentiated governance, with higher levels requiring stronger privacy protections. Educational standards including xAPI and Learning Record Stores can enable both personalization and student-controlled data portability.

The strategic implication: as Level 1–2 content becomes freely available, educational organizations must focus on Level 3+ capabilities to remain viable. Our framework provides a common vocabulary for evaluating products, establishing procurement criteria, and developing level-specific regulations.

Keywords: intelligent textbooks, educational technology, artificial intelligence, adaptive learning, privacy, classification framework, autonomous systems, large language models, xAPI, Learning Record Store, METR, exponential growth

1 Introduction

The convergence of artificial intelligence and educational publishing is creating a new category of learning resources that defy traditional classification. Products marketed as "intelligent," "adaptive," or "AI-powered" textbooks vary enormously in their actual capabilities—from simple keyword search to sophisticated conversational tutoring. This ambiguity creates significant challenges for educational institutions attempting to evaluate, procure, and regulate these emerging technologies.

The problem is not merely semantic. Without clear classification, educators cannot make informed decisions about which tools best serve their pedagogical goals. Administrators struggle to develop appropriate policies for data governance. Policymakers lack the precision needed to craft effective regulations. And students remain uninformed about what data they are sharing and how it will be used.

This paper proposes a solution: a five-level classification framework for intelligent textbooks, modeled on the successful SAE J3016 standard for autonomous vehicles. Just as that framework transformed discourse around self-driving cars—enabling clear communication between manufacturers, regulators, and consumers—we argue that a similar framework can bring much-needed clarity to the educational technology landscape.

The genesis of this framework was a blog post written to initiate discussion within the educational technology community [1]. The response was substantial: dozens of educators, technologists, and policymakers provided valuable feedback that has shaped and refined the framework presented here. This paper represents the synthesis of those conversations, formalized for broader academic discourse.

Our framework is motivated by three converging trends:

1. **Exponential AI Growth**: Large language models have progressed from experimental curiosities to production-ready tutoring systems in just a few years. Educational institutions need frameworks that can accommodate this pace of change.

2. **Proliferating Products**: The market is flooded with AI-enhanced educational products, each claiming unique capabilities. Standardized classification enables meaningful comparison.

3. **Privacy Imperatives**: Higher levels of textbook intelligence require increasingly detailed student data, raising questions that current regulatory frameworks were not designed to address.

The remainder of this paper is organized as follows. Section 2 examines the economics of AI-generated educational content, presenting evidence from METR research that content creation costs are dropping exponentially. Section 3 reviews related work in educational technology classification. Section 4 examines the SAE J3016 autonomous vehicle standard and extracts lessons for educational technology. Section 5 presents our five-level framework in detail. Section 6 analyzes the privacy inflection point at Level 3. Section 7 discusses relevant educational technology standards. Section 8 addresses implementation considerations. Section 9 discusses limitations and future work. Section 10 concludes.

2 The Economics of Educational Content Creation

A fundamental driver of the intelligent textbook framework is the rapidly declining cost of content creation. This section examines empirical evidence for this trend and its implications for educational publishers and institutions.

2.1 The METR Study: Measuring AI Task Capabilities

In March 2025, the Model Evaluation and Threat Research (METR) organization published groundbreaking research on measuring AI capabilities in terms of *task horizons*—the duration of human tasks that AI systems can complete with specified success probabilities [2].

The METR methodology represents a paradigm shift in AI evaluation. Rather than measuring narrow benchmarks (e.g., question answering accuracy, code completion), METR measures the **length of realistic human tasks** that an AI agent can complete autonomously with a 50% (or 80%) probability of success.

Table 1: AI Task Horizon Growth (50% Success Probability)

Model	Release Date	Horizon (min)	Horizon (hours)
GPT-2	Feb 2019	2.4	0.04
davinci-002	May 2020	8.9	0.15
GPT-3.5	Mar 2022	36.3	0.6
GPT-4	Mar 2023	321.8	5.4
GPT-4o	May 2024	550.2	9.2
Claude 3.5 Sonnet	Jun 2024	1,092.9	18.2
o1-preview	Sep 2024	1,325.7	22.1
Claude 3.7 Sonnet	Feb 2025	3,253.6	54.2
o3	Apr 2025	5,530.7	92.2
GPT-5	Aug 2025	8,239.1	137.3

2.2 The Seven-Month Doubling Time

Analysis of the METR data reveals a remarkably consistent exponential growth pattern, as shown in Figure 1. AI task completion capabilities have been **doubling approximately every seven months**. This growth rate can be expressed as:

$$H(t) = H_0 \times 2^{t/7} \tag{1}$$

where $H(t)$ is the task horizon in minutes at time t (months from baseline), H_0 is the baseline horizon, and the doubling time is 7 months.

To contextualize this growth rate: from February 2019 (GPT-2) to August 2025 (GPT-5), task horizons increased from 2.4 minutes to 8,239 minutes—a factor of approximately 3,400× in 78 months, or roughly 11.7 doublings ($78/7 \approx 11.1$).

2.3 Extrapolation to 2030

If the seven-month doubling time continues, AI capabilities by 2030 would be extraordinary. Table 2 shows the projected task horizons.

Table 2: Projected AI Task Horizons (2025–2030)

Date	Horizon (min)	Horizon (hours)	Horizon (days)
Aug 2025 (baseline)	8,239	137	5.7
Mar 2026 (+7 mo)	16,478	275	11.5
Oct 2026 (+14 mo)	32,956	549	22.9
May 2027 (+21 mo)	65,912	1,099	45.8
Dec 2027 (+28 mo)	131,824	2,197	91.5
Jul 2028 (+35 mo)	263,648	4,394	183
Jan 2030 (+53 mo)	1,318,000	21,970	915

By January 2030, if trends continue, AI systems could reliably complete tasks that would take a human **over two years** of continuous work. Even accounting for significant slowdown, capabilities measured in weeks or months of human-equivalent work appear plausible within this timeframe.

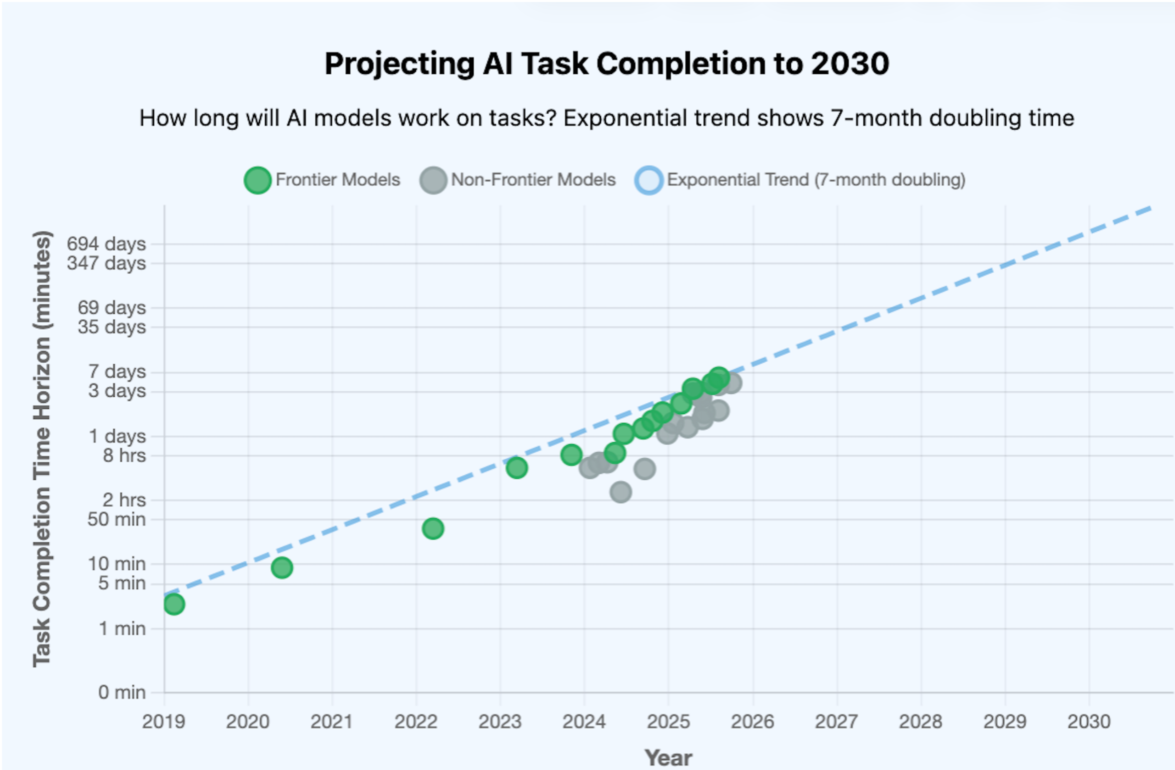


Figure 1: AI Task Horizon Growth (2019–2030). Historical data points (blue) show measured AI capabilities from METR evaluations. The exponential trendline (dashed) extrapolates the 7-month doubling time through 2030, projecting AI systems capable of completing tasks requiring months of human effort.

Important caveats: Exponential trends do not continue indefinitely. Physical limits, diminishing returns, data constraints, or fundamental algorithmic barriers could slow this growth. However, even a significant reduction in growth rate (e.g., 14-month doubling instead of 7-month) would still yield transformative capabilities by 2030.

2.4 Implications for Textbook Production Costs

The METR framework has direct implications for educational content creation. Consider the tasks involved in producing a Level 2 interactive textbook chapter:

- Writing 5,000 words of educational prose: 2–4 hours
- Creating 3 embedded quiz questions: 30–60 minutes
- Developing a simple interactive simulation: 4–8 hours
- Producing a concept diagram: 1–2 hours
- Editing and quality review: 2–4 hours

Total: approximately **10–20 hours** per chapter.

Current frontier models (2025) already have task horizons exceeding this duration. By 2027, AI systems may reliably produce entire textbooks—hundreds of pages with integrated simulations, assessments, and multimedia—in single autonomous sessions.

2.5 The Cost Trajectory

The economic implications are stark. Table 3 estimates the marginal cost of producing a Level 2 textbook chapter.

Table 3: Projected Cost per Level 2 Chapter

Year	Human Hours	AI Cost	Notes
2020	15	N/A	Fully human production
2023	10	\$50–100	AI assists with drafting
2025	3	\$10–20	AI produces first draft; human review
2027	0.5	\$1–5	AI produces complete draft; light review
2030	0.1	\$0.10–0.50	Near-autonomous production

At 2030 cost levels, producing a complete interactive textbook with 30 chapters would cost approximately **\$3–15**. Amortized across even a modest student population, this translates to **pennies per student per course**.

2.6 The Commoditization of Level 1–2 Content

This cost trajectory has profound strategic implications:

Level 1 (static) content becomes essentially free. Any organization can generate high-quality static educational text at negligible cost. The value of static content approaches zero.

Level 2 (interactive) content becomes a commodity. Interactive elements—simulations, quizzes, embedded videos—that once required specialized development will be producible on demand. The barrier to entry for creating engaging digital textbooks disappears.

Differentiation requires Level 3+. If any organization can produce Level 1–2 content at near-zero cost, competitive advantage must come from capabilities that AI cannot easily replicate or that require persistent student relationships:

- **Level 3 (Adaptive)**: Requires institutional data infrastructure, student authentication, and ongoing relationship management
- **Level 4 (Chatbot)**: Requires integration with institutional systems, oversight capabilities, and trust relationships
- **Level 5 (Autonomous)**: Requires deep institutional knowledge and comprehensive student support systems

2.7 Strategic Imperatives

This analysis yields clear strategic imperatives for educational organizations:

1. **Do not compete on content alone**. Organizations whose value proposition is "we produce good educational content" face existential risk as content production costs approach zero.

2. **Invest in data infrastructure**. The capabilities required for Level 3+ systems—student authentication, learning record management, privacy compliance—take years to build. Organizations should begin now.

3. **Build trust relationships**. Higher-level intelligent textbooks require students to share sensitive learning data. Institutions with established trust have a significant advantage.

4. **Develop integration capabilities**. Level 4+ systems must integrate with institutional LMS platforms, authentication systems, and support services. Technical integration capability becomes a competitive moat.

5. **Focus on what AI cannot easily provide**. Human mentorship, accountability structures, credential verification, and community remain valuable even as content becomes free.

2.8 The Penny-Per-Day Vision

Extrapolating current trends suggests a future where high-quality interactive educational content is available to every student at costs measured in pennies per day. A complete interactive textbook for a semester-long course might cost less than a single cup of coffee.

This vision has profound equity implications. Educational content that was once gatekept by expensive textbook publishers could become universally accessible. The question shifts from "Can students afford good textbooks?" to "How do we help students navigate abundant free resources?"

However, this abundance creates new challenges:

- Quality assurance becomes harder when anyone can produce professional-looking content
- Curation and recommendation become more valuable than production
- Institutional brands serve as trust signals in a sea of undifferentiated content
- The value of credentials—proof that learning occurred—increases as content access becomes universal

The organizations that thrive in this future will be those that recognize Level 1–2 content as table stakes and build their value proposition around the adaptive, interactive, and relationship-based capabilities of Levels 3–5.

3 Related Work

The classification of educational technologies has been approached from multiple perspectives in the literature.

3.1 Adaptive Learning Systems

Research on adaptive learning systems dates to the 1970s with early intelligent tutoring systems (ITS). Woolf [3] provides a comprehensive overview of ITS architectures, distinguishing between systems based on their student modeling capabilities. More recently, the rise of massive open online courses (MOOCs) has prompted new taxonomies focused on scalability and personalization [4].

3.2 Educational Technology Maturity Models

Several maturity models have been proposed for educational technology adoption. The Technology Integration Matrix (TIM) [5] classifies technology use along dimensions of student engagement and curricular integration. The SAMR model [6] (Substitution, Augmentation, Modification, Redefinition) focuses on how technology transforms learning activities. However, these models address technology *use* rather than technology *capabilities*.

3.3 AI in Education Frameworks

The emergence of AI in education has prompted new classification efforts. Holmes et al. [7] distinguish between AI applications that support learning management, learner support, and assessment. Luckin et al. [8] propose a framework centered on different types of intelligence that AI systems can exhibit in educational contexts. More recently, Haupt et al. [9] examine the deployment of large language models and AI coding assistants in signal processing education, addressing challenges including hallucination mitigation, fairness, and the development of "smart textbooks" that integrate AI capabilities while maintaining transparency and trustworthiness.

3.4 Privacy in Learning Analytics

The learning analytics community has extensively studied privacy implications of educational data collection. Slade and Prinsloo [10] outline ethical frameworks for learning analytics, while Rubel and Jones [11] analyze student privacy through the lens of contextual integrity. However, these works do not provide systematic classification of systems by their data requirements.

3.5 Gap in the Literature

While existing frameworks address aspects of educational technology classification, none provide a comprehensive, capability-based taxonomy analogous to the SAE J3016 standard. Our framework fills this gap by defining clear levels based on system capabilities while explicitly linking each level to its data requirements and privacy implications.

4 Lessons from Autonomous Vehicle Classification

The Society of Automotive Engineers' J3016 standard, first published in 2014, has become the definitive framework for classifying autonomous vehicle capabilities. Its success offers valuable lessons for educational technology classification.

4.1 The Problem J3016 Solved

Prior to J3016, the autonomous vehicle industry suffered from terminological chaos. Terms like "autopilot," "self-driving," "semi-autonomous," and "driver assist" were used interchangeably by manufacturers, creating confusion for consumers, regulators, and researchers alike. This ambiguity had serious consequences: consumers misunderstood system capabilities, leading to misuse and accidents; regulators struggled to craft appropriate requirements; and researchers lacked a common vocabulary for comparing systems.

4.2 The J3016 Framework

The SAE standard defines six levels of driving automation (0-5), based on who performs the dynamic driving task (human or system) and under what conditions:

Table 4: SAE J3016 Levels of Driving Automation

Level	Name	Description
0	No Automation	Human performs all driving tasks
1	Driver Assistance	System controls steering **or** acceleration; human monitors and controls all other tasks
2	Partial Automation	System controls steering **and** acceleration; human monitors environment and intervenes as needed
3	Conditional Automation	System handles all driving in specific conditions; human must intervene when requested
4	High Automation	System handles all driving in defined operational domains; no human intervention required within those domains
5	Full Automation	System handles all driving in all conditions; no human intervention ever required

4.3 Key Design Principles

Several design principles contributed to J3016's success:

Capability-based definition: Levels are defined by what the system can do, not by the technology used. This makes the framework technology-agnostic and resistant to obsolescence.

Clear responsibility allocation: Each level specifies who (human or system) is responsible for what tasks. This clarity enables appropriate liability frameworks.

Graduated expectations: The framework acknowledges that full automation is a distant goal, setting realistic expectations while validating intermediate achievements.

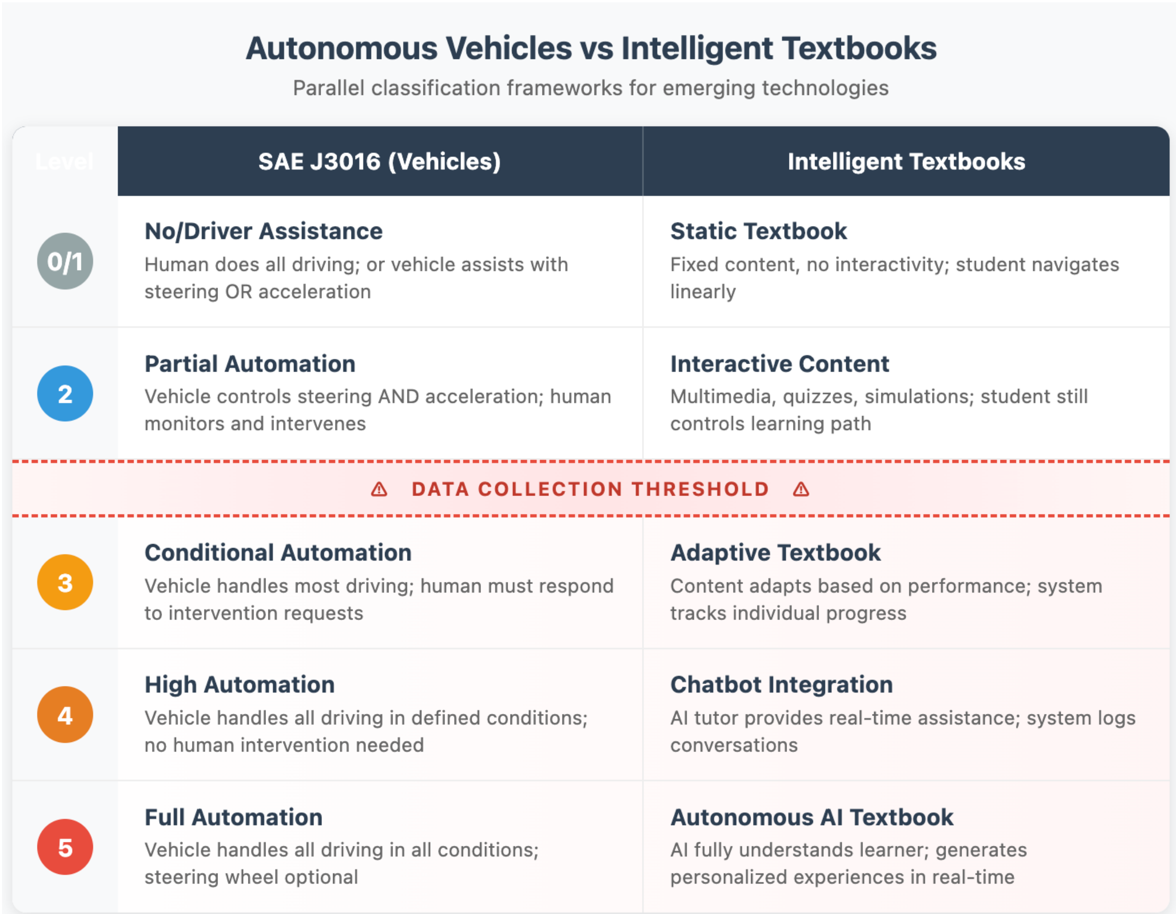
Regulatory compatibility: The framework has been adopted by regulators worldwide, enabling consistent policy development.

4.4 Application to Educational Technology

Figure 2 illustrates the parallel between SAE J3016 autonomous vehicle levels and our proposed intelligent textbook classification.

We apply these principles to intelligent textbook classification:

- Define levels by capability, not technology
- Specify who (student, instructor, or system) controls learning progression
- Acknowledge that full autonomy remains aspirational



Autonomous Vehicles vs Intelligent Textbooks

Parallel classification frameworks for emerging technologies

Level	SAE J3016 (Vehicles)	Intelligent Textbooks
0/1	**No/Driver Assistance** Human does all driving; or vehicle assists with steering OR acceleration	**Static Textbook** Fixed content, no interactivity; student navigates linearly
2	**Partial Automation** Vehicle controls steering AND acceleration; human monitors and intervenes	**Interactive Content** Multimedia, quizzes, simulations; student still controls learning path
	⚠ DATA COLLECTION THRESHOLD ⚠	
3	**Conditional Automation** Vehicle handles most driving; human must respond to intervention requests	**Adaptive Textbook** Content adapts based on performance; system tracks individual progress
4	**High Automation** Vehicle handles all driving in defined conditions; no human intervention needed	**Chatbot Integration** AI tutor provides real-time assistance; system logs conversations
5	**Full Automation** Vehicle handles all driving in all conditions; steering wheel optional	**Autonomous AI Textbook** AI fully understands learner; generates personalized experiences in real-time

Figure 2: Side-by-side comparison of SAE J3016 autonomous vehicle levels and intelligent textbook levels. Both frameworks progress from no automation/static content through increasing levels of system autonomy, with corresponding shifts in human oversight requirements.

- Enable graduated regulatory responses

A critical addition for educational technology is explicit consideration of data requirements at each level, given the sensitivity of student information.

5 The Five Levels of Intelligent Textbooks

We propose a five-level classification framework for intelligent textbooks, ranging from static content to fully autonomous AI-driven learning systems. Figure 3 provides a visual overview of this framework.

5.1 Level 1: Static Textbooks

Definition: Traditional printed or digital formats with fixed content and no interactive elements.

Characteristics:

- Composed of text and static images

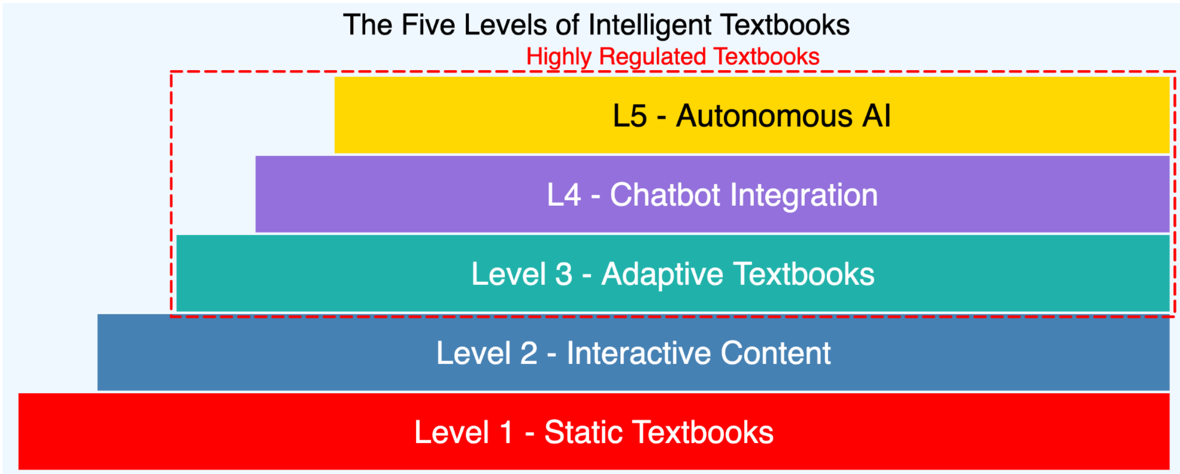


Figure 3: The Five Levels of Intelligent Textbooks. Each level represents increasing capabilities in personalization, interactivity, and AI integration, with corresponding increases in data requirements and privacy considerations.

- Linear progression through material
- No digital interactivity or personalization
- Content does not adapt to the learner

Data Requirements: None. Level 1 textbooks collect no student data.

Current Status: Over 90% of textbooks used by college students today remain at Level 1, including both physical books and static PDFs.

5.2 Level 2: Interactive Content Textbooks

Definition: Digital textbooks incorporating interactive elements that engage readers beyond passive consumption.

Characteristics:

- Keyword search functionality
- Internal and external hyperlinks
- Embedded videos and multimedia content
- Simple self-assessment quizzes with immediate feedback
- Interactive simulations (MicroSims) for concept visualization
- Learning graphs showing concept dependencies and prerequisites
- Detailed glossary with contextual linking
- Social sharing capabilities
- Optional anonymous usage analytics

Data Requirements: Minimal. Anonymous aggregate analytics may be collected (page views, time on page), but no individual student tracking is required for core functionality.

Evidence of Effectiveness: Research on interactive simulations demonstrates significant learning gains. The MicroSims framework [12] reports that interactive simulations can enhance conceptual understanding by 30–40% compared to conventional teaching methods, while AI-assisted creation tools dramatically reduce development costs.

Implementation: Can be achieved with static site generators (e.g., MkDocs), embedded JavaScript simulations, and optional analytics platforms.

5.3 Level 3: Adaptive Textbooks

Definition: Textbooks that dynamically adjust content presentation based on individual learner performance and behavior.

Characteristics:

- Personalized learning pathways through deterministic rules
- Algorithmic traversal of learning graphs (present at Level 2) combined with individual performance data
- Selection of content based on assessment performance
- Spaced repetition for knowledge reinforcement
- Continuous recording of concept mastery
- Prerequisite enforcement and remediation

Data Requirements: Significant. Systems must maintain individual learning histories including assessment results, time spent on concepts, error patterns, and mastery estimates.

Technical Requirements: Learning management system integration, graph databases for concept relationships, student authentication systems.

5.4 Level 4: Textbooks with Chatbots

Definition: Textbooks integrating intelligent conversational interfaces that provide real-time, personalized assistance.

Characteristics:

- Large Language Model (LLM) integration for tutoring
- Natural language question answering about content
- GraphRAG architecture combining retrieval and generation
- Real-time feedback on student questions
- Socratic dialogue capabilities
- Content recommendations based on conversation analysis

Data Requirements: High. Systems log complete conversation histories, including potentially sensitive questions students might not ask human instructors.

Technical Requirements: LLM infrastructure (cloud or local), embedding models, vector databases, careful prompt engineering to ensure accuracy.

5.5 Level 5: Autonomous AI Textbooks

Definition: Systems where AI fully understands individual learner needs and autonomously generates personalized learning experiences across all contexts.

Characteristics:

- Deep understanding of individual student knowledge states
- Real-time generation of customized lessons and explanations
- Autonomous assessment creation tailored to learner gaps
- Complete adaptability to learning styles and preferences
- Proactive intervention when confusion is detected
- Human-like tutoring capabilities without human oversight

Data Requirements: Very high. Systems require comprehensive learner profiles including cognitive patterns, behavioral indicators, and potentially affective states.

Current Status: Aspirational. No current systems reliably achieve Level 5 across all educational contexts. This level requires advances in AI reliability, interpretability, and privacy-preserving computation.

5.6 Summary Comparison

Table 5 summarizes the key distinctions between levels.

Table 5: Comparison of Intelligent Textbook Levels

Level	Type	Personalization	Data Required	AI Dependency
1	Static	None	None	None
2	Interactive	None	Minimal	Optional
3	Adaptive	Rule-based	Individual history	Algorithms
4	Chatbot	Conversational	Conversation logs	LLM
5	Autonomous	Full	Comprehensive	Advanced AI

6 The Privacy Threshold: Level 3 and Above

A critical finding of our analysis is the identification of Level 3 as a privacy inflection point. This section examines why this threshold matters and what it implies for governance. Figure 4 visualizes the dramatic escalation in data collection requirements that occurs at Level 3.

6.1 The Nature of the Threshold

Below Level 3, intelligent textbooks can function with minimal or no student-specific data:

- **Level 1** textbooks are static artifacts that collect nothing.
- **Level 2** textbooks may collect anonymous usage analytics, but personalization does not depend on individual tracking.

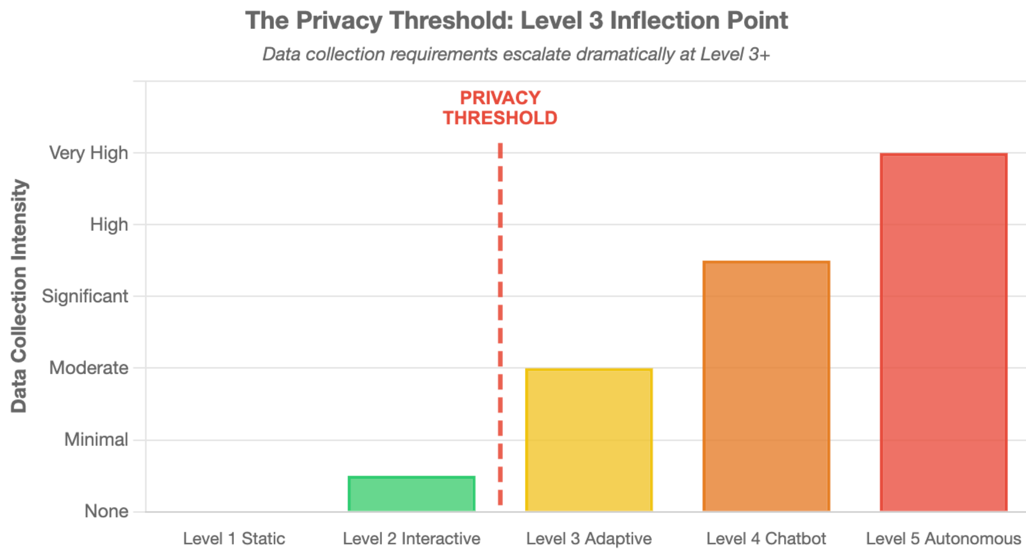


Figure 4: The Privacy Threshold: Data collection intensity by intelligent textbook level. Levels 1–2 (green) require minimal or no student data, while Levels 3–5 (yellow to red) require progressively more extensive data collection, triggering regulatory requirements under FERPA, COPPA, and GDPR.

At Level 3 and above, the fundamental value proposition changes. Adaptivity *requires* knowing how individual students are performing:

- **Level 3** systems track every assessment result, every wrong answer, every moment of hesitation.
- **Level 4** systems log complete conversations, including questions students might never ask a human.
- **Level 5** systems would require comprehensive cognitive profiling.

6.2 Regulatory Implications

This threshold has immediate regulatory consequences:

FERPA (Family Educational Rights and Privacy Act): In the United States, student education records are protected under FERPA. Level 3+ systems generate extensive education records, triggering compliance requirements around access, disclosure, and retention.

COPPA (Children's Online Privacy Protection Act): For K-12 education, COPPA imposes additional requirements on collecting data from children under 13, including parental consent mechanisms.

GDPR (General Data Protection Regulation): In the European Union, GDPR requires explicit consent, data minimization, and the right to erasure. Level 3+ systems must be designed with these requirements from inception.

State Laws: Emerging state-level student privacy laws (e.g., California's SOPIPA) add additional compliance layers.

6.3 Institutional Responsibilities

When institutions adopt Level 3+ textbooks, they assume new responsibilities:

1. **Data Custodianship**: The institution becomes responsible for protecting student learning data, requiring investment in security infrastructure.
2. **Vendor Management**: Contracts with textbook providers must address data ownership, retention, and breach notification.
3. **Audit Capabilities**: Institutions must be able to demonstrate compliance, requiring logging and reporting systems.
4. **Student Rights**: Students must be informed about data collection and provided mechanisms to access or delete their data.

6.4 Algorithmic Concerns

Beyond data privacy, Level 3+ systems raise algorithmic concerns:

Bias: Adaptive algorithms may perpetuate or amplify biases present in training data. A system that routes struggling students to easier content might create self-fulfilling prophecies.

Transparency: Students and instructors should understand why the system is making particular recommendations. Black-box adaptivity undermines trust and pedagogical autonomy.

Manipulation: Systems optimized for engagement metrics might prioritize addictive patterns over genuine learning.

6.5 The Trust Paradox

Level 4 systems present a particular trust challenge. Students may share vulnerabilities with an AI tutor that they would never reveal to a human instructor. Questions like "I don't understand anything in this chapter" or searches for basic concepts can feel exposing. This candor enables better tutoring but creates sensitive records.

Institutions must earn this trust through transparent policies, robust security, and genuine commitment to using data only for educational benefit.

6.6 Recommendations for Governance

We recommend differentiated governance based on level:

7 Educational Technology Standards

Existing educational technology standards provide important foundations for implementing Level 3+ intelligent textbooks while managing privacy concerns. Figure 5 illustrates the relationships between key standards in the educational technology ecosystem.

7.1 Experience API (xAPI)

The Experience API (xAPI), developed by the Advanced Distributed Learning (ADL) Initiative, provides a specification for collecting data about learning experiences. Key features relevant to intelligent textbooks include:

Table 6: Recommended Governance by Level

Level	Governance Requirements
1–2	Standard content review; no special data governance required
3	Privacy impact assessment; data retention policy; student notification; bias auditing
4	All Level 3 requirements plus: conversation data protections; AI accuracy monitoring; human oversight mechanisms
5	All Level 4 requirements plus: comprehensive ethical review; ongoing algorithmic auditing; clear human override capabilities

- **Activity Statements**: xAPI records learning activities as "Actor-Verb-Object" statements (e.g., "Student completed quiz"), providing a standardized vocabulary for learning events.
- **Portability**: Learning records can be transferred between systems, enabling students to maintain continuous learning histories across platforms.
- **Granularity**: xAPI can capture fine-grained interactions (individual question responses, time stamps) needed for Level 3+ adaptivity.

7.2 Learning Record Store (LRS)

The Learning Record Store is the data repository component of the xAPI ecosystem:

- **Centralized Storage**: LRS systems aggregate learning data from multiple sources, enabling cross-platform analytics.
- **Access Control**: LRS implementations can enforce role-based access, ensuring only authorized parties access student data.
- **Student Ownership**: Architectures where students control their own LRS enable data portability while maintaining privacy.

7.3 IEEE Learning Technology Standards

The IEEE Learning Technology Standards Committee (LTSC) has developed several relevant standards:

- **IEEE 1484.12 Learning Object Metadata (LOM)**: Provides vocabulary for describing educational content, enabling discovery and interoperability.
- **IEEE 1484.11 SCORM**: The Sharable Content Object Reference Model enables content packaging and sequencing, though it is being superseded by xAPI for adaptive applications.
- **IEEE P2881**: An emerging standard for learning engineering, which may provide frameworks for validating adaptive system effectiveness.

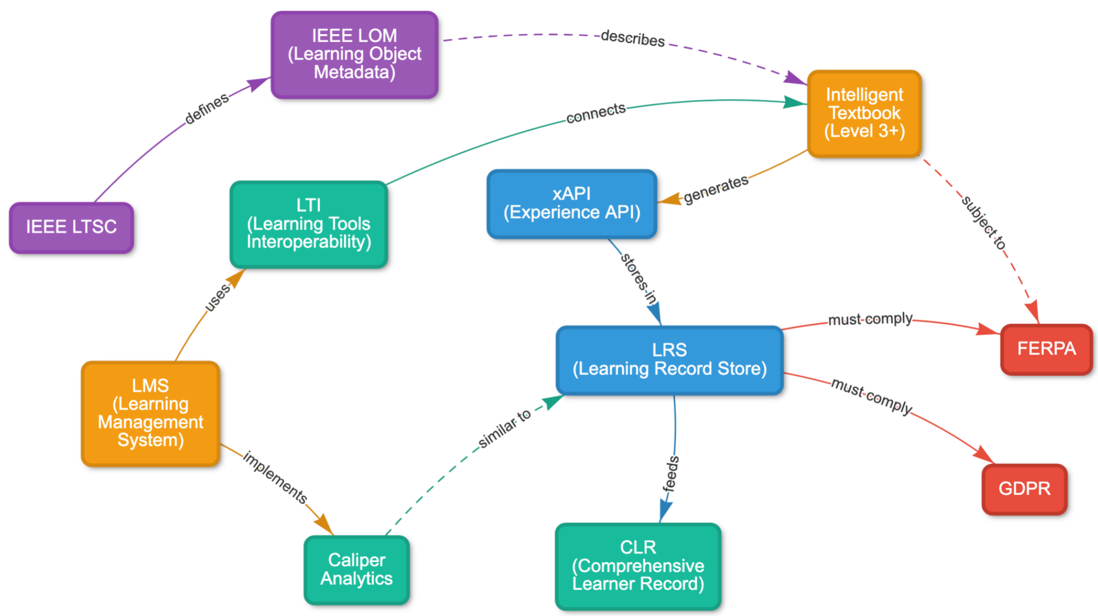


Figure 5: The Educational Technology Standards Ecosystem. This network diagram shows relationships between standards bodies (IEEE, IMS Global, ADL), specifications (xAPI, LTI, Caliper), data stores (LRS), and credential standards (CLR). Arrows indicate data flow and dependency relationships.

7.4 IMS Global Standards

IMS Global Learning Consortium provides additional relevant specifications:

- **LTI (Learning Tools Interoperability)**: Enables seamless integration of external tools with learning management systems.
- **Caliper Analytics**: A competing/complementary standard to xAPI for learning analytics, with stronger alignment to IMS ecosystem tools.
- **Comprehensive Learner Record (CLR)**: Emerging standard for portable credential and achievement records.

7.5 Standards and Privacy

These standards can support privacy in several ways:

Data Minimization: xAPI's granular vocabulary enables collecting only necessary data, avoiding over-collection.

Transparency: Standardized activity statements make it clear what is being collected, supporting informed consent.

Portability: Student-controlled LRS architectures enable data portability without vendor lock-in.

Interoperability: Standards enable switching between providers without losing learning history, reducing switching costs that might otherwise trap students in privacy-hostile ecosystems.

7.6 Gaps and Opportunities

Current standards do not fully address:

- Conversation data from Level 4 chatbot interactions
- Real-time streaming data for Level 5 cognitive modeling
- Privacy-preserving computation (federated learning, differential privacy)
- Algorithmic transparency and auditability

Future standards development should address these gaps as intelligent textbook capabilities advance.

8 Implementation Considerations

This section provides practical guidance for institutions considering intelligent textbook adoption at various levels.

8.1 Flexible Implementation Paths

Organizations need not progress sequentially through all five levels. Depending on specific needs and available technologies, viable strategies include:

Leapfrogging: Moving directly from Level 2 to Level 4 by adding chatbot capabilities to interactive content, bypassing the data infrastructure requirements of Level 3 adaptive systems.

Hybrid approaches: Implementing features from multiple levels simultaneously—for example, Level 2 interactivity with optional Level 4 chatbot assistance.

Domain-specific depth: Achieving higher levels in specific subject areas while maintaining lower levels elsewhere, based on where adaptivity provides greatest benefit.

8.2 Precision in Level Terminology

One of the primary benefits of a standardized classification framework is enabling precise communication about capabilities and requirements. When discussing educational technology initiatives, stakeholders should use level designations with care and consistency.

Regulatory Clarity: Precise level terminology streamlines governance discussions. For example: "Since learning graphs are part of a Level 2 textbook, we don't need regulatory approval from the student privacy review board." This clarity accelerates decision-making and reduces unnecessary compliance overhead for lower-level implementations.

The Level 2.99 Frontier: There remains substantial room for innovation at the boundary between Level 2 and Level 3—what we might informally call "Level 2.99." Consider a textbook that incorporates:

- **Knowledge graphs** storing concept dependencies and learning paths
- **In-browser graph algorithms** that recommend personalized learning sequences based on student-declared goals
- **Local storage** using browser-based mechanisms entirely controlled by each student

Such a system can provide sophisticated path recommendations without crossing the privacy threshold. The key distinction: these recommendations are computed locally, goals are not persistent across sessions (unless the student explicitly saves them), and no student-specific data resides on institutional servers.

Browser-Based Privacy Preservation: Modern web technologies—including IndexedDB, localStorage, and client-side JavaScript—enable surprisingly capable learning systems that keep all personalization data under student control. A student's browser becomes their personal learning record store, exportable and deletable at will. This architecture sidesteps many regulatory concerns while still enabling meaningful personalization within a session.

The 2.99 to 3.01 Transition: When students *do* want persistent, cross-device learning histories—and many will, for legitimate educational benefit—the challenge becomes crossing from Level 2.99 to Level 3.01 without data loss. Students who have built up valuable learning records in browser storage need a seamless path to server-side persistence when they're ready.

This transition requirement will drive a new industry of secure Software-as-a-Service Learning Record Store providers. These services must offer:

- Easy import of browser-based learning records
- Student-controlled data ownership and portability
- Institutional integration without institutional ownership
- Compliance with FERPA, COPPA, GDPR, and emerging state regulations
- Clear data retention and deletion policies

The organizations that build trusted bridges across the 2.99/3.01 boundary—enabling students to "graduate" their learning data from local to cloud storage on their own terms—will occupy a strategically valuable position in the educational technology ecosystem.

8.3 Level 2 Implementation

Level 2 represents an accessible entry point with significant benefits and minimal risk:

Technology Stack:

- Static site generators (MkDocs, Hugo, Jekyll)
- JavaScript simulation libraries (p5.js, D3.js)
- Learning graph visualization (vis.js, D3.js, Cytoscape.js)
- Embedded video platforms
- Optional analytics (privacy-respecting options like Plausible)

Cost Profile: Low. Open-source tools enable Level 2 implementation with minimal licensing costs. Primary investment is content development.

Timeline: Weeks to months for initial deployment; ongoing content enhancement.

8.4 Level 3 Implementation

Level 3 requires significant infrastructure investment:

Technology Stack:

- Learning Management System with adaptive capabilities
- Student authentication and identity management
- Learning Record Store (xAPI-compatible)
- Learning graphs (available from Level 2) with server-side traversal algorithms
- Assessment engine with item banking

Organizational Requirements:

- Privacy impact assessment
- Data governance policies
- Security infrastructure and monitoring
- Staff training on data handling

Cost Profile: Moderate to high. Licensing, infrastructure, and compliance costs accumulate.

8.5 Level 4 Implementation

Level 4 adds AI complexity:

Technology Stack:

- Large Language Model access (API or self-hosted)
- Vector database for retrieval augmentation
- Prompt engineering and guardrails
- Conversation logging and analysis
- Human oversight dashboard

Critical Considerations:

- **Accuracy**: LLMs can generate plausible but incorrect information. Retrieval-Augmented Generation (RAG) architectures constrain responses to verified content.
- **Cost management**: LLM API costs scale with usage. Tiered approaches using smaller models for simple queries can manage costs.
- **Privacy**: Conversation data may flow to third-party API providers. Self-hosted models or contractual protections may be necessary.

8.6 The Role of Human Instructors

While intelligent textbooks offer powerful capabilities, their effectiveness depends critically on the human instructors who deploy them. The most sophisticated adaptive system cannot replace the irreplaceable: a skilled educator who understands students as whole persons.

The Enduring Importance of Great Teachers: Research consistently demonstrates that teacher quality is the single most important school-based factor affecting student achievement. Excellent instructors inspire curiosity, build confidence, recognize when students are struggling emotionally as well as academically, and create the psychological safety necessary for intellectual risk-taking. No algorithm, regardless of sophistication, can fully replicate the human capacity to recognize a student's unspoken distress, to offer encouragement at precisely the right moment, or to model intellectual humility and passion for learning.

From "Sage on the Stage" to "Guide on the Side": The traditional model of direct instruction positions the teacher as the primary source of knowledge—the "Sage on the Stage" who delivers content while students passively receive. This model made sense when textbooks were static and information scarce. In an era of intelligent textbooks, however, the value proposition of instructors fundamentally shifts.

When Level 2–4 systems can deliver content, provide practice opportunities, answer questions, and adapt to individual pace, the instructor's comparative advantage moves from *content delivery* to *learning facilitation*. The modern instructor becomes the "Guide on the Side"—a mentor who:

- **Understands psychological needs**: Recognizes that learning is an emotional as well as cognitive process. Students arrive with anxiety, impostor syndrome, personal challenges, and varying degrees of confidence. Effective guides create environments where struggle is normalized and failure is framed as learning.
- **Provides metacognitive coaching**: Helps students develop awareness of their own learning processes. While an adaptive system might identify that a student struggles with a concept, a skilled instructor helps the student understand *why* they struggle and develop strategies for similar challenges.
- **Facilitates peer learning**: Creates opportunities for collaborative problem-solving and peer instruction that intelligent textbooks cannot orchestrate. Social learning remains a powerful complement to individual study.
- **Monitors holistic well-being**: Notices when a student's academic struggles signal deeper issues—mental health challenges, family difficulties, or identity development—and connects students with appropriate support.
- **Models intellectual virtues**: Demonstrates curiosity, persistence, humility, and the joy of discovery. Students learn not just content but *how to be learners* by observing skilled practitioners.

The Synergy of Human and Machine: The most effective implementations of intelligent textbooks leverage the complementary strengths of technology and human instructors. The textbook handles what it does well: consistent content delivery, unlimited patience for practice, immediate feedback, and data-driven adaptation. This frees instructors to focus on what they do uniquely well: building relationships, fostering motivation, addressing emotional barriers, and developing the whole person.

Institutions implementing higher-level intelligent textbooks should invest not only in technology but in developing instructors' facilitation skills. Professional development should emphasize coaching techniques, psychological awareness, and strategies for leveraging learning analytics to inform—not replace—human judgment.

The goal is not to eliminate instructors but to elevate their role from information transmitter to learning architect. In this vision, intelligent textbooks handle the "what" of learning while human guides attend to the "who" and "why."

8.7 Evaluation Criteria

When evaluating intelligent textbook products, institutions should assess:

1. **Claimed Level**: What level does the vendor claim? Is evidence provided?
2. **Data Requirements**: What student data is collected? Where is it stored? Who has access?
3. **Standards Compliance**: Does the product support xAPI, LTI, or other interoperability standards?
4. **Transparency**: Can instructors understand why the system makes specific recommendations?
5. **Exit Strategy**: Can student data be exported if the institution changes providers?

8.8 Measuring Success

Success metrics should align with level:

Table 7: Success Metrics by Level

Level	Key Metrics
1–2	Content quality; accessibility compliance; student satisfaction; usage analytics (if collected)
3	Learning outcome improvements; time-to-mastery; completion rates; prerequisite remediation effectiveness
4	Question resolution rate; conversation satisfaction; accuracy of responses; escalation to human rates
5	Comprehensive learning gains; transfer learning evidence; long-term retention; learner autonomy development

9 Discussion

9.1 Limitations of the Framework

Our five-level framework, like all classification systems, involves simplifications:

Discrete vs. Continuous: Real systems exist on a continuum. A textbook with sophisticated simulations but no adaptivity might be "Level 2.5." The discrete levels should be understood as reference points rather than rigid categories.

Domain Dependence: The optimal level may vary by subject matter. Procedural skills (mathematics, programming) may benefit more from Level 3 adaptivity than humanities subjects where multiple valid interpretations exist.

Cultural Context: Privacy expectations and regulatory environments vary globally. A system appropriate for one jurisdiction may be problematic in another.

Rapid Evolution: AI capabilities are advancing rapidly. Features currently requiring Level 4 infrastructure may become achievable at Level 2 costs within years.

9.2 Relationship to Learning Graphs

Our framework is closely related to the concept of learning graphs—structured representations of concept dependencies that enable navigation through educational content.

Learning graphs are a fundamental component of Level 2 intelligent textbooks. At this level, learning graphs serve as static knowledge structures that:

- Define prerequisite relationships between concepts
- Enable students to visualize the overall course structure
- Support non-linear navigation based on student interests
- Provide context for where each concept fits in the broader curriculum

The key distinction is that at Level 2, learning graphs inform *student-directed* exploration, while at Level 3 and above, they enable *system-directed* adaptation:

- **Level 2**: Learning graphs help students choose their own path; the system does not track which path was taken
- **Level 3**: Systems use learning graphs combined with individual performance data to sequence content and enforce prerequisites automatically
- **Level 4**: Systems use learning graphs to ground chatbot responses in verified concept relationships
- **Level 5**: Systems use learning graphs as the knowledge backbone for autonomous tutoring

The quality of the learning graph significantly impacts effectiveness at all levels. At Level 2, a well-structured graph helps students understand concept relationships; at Level 3+, it determines the quality of adaptive pathways. Poorly structured graphs lead to confusion at Level 2 and suboptimal learning paths at higher levels.

9.3 The Role of Human Instructors

Our framework should not be interpreted as advocating for replacement of human instructors. Rather, each level changes the instructor's role:

- **Level 1–2**: Instructor as content curator and discussion facilitator
- **Level 3**: Instructor as learning path supervisor and exception handler
- **Level 4**: Instructor as AI overseer and mentor for complex questions
- **Level 5**: Instructor as learning experience designer and pastoral supporter

Even at Level 5, human instructors remain essential for motivation, ethical guidance, and handling situations beyond AI capabilities.

9.4 Equity Considerations

Higher-level intelligent textbooks risk exacerbating educational inequities:

- **Access**: Level 3+ systems require reliable internet and modern devices, potentially disadvantaging students with limited technology access.
- **Data literacy**: Students must understand data collection to provide meaningful consent, requiring digital literacy education.
- **Algorithmic bias**: Systems trained on data from well-resourced institutions may perform poorly for underrepresented populations.

Institutions should assess equity implications before adoption and ensure lower-level alternatives remain available.

9.5 Future Research Directions

Several research directions emerge from this framework:

1. Empirical validation of level boundaries through user studies
2. Development of standardized assessment instruments for level classification
3. Privacy-preserving architectures for Level 3+ systems (federated learning, differential privacy)
4. Long-term learning outcome studies comparing levels
5. Cross-cultural adaptation of the framework

10 Conclusion

The rapid advancement of artificial intelligence is transforming educational content delivery. Without standardized frameworks for understanding these changes, educators, administrators, and policymakers struggle to make informed decisions about technology adoption.

This paper has proposed a five-level classification framework for intelligent textbooks, inspired by the SAE J3016 standard that successfully brought clarity to autonomous vehicle discourse. Our framework defines clear progression from static textbooks (Level 1) through interactive content (Level 2), adaptive systems (Level 3), chatbot integration (Level 4), to autonomous AI tutoring (Level 5).

A critical contribution of our analysis is the identification of Level 3 as a privacy inflection point. Below this threshold, intelligent textbooks can deliver significant educational value with minimal student data collection. At Level 3 and above, personalization requires increasingly detailed individual tracking, triggering regulatory obligations and institutional responsibilities that must be addressed proactively.

We have examined how existing educational technology standards—including xAPI, Learning Record Stores, and IEEE specifications—can provide interoperability foundations while supporting privacy through data portability and student control. However, gaps remain, particularly for conversational AI data and privacy-preserving computation.

The framework enables differentiated governance, with regulatory and institutional requirements scaled to the data intensity and AI autonomy of each level. This graduated approach avoids both over-regulation that stifles beneficial innovation and under-regulation that fails to protect student privacy.

As AI capabilities continue their exponential growth, the educational technology landscape will continue to evolve. The five-level framework provides a stable vocabulary for navigating this evolution—enabling clear communication, informed decision-making, and responsible innovation.

We invite the educational technology community to adopt, refine, and extend this framework. Just as SAE J3016 evolved through multiple revisions in response to technological and regulatory developments, we expect this framework to evolve as intelligent textbooks mature. The goal is not a permanent taxonomy but a useful starting point for the conversations that will shape the future of education.

> "All models are wrong, but some are useful." — George Box

We hope this model proves useful.

Acknowledgments

The author would like to thank the educational technology research community and the developers of open educational resources whose work has informed this framework.

References

[1] D. McCreary, "Five levels of intelligent textbooks." Medium, `https://dmccreary.medium.com/five-levels-of-intelligent-textbooks-b81a4c1525a0`, 2024. Introduces the five-level classification framework for intelligent textbooks.

[2] METR, "Measuring ai ability to complete long tasks." `https://metr.org/blog/2025-03-19-measuring-ai-ability-to-complete-long-tasks/`, March 2025. Model Evaluation and Threat Research.

[3] B. P. Woolf, *Building Intelligent Interactive Tutors: Student-centered Strategies for Revolutionizing E-learning.* Morgan Kaufmann, 2010.

[4] V. Aleven, E. A. McLaughlin, R. A. Glenn, and K. R. Koedinger, "Instruction based on adaptive learning technologies," in *Handbook of Research on Learning and Instruction*, pp. 522–560, Routledge, 2016.

[5] J. C. Harmes, J. L. Welsh, and R. J. Winkelman, "The technology integration matrix: Development and evolution of a research framework," *Journal of Research on Technology in Education*, vol. 48, no. 2, pp. 88–102, 2016.

[6] R. R. Puentedura, "Transformation, technology, and education." `http://hippasus.com/resources/tte/`, 2006.

[7] W. Holmes, M. Bialik, and C. Fadel, *Artificial Intelligence in Education: Promises and Implications for Teaching and Learning.* Center for Curriculum Redesign, 2019.

[8] R. Luckin, W. Holmes, M. Griffiths, and L. B. Forcier, *Intelligence Unleashed: An Argument for AI in Education.* Pearson Education, 2016.

[9] J. Haupt, Q. Lu, Y. Shen, J. Chen, Y. Dong, D. McCreary, M. Akçakaya, and G. B. Giannakis, "Deploying ai for signal processing education: Selected challenges and intriguing opportunities," *arXiv preprint arXiv:2509.08950*, 2025. Discusses smart textbooks, LLM integration, and AI deployment challenges in education.

[10] S. Slade and P. Prinsloo, "Learning analytics: Ethical issues and dilemmas," *American Behavioral Scientist*, vol. 57, no. 10, pp. 1510–1529, 2013.

[11] A. Rubel and K. M. Jones, "Student privacy in learning analytics: An information ethics perspective," *The Information Society*, vol. 32, no. 2, pp. 143–159, 2016.

[12] V. Lockhart, D. McCreary, and T. A. Peterson, "Microsims: A framework for ai-generated, scalable educational simulations with universal embedding and adaptive learning support," *arXiv preprint arXiv:2511.19864*, 2025. Demonstrates 30-40% learning gains from interactive simulations; framework for AI-assisted simulation creation.