

## Case Study

# Assessing and Evaluating AI Tools and Features for K-12 Learning

Magic  
EdTech

### The Client

The client is a leading provider of accredited, tuition-free online public schools across multiple US states. They offer a comprehensive, standards-aligned curriculum through an interactive platform with live classes, one-on-one sessions with certified teachers, and self-paced learning, all while emphasizing strong teacher-student interactions and support services.

### The Challenge

The client was launching two high-profile AI features: a custom assessment generator for teachers and an AI 'tutor' for students. They needed experienced educators and subject matter experts in ELA, biology, and world history to assess the new features' performance and the accuracy, depth, tone, and appropriateness of their outputs, ensuring alignment with their focus on strong teacher-student interactions.

### Critical Success Parameters

- ✓ Validate the custom assessment generator for ELA teachers produces accurate, unbiased test questions in various formats (multiple choice, multiple select, true/false, etc.), with plausible distractors and proper Depth of Knowledge calibration.
- ✓ Ensure the AI tutor for world history and biology answers all student questions accurately, at the appropriate level, and in a supportive, approachable, and professional tone.
- ✓ Utilize the validity and persuasiveness of SME evaluations and feedback as key metrics for the company's "go/no go" decision.

### Our Approach

- ✓ Assigned three SMEs from our content development/curriculum team (former educators, writers, and subject matter experts all) —one each for ELA, world history, and biology to the project.
- ✓ Conducted daily one-hour meetings with the client's development and QA teams to review feature status and share SME feedback.
- ✓ Governed each day's SME evaluations with key questions on feature performance and used meetings to set agendas and distribute assignments.



### Key Result Highlights

Assigned **3** subject matter experts/former educators to evaluate ELA, world history, and biology features.

Tested **2** AI-driven features: a teacher assessment generator and a student "tutor."

Facilitated the client's **1** go/no-go decision with Magic's feedback.

Analyzed **100s** of "tutor" responses and custom ELA questions.