

The educator's guide to **evaluating AI**

Balancing cost, safety and educational impact

Contents

- Introduction..... 3
- Five areas to consider when choosing AI solutions in education..... 4
 - Educational Efficacy - Ensuring high-quality, reliable outputs..... 4
 - Flexibility and breadth of capabilities that enhance learning..... 5
 - Enabling an iterative feedback loop 5
 - Handling student data and privacy 6
 - Pricing and educational value 6
- Conclusion.....7

Introduction

There's a lot of buzz about the enormous and compelling promise of AI in education. From supporting teachers by automating administrative tasks to creating personalised learning journeys or generating insights from predictive analytics, AI's ability to complement rather than replace conventional teaching techniques has truly captured the attention of the education community.

Currently, free AI tools are readily available and widely used by teachers for various tasks, such as Chat GPT or AI features embedded in search engines. Paid solutions sit at the other end of the scale and offer more refined and education-specific functionalities.

When evaluating AI tools for education, it's essential to be clear on the outcomes you're looking for and how to choose the right AI tool to achieve that outcome. In particular, privacy and security should be top of mind. You need to know that the AI tool you choose won't do harm, introduce bias, misinform students, or misuse student data. This guide will help.

Not all AI tools are created equal, and that's where some definitions can help

- **Natural Language Processing (NLP)** focuses on the interaction between computers and human language. It's used for tasks such as language translation, sentiment analysis, and text summarisation.
- **Machine Learning** is the broader concept underpinning much of AI, where systems improve their performance on a task over time by learning from data.
- **Generative AI** refers to algorithms capable of creating new content, such as text, images, or music, based on the data they've been trained on.
- **GPT (Generative Pre-trained Transformers)** if you've ever wondered what the GPT in ChatGPT actually means, it refers to a breakthrough in AI research that gives applications the ability to create human-like text and content (images, music, etc) while also answering questions in a conversational manner.

- **Large Language Models (LLMs)** are a type of advanced NLP that can combine multiple learning models and vast amounts of data to suit a specific training goal. For example, different types of LLM can focus on language, sentiment analysis, answering questions or summarising – all valuable capabilities in the education space.
- **Small Language Models (SLMs)** are similar to LLMs, but as the name implies, they require less training data and are typically suited for single, specific tasks. e.g. answering customer questions about a specific product. Fun fact, anything with less than 30 billion pieces of data (or parameters) is considered an SLM, while recent LLMs have clocked in at 2 trillion parameters. For this reason, LLMs consume a lot more computing resources than SLMs, making them more expensive to run but also more adept at complex, sophisticated tasks.

We know first-hand that creating effective AI tools requires extensive research, development, and training data. That time and effort directly correlate to the differences in functionality and effectiveness of free and paid solutions. While there are always exceptions, AI tools at the free end of the scale generally use SLMs and focus on more straightforward capabilities, while paid AI tools use LLMs and encompass quite advanced capabilities that can handle work with greater nuance and variability.

In other words, the more advanced the AI, the greater the investment to ensure it can deliver precise, safe, reliable outcomes.

With this in mind, we have developed a mini buyer's guide to help educators evaluate AI offerings in education, ensuring they choose AI tools that align with their specific needs, goals, and budgets.

Five areas to consider when choosing AI solutions in education

AI frameworks in education should be built on five interlinked drivers that guide their practical use and implementation. By considering these factors, educators can better understand the essential and emerging features that can be leveraged to enhance teaching and learning experiences.



Educational Efficacy - Ensuring high-quality, reliable outputs

The first point may seem obvious, but we all tend to get distracted by flashy/showy features, even if they don't add any real value to the student learning experience. You must be clear about your use case and the AI tool's method of implementation to ensure the system actually supports student learning.

This is where the choice of LLM or SLM as the underlying model is critical. Unlike smaller models, advanced LLMs are trained on vast amounts of data. This breadth of data allows them to understand and interpret complex, open-ended questions and provide accurate and contextually relevant feedback to the specific content and learning objectives - helping students grasp underlying concepts rather than just surface-level correctness.

For example, if a student writes an essay that contradicts the theme of a book like "1984" by George Orwell - perhaps suggesting it's an uplifting and positive story - an advanced LLM can detect this mismatch. It recognises the broader context of the book, providing feedback that corrects the misunderstanding. This is something less sophisticated models would likely miss entirely.

Advanced AI tools in education often include additional scaffolding and processing layers that guide the AI's decision-making process. By incorporating these scaffolds, AI can provide additional insightful, educationally valuable feedback that pushes students to think more deeply.



What to look for:

Schools should prioritise tools that use advanced LLMs capable of delivering accurate, contextually relevant feedback. Look for AI systems with scaffolding features so students can engage in deeper learning and critical thinking. The goal should be to choose tools that support primary learning outcomes and a richer, more nuanced understanding of the subject matter.



Flexibility and breadth of capabilities that enhance learning

Advanced LLM-based AI systems also offer remarkable flexibility. They can handle a wide range of question types, from simple multiple-choice to complex, open-ended tasks. In educational settings, where diverse question types are used to assess various levels of student understanding and critical thinking, a single system with many capabilities eliminates the need to manage and toggle through multiple systems.

Whereas basic or freemium models are usually limited to straightforward questions, advanced LLMs can engage with more complex prompts that require higher-order thinking and analysis. For example, comprehensive AI systems can evaluate a student's response to an open-ended essay question, assess the coherence of their argument, and provide nuanced feedback that helps refine their reasoning.

What to look for:

Look for AI systems capable of providing meaningful feedback on complex, open-ended questions and those that can support higher-order thinking and critical analysis. Regardless of which model you choose, always remember that AI isn't a "set-and-forget" replacement for the skill of the teacher. AI helps teachers distribute the load and scale their efforts, but there should be human oversight at some point to correct misunderstandings and ensure learning stays on track. Tools that incorporate some form of student feedback on the analysis and provide visibility on these responses make it easier for teachers to intervene and close that feedback loop.



Enabling an iterative feedback loop

An iterative feedback loop is a powerful feature of advanced AI systems, allowing students to revise their work based on specific, actionable insights. Each round of feedback builds on the previous one, helping students gradually enhance their skills and knowledge.

In contrast to one-off feedback systems that leave students with little opportunity to improve, this approach more closely mirrors natural learning. By revising their work in response to detailed feedback, students can make more meaningful progress with each iteration and never miss the deeper learning that comes from continuous refinement.

What to look for:

Avoid tools that offer one-off feedback, as they limit opportunities for deeper learning and refinement. The ideal AI system should foster a growth-oriented learning environment, encouraging students to engage in a cycle of feedback and improvement.



Handling student data and privacy

All AI models are trained on datasets, but for education there are obvious concerns around how student data is used, especially regarding consent and privacy. Within educational platforms, data often includes student responses and interactions, and so in some cases, student data may be automatically included in training datasets. To manage this, there would need to be explicit knowledge and consent from students and/or their guardians across the different use cases.

There is also data sovereignty - the idea that data should be stored and processed within the same country as the students - to consider. In Australia, student data should be stored within Australian borders rather than being sent overseas. This reduces the risk of it being subject to foreign laws or less stringent privacy protections and provides an additional layer of security and compliance with local regulations.

What to look for:

Given the real implications for trust and safety in educational environments, schools and educators must be vigilant about how AI tools handle their data. They should demand transparency from providers about their data privacy practices to ensure stringent alignment with existing school policies and national standards on data protection.



Pricing and educational value

While 'free' AI tools may seem like the most appealing route, they come with trade-offs and compromises like smaller datasets, less sophisticated algorithms, and limited processing power. This translates to a lower depth and quality of educational support along with other hidden costs, such as the potential for student data to be used without consent for further model training.

Advanced AI systems demand vast amounts of data for training, robust computational power, and continual updates to maintain their effectiveness. These factors contribute to the higher costs of premium AI solutions but also mean they can provide accurate, relevant, incremental, and contextually appropriate feedback to students.

What to look for:

Schools should carefully evaluate the value proposition of both free and premium offerings. Look for transparency in pricing and a clear explanation of what you are paying for - whether that's advanced capabilities, reliability, or stronger data privacy protections. Schools should be particularly cautious of free tools that cut corners on quality, or use student data in ways that compromise privacy. The right AI tool should strike a balance between cost and the quality of educational outcomes it delivers, ensuring that it meets the specific needs of your institution.

Conclusion

The majority of teachers care most about whether AI is safe and genuinely helps students learn, so selecting the right tool is not just about cost or features; it's about understanding and quantifying the value it brings to the educational environment through high-quality, iterative learning tools.

This is why it's so important that the evaluation process prioritises educational efficacy, safety, privacy and long-term student outcomes over just the upfront costs. This importance is being recognised in the development of accreditations such as the ST4S (Safer Technologies for Tools) Readiness Check in Australia and NZ.



While this is a self-assessment tool for suppliers to check how their product is doing against the nationally agreed privacy and security assessment framework for K-12 education, it can also be another helpful lens in conjunction with the contents of this guide when doing your own evaluations.

When assessing options, ask yourself:

1. How does the AI tool ensure high-quality responses to minimise misinformation and biases?
2. Does it provide contextually relevant feedback?
3. Does the AI system offer flexibility in handling a wide range of question types, including open-ended and complex tasks?
4. Is there an iterative feedback loop that allows for continuous student improvement?
5. How does the tool handle student data and ensure privacy? Is data stored locally in compliance with national regulations?
6. What is the true cost of the AI tool, and how does it balance cost with the quality of educational outcomes?

Education Perfect meets the criteria outlined in this guide by providing a comprehensive suite of AI tools designed to enhance student learning, ensure data privacy, and deliver consistent, high-quality outputs. By singularly focusing on educational outcomes, and working back from there, Education Perfect's AI solutions have been carefully developed to align closely with school's goals and provide the best possible support for their students and teachers.

Get in touch

🌐 educationperfect.com **in** [educationperfect](https://educationperfect.com)

f [educationperfectnz](https://www.facebook.com/educationperfectnz) **@** [educationperfect](https://www.instagram.com/educationperfect)

X [epforlearning](https://twitter.com/epforlearning)

