Guidelines for EdTech Product Developers for AIEd Development and Deployment in Singapore

MOE'S APPROACH TO AI IN EDUCATION

1. MOE's approach to AI in education puts students in the centre and pedagogy at the fore. MOE emphasises responsible and age- and developmentally- appropriate use of AI products in teaching and learning as students learn about AI, learn to use AI, learn with AI and learn beyond AI.

OVERVIEW OF GUIDELINES

2. This set of guidelines is applicable for AI products that seek to be:
   a. **whitelisted for schools' use**, which allows SLS users to more seamlessly access the AI product as an external resource for teaching and learning, or to embed it into a module as an inline frame (iframe);
   b. **integrated with SLS**, which entails single sign-on (SSO) access to the product as a feature of SLS and data sharing via an API connection.

3. These guidelines[1] are intended to complement AI governance regulations and guidelines across the broader governance landscape, such as the Personal Data Protection Act, the Personal Data Protection Commission's advisory guidelines on use of personal data in AI recommendation and decision systems, the Generative AI Model Governance Framework by IMDA and other guidelines for industry developers.

4. These guidelines are not intended to function as a checklist where all criteria must be met in full. Developers are encouraged to consider these guidelines holistically and to demonstrate how their product addresses the considerations most relevant to their use case and product category.

5. A range of AI products are available in general education. As these tools serve different purposes and utilise varying AI

[1] Key documents referenced in the development of these guidelines include the UNICEF Learning Cabinet's Edtech for Good Framework, the UK DfE's Generative AI Product Safety Standards, and the ISTE Teacher Ready Edtech Evaluation Tool.

architectures and systems, their associated risks differ accordingly. These tools can be grouped into four categories, and this set of guidelines apply to all four categories.

Category	**Notes**
Non-generative *general-purpose* AI tools, such as text-to-speech and speech-to-text systems	Predictive or descriptive AI. Used in education but could also be used for any other purpose.
Non-generative *educational* AI tools, including intelligent tutoring systems or early warning systems	Developed for educational purposes. Usually utilised for prediction or assessment.
Generative *general-purpose* AI tools, e.g. LLMs such as ChatGPT	AI that can generate new outputs. Developed for general use and might lack educational safeguards.
Generative *educational* AI tools, e.g. Khanmigo or the Socratic Playground	Often contains educational safeguards such as configurations to refrain from giving solutions and support learners to obtain the solution on their own.

6. All AIEd products should avoid the following:
   a. Use of AI for companionship or to substitute human relationships and interactions *(e.g. AI personas designed to form relational bonds with students)*
   b. Use of AI to provide socio-emotional support
   c. Use of AI for therapy
   d. Use of AI for student well-being matters

7. For these guidelines,
   a. an "AI tool" or "AI product" is any software, feature or service that uses a large language or AI model to infer from data and produce content, predictions, classifications, recommendations, scores or decisions that support or influence an educational task or interaction. This includes student-facing and teacher-facing tools, embedded or back-end AI features, agentic functions and AI accessed through third-party APIs. It includes both generative and non-generative AI, whether developed specifically for education or for general use.
   b. "AIEd product" refers to an AI tool used for teaching, learning, assessment, learner support or educational decision-making.
   c. "AI output" includes content, feedback, scores, predictions, recommendations, learning pathways, alerts or actions produced or materially influenced by the AI system.

For clarification or feedback on these guidelines, please reach out via email to SLS_Partnerships@moe.gov.sg.

QUALITY AIED PRODUCTS SHOULD EMBODY THE FOLLOWING CONSIDERATIONS

Elaboration for each guideline can be accessed by clicking on the icon next to the guideline.

<table>
<tr><td rowspan="3">Guidelines for Selection and Use</td><td colspan="2">Overarching Considerations</td></tr>
<tr><td colspan="2">[O1] Represent Singapore's diverse cultural contexts and ethnicities, and align to MOE's mission and desired outcomes of education
[O2] Able to be accessed by all learners across socioeconomic backgrounds or languages
[O3] Able to be used by and to benefit diverse groups of learners
[O4] Designed according to age-appropriate guidelines aligned to MOE recommendations on AI/EdTech use and exposure, based on what is appropriate for each developmental stage
[O5] Provide teacher-in-the-loop or teacher-over-the-loop features or functionalities</td></tr>
<tr><td>Pedagogical Considerations</td><td>Safety Considerations</td></tr>
<tr><td>Additional Guidelines for Deployment (i.e. through SLS integration)</td><td>[P1] Preserve students' and teachers' choice and control over important personal and educational decisions
[P2] Adopt evidence-based design approaches to ensure educational value
[P3] Evaluate products with clear metrics for learning effectiveness or learning outcomes</td><td>[E1] Establish clear protocols to ensure data privacy and security
[E2] Put in place safeguards to protect learners' privacy and well-being
[E3] Explain transparently when and where AI (including agents or back-end decision-making AI systems) is used, and how these systems operate (i.e. data used, trustworthiness)
[E4] Demonstrate a high standard of accuracy and as little bias as possible to avoid outcomes which run counter to MOE values and MOE's position on AI use in Education</td></tr>
</table>

RATIONALE FOR EACH GROUP OF CONSIDERATIONS

Overarching Considerations

These considerations were designed to address the central question of how AIEd can be integrated in ways that respond to diverse local needs and contribute to more inclusive teaching and learning practices. They are aligned with the MOE AIEd Framework, which builds on IMDA's Model AI Governance Framework to address education-specific use and risks, and incorporates key beliefs from the Ethos of the Teaching Profession.

Pedagogical Considerations

These considerations were intended to support design, development and customisation of AI tools in ways that are aligned with good pedagogy and ensure due consideration given to the roles of teachers and students when teaching and learning with AI tools. They were developed with reference to learning sciences and instructional sciences research evidence while incorporating local guidance including the Singapore Teaching Practice (STP) and e-Pedagogy.

Safety Considerations

These considerations were crafted following a review of existing regulatory frameworks such as the PDPA, government mandates (i.e. the PMO circular), and safety guidelines for AI product development in Singapore. Where possible, the original language and intent of these documents were retained. The examples raised were chosen to reflect the safety considerations most relevant and necessary for AIEd use cases.

Elaboration of Guidelines for AIEd Development and Deployment in Singapore

[O1] Represent Singapore's diverse cultural contexts and align to MOE's mission and desired outcomes of education

Description	Examples (non-exhaustive)	Useful References
[O1.1] Fair and appropriate representation of Singapore's diverse cultural contexts and ethnicities	• Learning materials generated can be prompted to reflect local culture and traditions • AI output (including generated learning materials) reflects Singapore's local culture, traditions and norms. • AI output demonstrates sensitivity towards Singapore's diverse ethnicities and religious practices and avoids content that could be considered disrespectful or inappropriate in a multi-religious context	1. Overview of Singapore's Education System 2. MOE Desired Outcomes of Education 3. Singapore Curriculum Philosophy 4. Subjects and syllabuses in Singapore classrooms
[O1.2] Ensure alignment to local curriculum and MOE's desired outcomes of education, i.e., • aligned to Singapore's subject syllabuses for academic subject areas • aligned to social emotional competencies and values for non-academic subject areas	• For adaptive learning systems, there is flexibility by system users to adjust learning content to be mapped to the learning progression and learning outcomes in the syllabus documents • Ability to adjust and use knowledge bases that are aligned to Singapore's syllabuses • Follow an active learning approach to construct knowledge, encourage students' holistic development, and promote learning beyond grades	
[O1.3] Able to provide a solution to target a local pain point or improve an existing issue within Singapore's education landscape	• Support low progress learners as understood by Singaporean educators • Enable development of 21CC as defined by MOE • Provide enrichment or extension pathways to stretch high-ability learners	

[O2] Able to be accessed by all learners

Description	Examples (non-exhaustive)	Useful References
[O2.1] Ensure equitable access to digital tools through reasonable pricing	• Provide different pricing levels (including a free tier) to allow learners from all socio-economic backgrounds to access the product	NIL
[O2.2] Enable access and use of the tool in Singapore's four official languages, where appropriate	• Ensure non-English languages used in schools are sufficiently represented and provided for at interface level, where appropriate • Design for multi-language user interaction and experience • Be transparent about variations in output quality and limitations for AI-generated content in Mother Tongue languages	

[O3] Able to be used by and to benefit diverse groups of learners

Description	Examples (non-exhaustive)	Useful References
[O3.1] Cater for learner variability in Singapore mainstream schools	• Works towards WCAG 2.1 standards to support users with additional learning needs, such as by embedding accessibility features such as text-to-speech and visual mind-mapping features to support users with additional learning needs like dyslexia or mild autism spectrum disorder • Ability for users to adjust output to cater to different ability levels	1. WCAG 2.1 Standards 2. Accessibility Checklist by GovTech
[O3.2] Ensure that students are not disadvantaged or sidelined based on characteristics like race, gender, disability or socioeconomic background	• Product does not make recommendations or change its output qualitatively based on race, gender, religion, or socioeconomic background • Product output does not reinforce stereotypes • Data collected does NOT include sensitive personal data such as race	NIL
[O3.3] Strive towards equitable impact of product use on all learners	• Adopt universal and participatory design approaches so that no group of learners is disadvantaged or sidelined based on characteristics like race, gender, disability or socioeconomic background	NIL

[O4] Designed according to age-appropriate guidelines or follows MOE recommendations on AI/EdTech use and exposure according to the appropriate developmental stages

Description	Examples (non-exhaustive)	Useful References
[O4.1] Ensure that AI features are designed according to learner developmental readiness and that their use will not negatively impact development (i.e. cognitive, social, emotional, executive function) in young learners	• Shows clear age limits; AI use is pegged to maturity or student readiness levels • Where younger students are allowed to use the product, there are distinctions in access to AI functions or restricted AI functions (i.e. through age-gated content, or a tiered permission system that unlocks AI features as students progress through age bands or year levels).	1. MOE's Parent Infographic: "AI and your child" (see p.3 for age-appropriate use)

[O5] Incorporate or provide teacher-in-the-loop or teacher-over-the-loop features or functionalities

Description	Examples (non-exhaustive)	Useful References
[O5.1] Ensure that teachers are involved in or able to review and edit AI decision-making and interactions with students	• AI outputs can be reviewed, edited or overridden by teachers *before* it is shown to students, or *during* the student-AI interaction • AI outputs are contextualisable by teachers • Teachers can review student-AI interactions post-hoc	NIL
[O5.2] Include features that aid or allow teachers to evaluate AI output for trustworthiness/ reliability	• Provide uncertainty or confidence indicators (i.e. classification confidence scores) to show teachers (and students) the likely trustworthiness of the AI output	NIL

[P1] Preserve students' and teachers' choice and control over important personal and professional decisions

Description	Examples (non-exhaustive)	Useful References
[P1.1] Maintain the decision-making authority of users over what, when or how they wish to learn or teach	• Input-seeking process features and overriding mechanisms that prioritise human involvement • Include technical mechanisms to disable or bypass AI features at teacher or school level	1. Hattie, J. (2023). *Visible learning: The sequel: A synthesis of over 2,100 meta-analyses relating to achievement*. Routledge. https://doi.org/10.4324/9781003380542
[P1.2] Support users in exercising sound judgement on how to use the AI product in an educationally meaningful manner	• Present relevant learning data or feedback in ways that can be easily understood • Offer explanations of how AI output is achieved to allow users to decide if and how to best utilise these outputs • Offer teacher training sessions or use cases to help users utilise their product in educationally sound ways • Appreciates and attempts to accommodate different readiness to learn, e.g., low/ mid/ high-progress students, or students of varying mastery	
[P1.3] Anchor AI features and product design on evidence-based pedagogical models that promote deep learning and 21st Century Competencies including social and emotional competencies and critical and independent thinking	• Integrate reflection checkpoints into the learning pathway or other relevant and evidence-based features to facilitate metacognition • Design AI scaffolds that progressively fade as learner competence grows • Exclude AI functionalities and scaffolds designed to create dependency or habitual reliance upon technology • Appreciate and attempt to accommodate different readiness to learn, e.g., low/ mid/ high-progress students, or students of varying mastery • Product incorporates a specific pedagogical model to train, prompt or scaffold the AI (on content design, facilitation, scaffolds, analytics, mastery). It is not a problem solver and avoids giving direct answers. • Product is anchored on or include features based on any of the following pedagogical models: ○ Purposeful engagement ○ Constructivism ○ Gamification ○ Collaboration ○ Feedback models ○ Multimodality	1. The Singapore Teaching Practice 2. OECD TALIS Report on Singapore (2024) 3. Brookings Institution Report: "A new direction for students in an AI world: Prosper, prepare, protect"

[P2] Use relevant research findings to guide product design to ensure educational value (Additional Guideline for Integration)

Description	Examples (non-exhaustive)	Useful References
[P2.1] Promote design and customisation of learning activities in ways that align with research-based best practices rooted in learning sciences	• Include recall tasks to facilitate knowledge retention • Create opportunities for reflection to encourage metacognition	1. Koedinger, K. R., Booth, J. L., & Klahr, D. (2013). Instructional complexity and the science to constrain it. *Science*, *342*(6161), 935-937.
[P2.2] Document the educational value of the product, including the learning theories and learning sciences research findings behind its design	• Provide evidence of educational value, which may take various forms depending on the developer's stage of development, including published research, internal pilot data, educator feedback, or third-party evaluations • Produce whitepapers that show evidence of efficacy as well as underpinning research and design principles • Iterate design through feedback, data and (educator) collaboration	

[P3] Evaluate products with clear metrics for learning effectiveness or learning outcomes (Additional Guideline for Integration)

Description	Examples (non-exhaustive)	Useful References
[P3.1] Provide an overview of how various product features are designed to achieve identified learning objectives, and how impact is measured	• Articulates impact measurement through theory of change, logic model or whitepaper • Provide impact analysis through pre/post assessments, usage analytics tied to learning outcomes, or teacher-reported observations	1. EdTech for Good Framework: Evaluation Criteria (p.13-21)
[P3.2] Provide documented evidence that demonstrates that your product leads to improved intended learning outcomes that the tool is designed for	• Articulates clear learning outcomes, which can be in the form of knowledge or skills acquisition • Obtains third party evaluations or certifications, such as International Society for Technology in Education (ISTE) SEAL	

[E1] Establish clear protocols to ensure data privacy and security

**: Examples have been adapted to MOE's context and regulations*
^: Examples/Descriptions have been adopted verbatim from relevant WOG guidelines

Description	Examples (non-exhaustive)	Useful References
[E1.1] Comply with Singapore's data privacy regulations and best practices, i.e. Personal Data Protection Act (PDPA) and Singapore Standard for Data Protection, and take measures to protect and keep user data secure	• *Apply stringent standards of protection for students' personal data, collecting, using and disclosing it only for appropriate, notified educational purposes with valid parental/school consent where required, and avoiding harmful or inappropriate profiling. ^Ensure that policies are in place to manage account and access control, backup and retention, and password security • ^Maintain accountability for data breaches by informing relevant parties • ^Put de-identification and other relevant processes in place to protect and keep user data anonymous • ^API services or self-hosted models provided must comply with the hosting and infrastructure requirements set out for the chosen mode of deployment	1. Advisory guidelines on the PDPA for children's personal data in the digital environment 2. PDPC | Advisory Guidelines on use of Personal Data in AI Recommendation and Decision Systems 3. Singapore Standards for Data Protection
[E1.2] State the purposes and limits of data collection, particularly unobtrusive data collection	• *Explanations are easy to find and presented in plain language to help users (both students and teachers) understand how they may be tracked or measured and how this data is used to support teaching and learning	
[E1.3] Ensure that data retention, assessment, and use for AI training is transparent and abides by contractual agreements and/or terms of use	• *Do not store user data without user permission or knowledge • ^Do not use user data without user permission or knowledge for training, or for auxiliary profit-making purposes, such as to be sold or shared beyond the user community • ^Disclose all third-party engagement • *Share definition of data ownership and portability	

[E1] Establish clear protocols to ensure data privacy and security (cont'd)

**: Examples have been adapted to MOE's context and regulations*
^: Examples/Descriptions have been adopted verbatim from relevant WOG guidelines

Description	• Examples (non-exhaustive)	Useful References
[E1.4] Take measures to process data and/or mitigate outputs that reproduce particular or culturally/contextually inappropriate training data	• *Review training data for information that could serve violent or terrorist incidents, or contravene intellectual property rights, and where appropriate, remove it. • *Implement measures to prevent, flag, or take other action in response to objectionable outputs, i.e. plagiarized, licensed or copyrighted content, or material that contravenes MOE values or age guidelines)	See above

[E2] Put in place safeguards to ensure product safety and reliability, and protect learners' overall well-being

**: Examples have been adapted to MOE's context and regulations*
^: Examples/Descriptions have been adopted verbatim from relevant WOG guidelines

Description	Examples (non-exhaustive)	Useful References
[E2.1] Ensure transparency around use of AI and AI interactions, maintain clear accountability and risk management procedures	• ^Embed visual cues to help users identify when they are interacting with AI • ^Provide adequate opportunities for users to provide feedback especially when the AI system does not perform as expected • ^Make clear at which point and how AI is used for decision making during the workflow/product pipeline	1. AI Verify Testing framework for Traditional and Generative AI 2. IMDA's Model AI Governance Framework (2020) 3. IMDA's Starter Kit for Testing LLM-Based Applications for Safety and Reliability
[E2.2] Ensure safe user interactions with the AI system, i.e. by embedding guardrails such as content filtering or monitoring mechanisms	• ^Block inappropriate or harmful content for children and youths (< 18) • ^Identify and flag potentially sensitive or harmful information disclosures from students (*e.g. expressions of distress, self-harm, illegal actions that could harm others, or abuse*) • *Put into place measures to ensure that students do not form a dependency upon the tool that could be developmentally detrimental or crippling (*e.g. design features that encourage independent attempts before AI assistance; set usage limits; include teacher dashboards that surface patterns of over-reliance for follow-up*)	
[E2.3] Conduct safety testing and evaluations (e.g. hallucination, bias, data leakage, undesirable content, vulnerability to adversarial prompts)	• ^Share results of safety performance of AI model/tool against appropriate benchmarks (i.e. child-specific safety benchmarks), especially for tools that interact directly and at length with students • ^Regular monitoring of how the tool stands against adversarial prompting, red-teaming etc.	

[E3] Explain transparently when and where AI (including agents or back-end decision-making AI systems) is used, and how these systems operate (i.e. data used, trustworthiness), and the limitations and risks of these systems (Additional Guideline for Integration)

**: Examples have been adapted to MOE's context and regulations*
^: Examples/Descriptions have been adopted verbatim from relevant WOG guidelines

Description	Examples (non-exhaustive)	Useful References
[E3.1] Ensure that users are sufficiently informed of residual risks and limitations of generative AI, and guide the appropriate use of their applications accordingly	• *Make explicit how AIEd system outputs are generated by explaining the logic behind the creation of learning pathway recommendations and learner feedback etc. • ^Provide clear Terms of Use and privacy statements (with easily interpretable explanations of AI limitations, residual risks, and user responsibilities) accessible from the main interface, with mandatory acknowledgement before access is granted. • ^Offer educational resources such as guides, playbooks, or workshops to promote responsible use and best practices.	1. AI Verify Testing framework for Traditional and Generative AI 2. IMDA's Model AI Governance Framework for Agentic AI (for guidance on enabling end user responsibility for agent use via transparency, see section 2.4, p 22) 3. IMDA's Transparency Guidelines
[E3.2] ^Provide voluntary disclosures about steps taken to mitigate the risks of AI in the product design and development	• Explain if and how human-in-the-loop processes have been implemented during the AI lifecycle • Disclose evidence of implicit/explicit bias and limitations around bias management to teachers and deployers	
[E3.3] ^Make known key information about AIEd system to enable deployers and/or users to interpret system outputs and enable them to use it appropriately	• Data used (in the context of how user data will be used and protected) • Model properties (LLM model and any changes from previously specified models) • System architecture (i.e. where and at what points AI is being used) • Capabilities and limitations • Evaluation results	

[E4] Demonstrate a high standard of accuracy and bias management (Additional Guideline for Integration)

**: Examples have been adapted to MOE's context and regulations*
^: Examples/Descriptions have been adopted verbatim from relevant WOG guidelines

Description	Examples (non-exhaustive)	Useful References
[E4.1] Collect and train AI on sufficiently representative data e.g., well-labelled and carefully curated local (and/or contextually relevant) data is used to train the AI ^	• *Use education-specific datasets to provide comprehensive training of large models for education and lay a foundation for more relevant and personalised outputs in various aspects of education • ^Undertake data quality control measures including annotating training datasets consistently and accurately, and using data analysis tools to facilitate data cleaning (e.g., debiasing and removing inappropriate content)	1. MOE Position Statement on AI Use in Education 2. Examples of different types of education-specific datasets (see pages 7 to 10 of Large Model for Education – Overall Reference Framework Alliance Standard) 3. IMDA's Starter Kit for Testing LLM-Based Applications for Safety and Reliability (see section on bias)
[E4.2] Test and minimise the likelihood of users encountering biases and/or provide mitigation strategies for implicit biases	• ^Adopt a systematic approach to evaluate the AI model's ability to perform intended tasks and functions across different scenarios • *Conduct regular, documented bias audits (i.e. on language, culture, gender, SES, SEN, etc) with explicit evidencing of bias, clear acknowledgement and mitigation strategies beyond the AI system	
[E4.3] Monitor AIEd system outputs regularly and rectify any identified errors to enhance system accuracy, ensure alignment with its intended purpose and mitigate against harm to students	• *Maintain oversight of content or output generated by the models to minimise errors and negative impacts on users' physical, socio-emotional and cognitive development • ^Put measures in place to guard against model drift and to ensure that the system outputs remain accurate and performs as intended, i.e. by acting on user reports • ^Evaluate outputs against established benchmarks or expert-defined 'ground truth', running sufficient iterations across varied scenarios. An independent AI model may also be used to cross-check outputs and flag errors.	