生成式AI對大學教育之衝擊與轉型

吳誠文
南臺科技大學校長

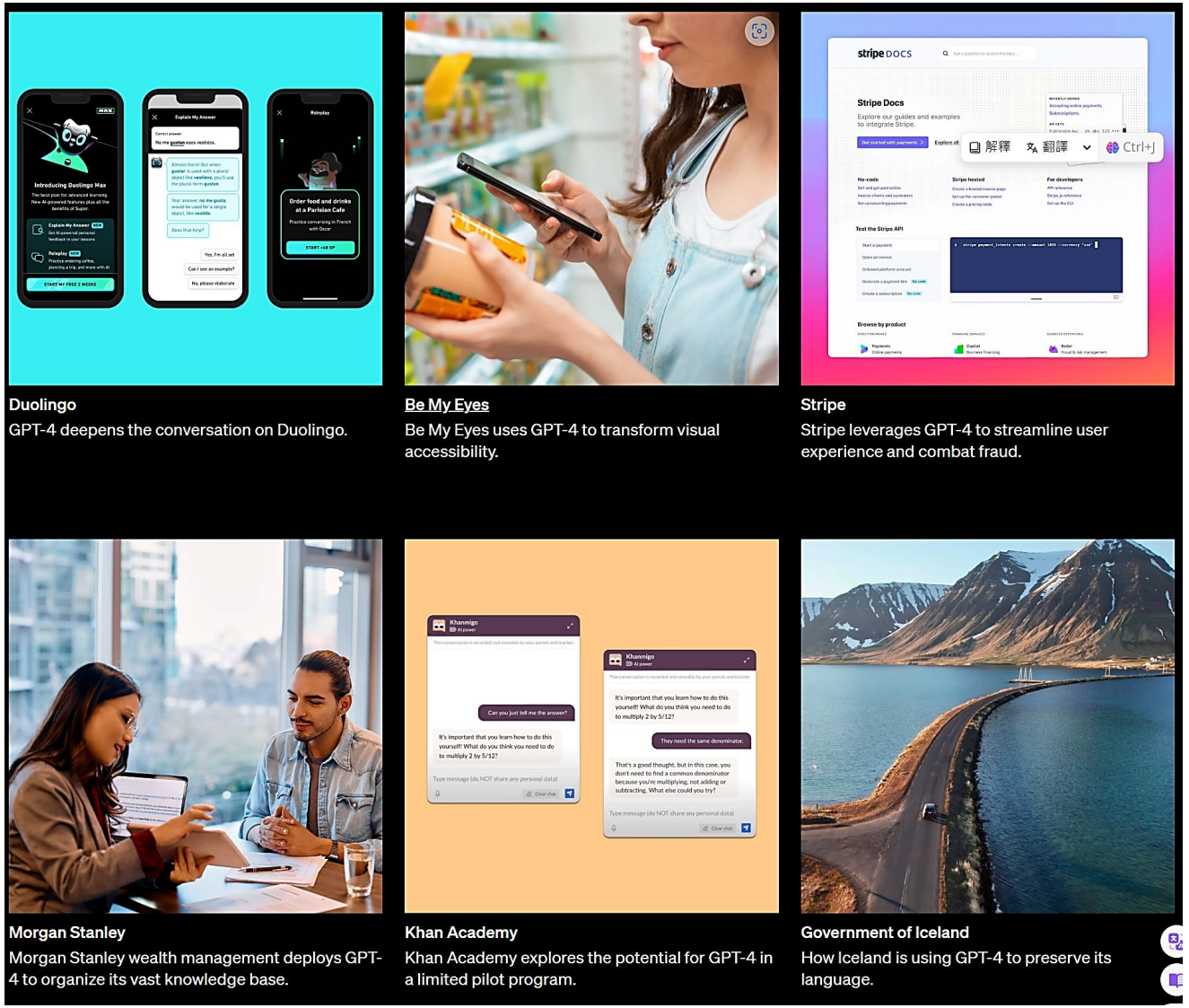
113年全國大專校院校長會議
2024 National Conference
for the University Presidents of Taiwan

Generative (生成式) AI

- A subset of AI that focuses on producing content that is creative, original, and often indistinguishable from human-created content
  1. Generative Adversarial Network (GAN)
  2. Variational Autoencoder (VAE)
  3. Autoregressive model
- GAN comprises two neural networks that work together to create data, images, text, and more
  - The **generator**: learns to produce realistic fake data from a random seed
  - The **discriminator**: learns to distinguish the fake data from realistic data
  - The discriminator penalizes the generator if it produces implausible results
- This technology has proven to be a game-changer in various fields, and education is no exception

ChatGPT and GPT-4

- GPT-4 was trained on Microsoft Azure AI supercomputers
  - Azure's AI-optimized infrastructure also delivers GPT-4 to users
- GPT-4 is available on ChatGPT Plus and as an API for developers
  - It still has many known limitations, such as *social biases* 社會偏見, *hallucinations* 幻覺, and *adversarial prompts* 對抗性提示
- There have been organizations building innovative products with GPT-4



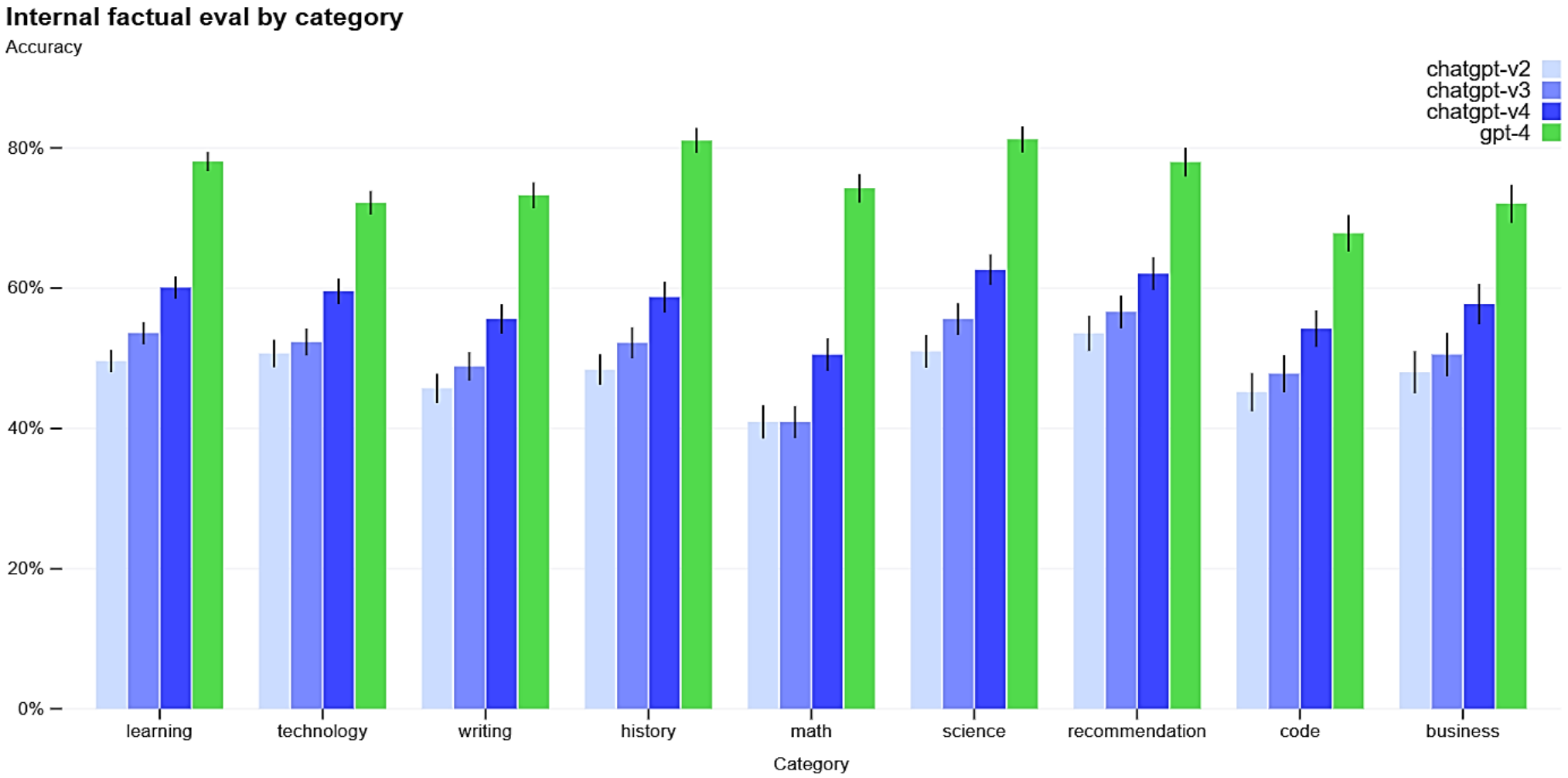
Source: openai.com/gpt-4 (Dec. 2023)

OpenAI的GPT-n系列AI模型

Model	Architecture	Par count	Training data	Release date	Training cost
GPT-1	12-level, 12-headed Transformer decoder (no encoder), followed by linear-softmax	**117M**	BookCorpus: 4.5 GB of text, from 7,000 unpublished books	June 11, 2018	**1 month on 8 GPUs (1.7e19 FLOPs)**
GPT-2	GPT-1, but with modified normalization	**1.5B**	WebText: 40 GB of text, 8M documents, from 45M webpages upvoted on Reddit	Feb. 14, 2019 (limited) Nov. 5, 2019 (full)	**1.5e21 FLOPs**
GPT-3	GPT-2, but with modification to allow larger scaling	**175B**	499B tokens consisting of CommonCrawl (570 GB), WebText, English Wikipedia, and two books corpora (Books1 and Books2)	May 28, 2020	**3.1e23 FLOPs**
GPT-3.5	NA	**175B**	NA	Mar. 15, 2022	**NA**
GPT-4	Trained with both text prediction and RLHF; accepts both text and images as input	**≈170T**	NA	Mar. 14, 2023	**NA, estimated 2.1e25 FLOP**

Source: en.wikipedia.org/wiki/Generative_pre-trained_transformer

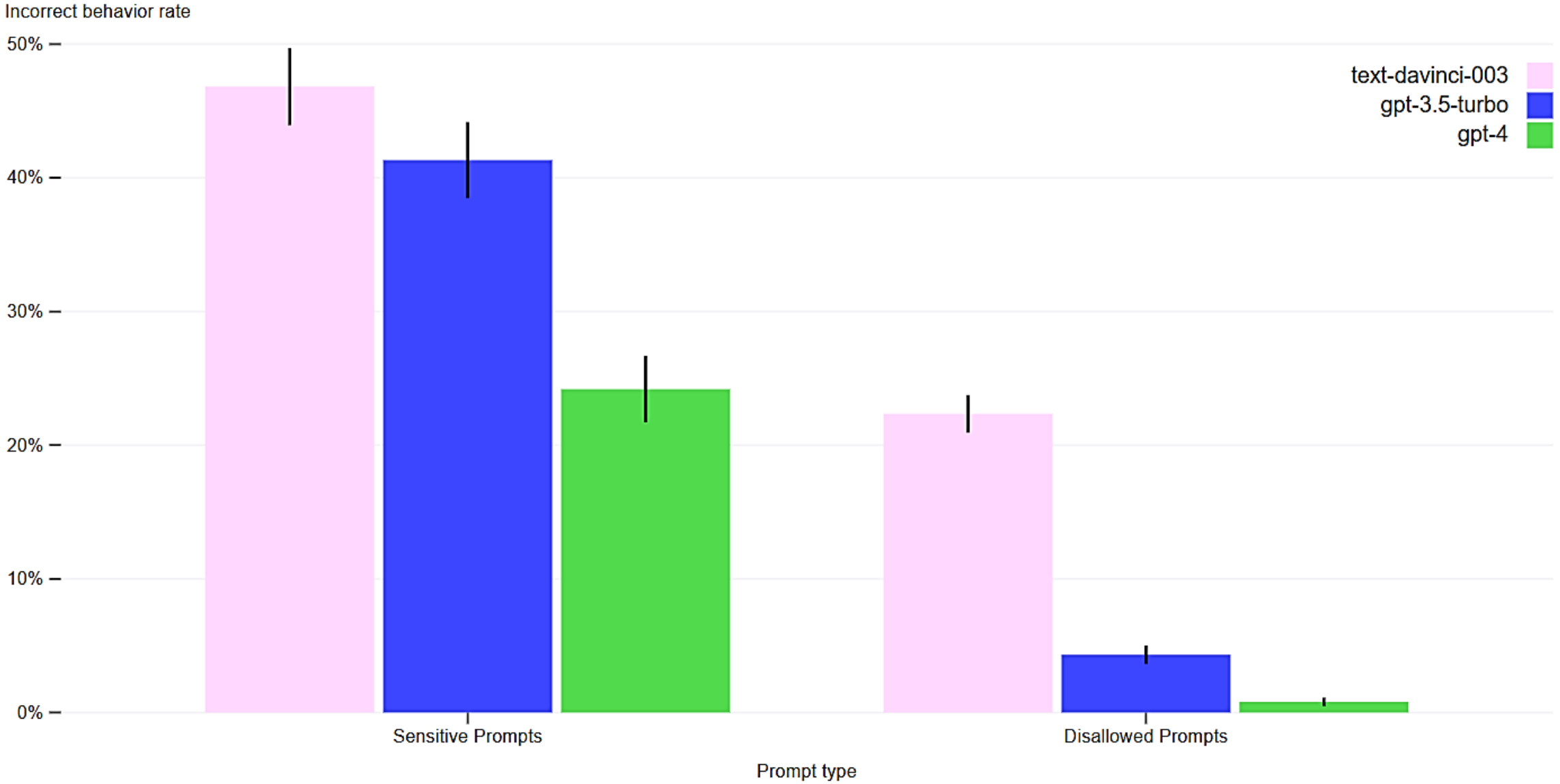
Factual Evaluation of the GPTs



On nine categories of internal adversarially-designed factual evals, we compare GPT-4 (green) to the first three ChatGPT versions. There are significant gains across all topics. An accuracy of 1.0 means the model's answers are judged to be in agreement with human ideal responses for all questions in the eval.

Source: openai.com/research/gpt-4 (Dec. 2023)

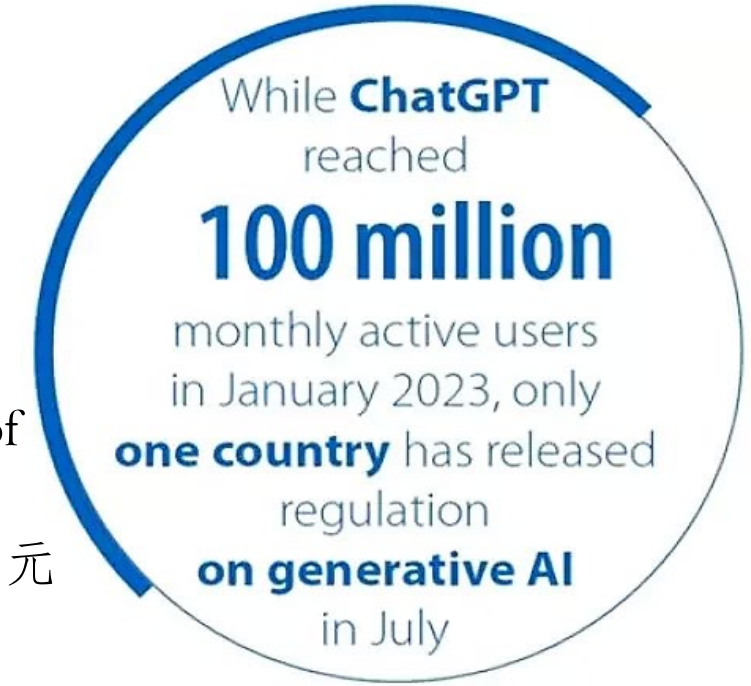
Incorrect behavior rate on disallowed and sensitive content



Source: openai.com/research/gpt-4 (Dec. 2023)

UNESCO Guidelines for AI in Education

1. Promote inclusion, equity, linguistic and cultural diversity 促進包容、公平、語言和文化多樣性
2. Protect human agency 保護人類代理人
3. Monitor and validate GenAI systems for education 監控和驗證GenAI教育系統
4. Develop AI competencies including GenAI-related skills for learners 培養人工智慧能力，包括學習者的 GenAI 相關技能
5. Build capacity for teachers and researchers to make proper use of GenAI 培養教師和研究人員正確使用GenAI的能力
6. Promote plural opinions and plural expressions of ideas 促進多元意見和多元思想表達
7. Test locally relevant application models and build a cumulative evidence base 測試與本地相關的應用程式模型並建立累積的證據庫
8. Review long-term implications in intersectoral and interdisciplinary manner 以跨部門和跨學科的方式審查長期影響



Ref: UNESCO, Guidance for generative AI in education and research, Sept. 2023

Co-designing Uses of GenAI to Support Teaching

Potential but unproven uses	Appropriate domains of knowledge or problems	Expected outcomes	Appropriate GenAI tools and comparative advantages	Requirements for the users	Required human pedagogical methods and example prompts	Possible risks
Curriculum or course co-designer	Conceptual knowledge on certain teaching topics and procedural knowledge on teaching methodologies.	Assisting with the curriculum and lesson design process, including outlining or extending views on key areas of the target topic and defining the curriculum structure. It may also help teachers prepare tests and exams by offering examples of questions and rubrics for evaluation. **Potential transformation: AI-generated curriculum**	Starting with the list in Section 1.2, assess whether the GenAI tools are locally accessible, open source, rigorously tested or validated by authorities. Further consider the advantages and challenges of any particular GenAI tool, and ensure that it properly addresses specific human needs.	The teachers must understand and carefully specify what they want the curriculum, courses, lessons, or tests to cover and achieve, whether they want to address procedural or conceptual knowledge, and what teaching theory they wish to apply.	Questions to GenAI on suggesting the structure and examples of factual knowledge on topic(s), suggesting teaching methods and processes for topics or problems, or creating course packages or lesson plans based on topic(s) and formatting. Human curriculum designers need to verify the factual knowledge and check the appropriateness of the suggested course packages.	The risk of GenAI imposing dominant norms and pedagogical methods is high. It may inadvertently perpetuate exclusionary practices in favour of the already data-rich groups and reinforce inequalities in access to relevant and high-quality educational opportunities, disadvantaging data-poor groups.
Generative chatbot as teaching assistant	Conceptual knowledge across multiple domains in well-structured problems.	Providing individualized support, answering questions and identifying resources. **Potential transformation: Generative twins of teachers' assistants**	Starting with the list in Section 1.2, assess whether the GenAI tools are locally accessible, open source, rigorously tested or validated by authorities. Further consider the advantages and challenges of any particular GenAI tool, and ensure that it properly addresses specific human needs.	It supports teachers but targets learners directly, so this requires learners to have sufficient prior knowledge, abilities and metacognitive skills to the verify the outputs of GenAI and notice the misinformation. Thus it might be more appropriate for learners in higher education.	Requires the teachers to understand the problems clearly, to monitor the conversation and help learners to verify dubious answers provided by GenAI.	Based on the current capabilities of GenAI models, educational institutions need to guarantee human supervision of the responses provided by GenAI tools, being alert to the risk of misinformation. It may also limit learners' access to human guidance and support, hindering the development of a strong teacher-student relationship, which is especially concerning for children.

Ref: UNESCO, Guidance for generative AI in education and research, Sept. 2023

Data, Information, and Intelligence