

Examining, teaching, and learning in the age of generative AI

A first assessment of the consequences for higher education
by **Thomas Bieger** and **Martin Kolmar**





A typical and intuitive first reaction of educators to new technologies challenging traditional ways of teaching is often to ban them or to constrain their use

Many educators and managers of IHEs may have undergone a similar experience in December 2022 - often alerted by their teenage kids, they became aware of advances in GAI with the potential to disrupt teaching and examination. A lot has happened since then, as GAI promises to be the “next big thing,” potentially revolutionising everything from (white collar) work to the way we organise information and relate to each other. Although it is not clear at this time exactly what will survive the hype, the potentially far-reaching consequences of this new technology make it necessary for IHEs to develop early strategies for how best to respond to it. Chatbots like ChatGPT or GPT 4 raise the question of how GAI will change or add to the competencies our graduates will need in the future world of work, what pedagogical and didactical formats will be required to teach these competencies, and ultimately how it will affect the overall strategies of IHEs.

How to react to GAI in the classroom?

A typical and intuitive first reaction of educators to new technologies challenging traditional ways of teaching is often to ban them or to constrain their use. One example is the time when calculators became available, but students were still forced to solve math problems by hand. This reaction is flawed. If the goal of education is to prepare students for real (work) life, and if GAI promises to be an essential part of it, we must find ways to integrate it into our programs. IHEs have an obligation to help students develop the necessary skills to use these tools productively—and to understand their limitations, including the associated ethical challenges.

Competencies for the new workplace

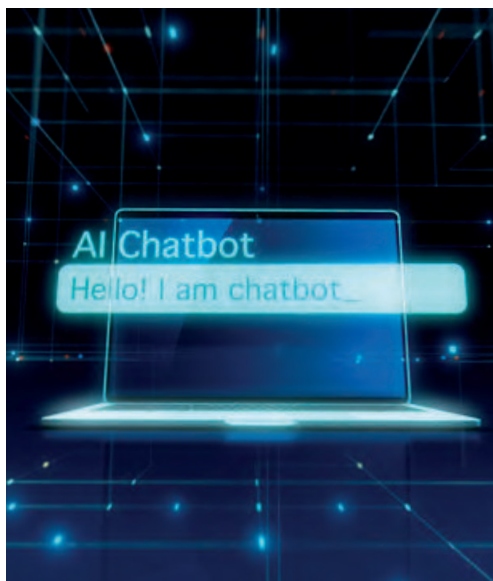
IHEs, therefore, need to identify the knowledge, skills, competencies, and attitudes required—and place them at the centre of their curricula, teaching, and examination methods. To do so, three fundamental questions must be addressed:

- First, we need an idea about the impact of GAI on the future of work, how it will change the way industries organise their value chains and create new, changing or even eradicating existing professions in this process. Program managers must reassess the skills and competencies necessary for their students to flourish in their future work life. Humans cannot compete with machines in areas where machines are designed to excel; education must equip students with the competencies to create value beyond the means of AI.
- Second, despite the growing body of research, we do not yet fully understand how students learn and develop their skills and personalities effectively in environments that blend digital tools with traditional teacher-to-learner formats. However, we need a robust understanding of the optimal mix of human and technological support to students for developing programs effectively.
- Third, as debates on reforms of management education have become a central theme of business schools over the past decade, the challenges in teaching and assessment necessitated by GAI should be integrated into this overarching discussion. Skills, competencies, and personality traits that remain relevant over long periods and in new, changing, and unknown contexts have to be identified and developed.

We will talk about a very narrow class of GAI to assess its implications on higher education. Chatbots like ChatGPT that are large-language models (LLMs). They are trained on huge quantities of text data to identify the most likely contexts in which phrases are used; what LLMs do is basically sequence prediction. This property leads to the now well-understood phenomenon of hallucination, i.e., that LLMs “invent” false texts or literature references because truth and probability are generally different.

This fact has important implications. A consensus seems to be emerging that LLMs are helpful for people qualified to ask the right questions and competently evaluate the output, but much less so for people without these evaluative competencies. If the GAI tool finds the closest association with the prompt, generic or nonsensical output is often a result of “bad” prompting. Some of these issues can therefore be resolved by learning how to “prompt well.” To use the potential of LLMs, the user needs *sufficient expertise to evaluate its output* and to improve the prompts from there.

There are, however, important exemptions from this rule. For example, it turns out that LLMs provide great opportunities to support students in learning to code. Unlike natural languages, artificial languages like *Python* exhibit its high degree of syntactic and semantic precision. Therefore, LLMs are exceptionally suitable for supporting learning by generating, optimising, and correcting code. The same is true for mathematical problems. Contrary to initial belief, LLMs are not necessarily “bad at math,” since (for standardised problems in teaching at least) one can often take the detour of letting LLMs generate code, which then contains the solution.



Which brings us to the hard problem for teaching and examining: If LLMs are most useful as support systems for people who already have the competencies to evaluate the output generated, how can we make sure and assess whether students acquire these competencies if they can fake them by using LLMs? This problem forces us to think about and reimagine how we teach and evaluate. Thereby we have to consider other, normative challenges that result from this specific form of text generation: How can we ensure that there is no uniformisation of theoretical and empirical interpretations of reality, given that LLMs “mainstream” text in the way described above? How can we ensure that all credible scientific views are correctly reflected, including heterodox ones? And how can we ensure that students are able to distinguish between valid arguments and hallucinations, especially if they are still in the process of developing the evaluative competencies mentioned above? Critical thinking in all facets becomes more important than ever.





Consequences for pedagogy, didactics, and examination

Many universities felt pressured to react quickly and develop guidelines and best practices to minimise the risk of students handing in essays generated by GAI tools. Examples have been: to “customise” writing assignments, break major assignments into smaller, individually graded chunks, prioritise on-campus exams, test assignments by grading the output generated by a chatbot, require heavy citations, and return to time-honoured oral exams, etc.

These “quick fixes” were mainly driven by the fact that the new technology became available during the lecture and examination period, which created the need to act, leaving the impression that LLMs pose a threat rather than an opportunity. The problem is that if the fire brigade is out, sustainable long-term solutions are rarely achieved. For example, before we rush back to oral exams, we should remember the phenomenon of examiner bias. Or, to give another example, it seems clear that the responsible use of GAI as part of academic integrity requires adequate standards of use—which is a challenge since, for example, the traditional concept of plagiarism does not readily capture the new phenomenon.

The deeper problem seems to be that as LLMs can generate exam-passing texts this reveals what kind of competencies we are implicitly expecting from our students. If what LLMs do is sequence prediction, and if sequence prediction passes exams, we must ask ourselves if this is really what we should be expecting from our students. If LLMs excel at generating good essays, we learn that we are expecting *mainstreaming* from our students. Certain examination formats just invite students to “blindly” memorise theories. If the exam questions are in addition very generic, it is little wonder that LLMs pass exams. It seems necessary to use the challenge posed by LLMs to better understand whether our exams consistently align with the competencies we want to develop and the teaching formats we are using.



PICTURE COURTESY OF UNIVERSITY OF ST GALLEN

GAI

It seems clear that the responsible use of GAI as part of academic integrity requires adequate standards of use



The fast rate of technological progress requires a continuous redefinition of teachers' qualifications

A new focus on teaching in faculty management

If these conjectures are correct, enabling a reflective learning process will become an even more important critical success factor for business schools. Academic careers are still almost exclusively built on research credentials. This time-approved model of selection has been adequate for as long as universities were the more or less exclusive access points for knowledge and content was decisive. The way we are teaching did not change very much over the years as its primary role was to give students access to knowledge. Digitalisation and, even more, the emergence of GAI changes that picture, as—except for fundamental research—access to knowledge became ubiquitous for everyone with access to the internet. What becomes increasingly important is no longer *what* we teach but *how* we teach it. But at present, faculty is not usually selected to excel in this dimension. Hence, we must reassess the necessary qualifications for academic teachers, train the existing faculty to “teach up” to the new challenges, and rethink the criteria for hiring new faculty. The ability to foster the development of *epistemic, social, and personal virtues* like curiosity, critical thinking, sociability, responsibility, intrinsic motivation, and resilience are key qualities of good teaching in interaction with digital tools.



More and more universities offer separate career paths for teaching and research. If our analysis is correct, these teaching tracks must be more than second-class alternatives, they should focus on a unique blend of research and teaching skills. The fast rate of technological progress requires a continuous redefinition of teachers' qualifications. Therefore, the 'teacher' career path requires the ongoing reassessment of the best teaching and examination formats based on empirical evidence. Thus, universities should not only engage in financial investments in these tracks but also actively search for qualified personalities, and so create a culture of learning and critical reflection on the best teaching and examination techniques. Moreover, a whole ecosystem, including organisational support for experiments, labs, and staff for technical support, will be an essential element in this process.





PICTURE COURTESY OF UNIVERSITY OF ST. GALLEN



Strategies of IHEs challenged

The use of GAI has the potential to further increase the gap between low-cost IHEs focusing on teaching basic skills and competencies and IHEs that can invest in a unique blend of research excellence and high-quality teaching to enable their graduates to deliver value beyond the capabilities of machines. To an extent this has already been driven by the high costs of funding basic research, e-learning, and other developments which disrupt the traditional academic 'value chain'. To qualify students to make valuable societal contributions requires teaching and examination formats that are more interactive, individualised, and focus on personality development. Although digitalisation, including LLMs, allows support and even replacement of some traditional teaching and examination

formats, as long as education is based not only on the *knowing*, but also the *doing* and *being* dimensions of learning, then at least, for the time being, humans will be enablers of these learning processes. These tools will not replace human beings in education and will not necessarily mark 'the end of the college essay', but they make it necessary to reassess their most productive roles.



About the Authors

Thomas Bieger is Professor of Business Administration and Director of the Institute for Systemic Management and Public Governance at the University of St. Gallen. He served as the university's President between 2011 and 2020.

Martin Kolmar is Professor of Economics and Director of the Institute for Business Ethics at the University of St. Gallen.