Hochschulforum Digitalisierung

WORKING PAPER NO. 86 / MARCH 2025

Artificial intelligence: Principles for action in university teaching

Results of the working group "Artificial Intelligence: Essential competences at universities"

Alexander Filipović, Aljoscha Burchardt, Simon Hirsbrunner, Antje Michel, Anna Puzio, Gabi Reinmann, Philipp Schaumann, Anja-Lisa Schroll, Ulrike Tippe, Martin Wan, Nicolaus Wilder

Working Paper No. 86 / March 2025

Artificial intelligence: Principles for action in university teaching

Results of the working group "Artificial Intelligence: Essential competences at universities"

Authors:

Alexander Filipović,
Aljoscha Burchardt,
Simon David Hirsbrunner,
Antje Michel,
Anna Puzio,
Gabi Reinmann,
Philipp Schaumann,
Anja-Lisa Schroll,
Ulrike Tippe,
Martin Wan,
Nicolaus Wilder

The Higher Education Forum on Digitalisation

As a nationwide think and do tank, the Higher Education Forum on Digitalisation (HFD) brings together a broad community around the digital transformation at universities, makes developments visible and tests innovative solutions. To this end, stakeholders from universities, politics, business and society are networked.

Established in 2014, HFD is a joint initiative of Stifterverband, the CHE Centre for Higher Education and the German Rectors' Conference (HRK). It is funded by the Federal Ministry of Education and Research (BMBF).

STIFTERVERBAND
CHE Centrum für Hochschulentwicklung
HRK Hochschulrektorenkonferenz
GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung

Inhaltsverzeichnis

1 Introduction

Since **generative artificial intelligence (AI)** tools became known to a wider public as serious developments with the publication of ChatGPT in November 2022, their development accelerating significantly in the months that followed, immense effects have emerged for the university sector. Political, economic, social and educational dimensions are recognisable.

Questions about the origin of information, its reliability, its bias and partiality and its ability to be cited have already been discussed many times. Data protection, data security and copyright issues associated with AI are also becoming increasingly urgent matters for discussion. In addition to the **effects on studying, teaching and assessment**, the sometimes unintentional, sometimes deliberate exploitation and thus devaluation of academic work must also be taken into account.

In contrast, the question of how to deal with intentional and unintentional manipulation of data and, as a result, knowledge has not yet been sufficiently analysed. Both the **implicit biases inherent in the system,** depending on specific language models or the underlying data, and deliberate distortions such as targeted propaganda or the illusion of a social/human counterpart inherent in AI chatbots are problematic: since natural language is used to interact with chatbots, we are confronted with the phenomenon that human-machine interaction can actually easily be ascribed characteristics of human-human interaction. On the one hand, this raises **questions of social interaction** and, on the other, it necessitates a debate on genuinely human capabilities and the inherent technological limits of artificial intelligence.

These issues are of great and growing **importance for a variety of social contexts** beyond that of higher education. Development is no longer dependent on individual tools or applications, but has now given rise to an extensive digital ecosystem. To the extent that (generative) artificial intelligence permeates all areas of life, knowledge of how it works on a fundamental level, in particular a knowledge of the associated value chains, is becoming increasingly important. A comprehensive **understanding of the shape, mechanisms of action and key drivers of** these digital ecosystems is therefore a prerequisite for their adequate utilisation, active design and further development, as well as for a critical scholarly examination of them.

The task of universities is to enable their graduates to act as mature citizens in this world. In particular, they are faced with the challenge of teaching their members the skills that will enable them to make informed decisions in an environment increasingly characterised by AI systems, so that they can **participate in society both now and in the future**. This includes the ability to judge where AI can be used appropriately and for which applications and tasks it is not suitable. Only if all university members, regardless of their field, understand the basic principles and limitations of AI can misjudgements be prevented and its potential appropriately exploited, contributing to a constructive social debate on new technologies.

From December 2023 to December 2024, the **working group "Artificial Intelligence: Essential competences at universities"** engaged with these issues. The objective of the work-

ing group went beyond practical advice on usage, recommendations for action or practical examples of individual, current applications such as ChatGPT, as these are in danger of quickly becoming obsolete in view of the dynamism and expected diversity of developments in this processing format.

Rather, within the framework of classic think tank work, the aim was to fundamentally address the question of how a **competent, reflective and academically sound approach to AI** can be institutionally anchored and concretely implemented at universities. The aim of these considerations should be to identify essential competences in dealing with AI for university members at all levels.

As a first step, the question of what goals higher education has in mind was discussed. In a second step, the values that can be assigned to these objectives of higher education were derived from this. Thirdly, with regard to artificial intelligence, the extent to which these **values and goals** change **as a result of AI** was analysed. Based on these preliminary considerations, the final step was to develop perspectives for action for universities in the context of AI as well as opportunities for skills development.

The working group worked in **the form of scenarios** in order to raise these considerations to a practical and illustrative level. The scenarios deal in particular with the topics of integrity, autonomy and research. At the same time, in a dynamically developing field such as artificial intelligence, these scenarios do not claim to be a conclusive observation that is difficult to make, but rather invite reflection.

The working group was chaired by Prof Dr Alexander Filipović (Chair of Christian Social Ethics, Faculty of Catholic Theology at the University of Vienna).

The members of the working group were

- Dr Aljoscha Burchardt: Deputy Site Spokesperson, German Research Center for Artificial Intelligence / AI Campus
- Dr Simon Hirsbrunner: Team Leader, International Center for Ethics in the Sciences and Humanities, University of Tübingen
- Prof Dr Antje Michel: Professor of Media Didactics and Knowledge Transfer, FH Potsdam
- Dr Anna Puzio: Ethics of Socially Disruptive Technologies (ESDiT), University of Twente
- Prof Dr Gabi Reinmann: University Didactician, University of Hamburg
- Dr Philipp Schaumann: Chair, KMK Working Group on Artificial Intelligence / Ministry of Science and Culture of Lower Saxony
- Prof Dr Ulrike Tippe: President of TH Wildau / HRK Vice-President for Digitalisation and Academic Continuing Education
- Dr Nicolaus Wilder: Research Assistant, Institute of Education, CAU Kiel

The working group was supervised by Martin Wan, Anja-Lisa Schroll and Stella Berendes from the office of the German Rectors' Conference in Bonn as part of the Higher Education Forum on Digitalisation.

2 The values and goals of higher education

2.1 What are the goals of higher education?

According to Huber (1983), universities have[1] three core functions[2]. They are:

- Educational institution: They offer comprehensive training and qualifications and thus serve a **practical**purpose.
- Research system: They generate new **knowledge** and nurture young researchers.
- Living and working environment: They provide **people** with education and free scope for research.

These three functions – related to practice, research and the individual – form the point of reference for the university as an institution and for its teaching. In 2015, the German Council of Science and Humanities formulated this triad as the goals of higher education:

*"Three central dimensions span the space of higher education goals: **(specialised) science, personal development and labour market preparation**. The dimension of (specialised) science is determined in particular by qualification objectives that are designed to enable students to select, apply and adapt scientific methods to suit the situation and to deal independently and critically with scientific findings. Objectives that promote the development of a professional identity and an academic and professional ethos or prepare students to take on responsibility in their profession and in society can be categorised primarily under the dimension of personal development. Finally, the labour market preparation dimension concerns the qualification of students, which is directly and specifically geared towards working life after graduation – within or outside academia."*[3]

2.2 What values underlie the goals of higher education?

Universities not only pursue educational goals in the narrower sense, but also fulfil a variety

1 Huber, Ludwig, Hochschuldidaktik als Theorie der Bildung und Ausbildung, in: idem. (ed.), Ausbildung und Sozialisation in der Hochschule (= Enzyklopädie Erziehungswissenschaften 10), Stuttgart 1983, 127-129.

2 Cf. Reinmann, Gabi; Watanabe, Alice, KI in der universitären Lehre, in: Schreiber, Gerhard; Ohly, Lukas (eds.), KI: Text. Diskurse über KI-Textgeneratoren, Berlin 2024, 29-46, 31 f.

3 German Council of Science and Humanities, Recommendations on the relationship between higher education and the labour market. Second part of the recommendation on the qualification of skilled workers (Drs. 4925-15), Bielefeld 2015, retrieved on 21 January 2025 at: https://www.wissenschaftsrat.de/download/archiv/4925-15.pdf, 9.

of social functions. For example, universities can help to promote social cohesion and safeguard the free and democratic basic order. One example of this is the 'unity of research and teaching', which is constitutive of the Humboldtian idea of European universities. The overarching goals of universities and higher education can be assigned **values** that are closely linked to these target perspectives.

- **Research:** As a central guiding value, 'academic integrity' is at the centre as an expression of the scientific ethos. This generic term refers to "respectful treatment of each other, study participants, animals, cultural assets and the environment" as well as responsibility in view of the freedom of research,[4] which also includes social responsibility. This in turn is achieved through values that characterise the research process. These include the quality or soundness of research processes, such as compliance with subject-specific standards and established methods, clear responsibilities for research work processes, inclusion of existing research achievements (state of research), adherence to legal and ethical requirements (compliance), comprehensive publication of all research results and relevant information (transparency, principle of openness), recognition of authorship and authors, confidentiality and neutrality in reviews.[5] If we focus on the specific educational tasks of universities in the context of science, then maturity and critical thinking in dealing with scientific findings are central values.

- **Personal development:** Today, the education of individuals is primarily related to autonomy in the sense of individual self-determination. In their self-determination theory, Deci and Ryan have identified three basic psychological needs and they have proven in empirical studies over several decades that these are very stable: experience of competence, experience of autonomy and social integration. In this sense, a degree programme should enable students to experience competence and autonomy and offer potential and experience of social integration. Furthermore, it can be assumed that real opportunities for participation in the institution, transparency and fairness are prerequisites for a bonding experience that ultimately produces responsible subjects who are able and willing to shape modern democracy and global contexts. In view of the diversity of cultures, lifestyles and ideas of the good life, values such as respect and tolerance are linked to the goal of personal development for social cohesion.[6]

- **Labour market preparation:** Values associated with the labour market orientation of universities include opportunities for participation in social processes through work for students and, as a social institution, fairness in the sense of fair opportunities on the labour market for graduates. These opportunities for participation also mean that stu-

4 German Research Foundation, Guidelines for Safeguarding Good Research Practice. Code of Conduct, Bonn 2022 (corrected version 1.1), retrieved on 21 January 2025 at: https://zenodo.org/records/6472827/files/kodex_leitlinien_gwp_dfg.1.1.pdf, 7.

5 Cf. ibid.

6 Cf. Deci, Edward L.; Ryan, Richard M., Self-determination theory: A macrotheory of human motivation, development, and health, in: Canadian Psychology 49 (2008) No. 3, 182-185.

dents are able to shape society through their work (participation, maturity) and take on responsibility thanks to the strong labour market orientation of their education.

•

3 Values and objectives changed by AI

Once these values and objectives have been identified, the question arises as to the extent to which they will be challenged or changed by artificial intelligence.

- **Are individual values or goals of higher education being reinterpreted against the background of technical innovations?**
- **Are the conditions for realising the goals and values of higher education changing in the context of AI technologies?**
- **How can the goals and values of higher education be strengthened through the responsible use of AI?**

The working group approached these questions using scenarios that are examples of everyday university life. We developed the scenarios in four steps, using existing techniques from **futurology and design research**[7]. Together with all members of the working group, we first assigned the values of higher education presented above to different teaching and research situations, identified drivers and obstacles and worked out a fundamental area of tension in each case based on the expertise of the think tank members. A team of two worked on these problems, which served as the starting point for the **scenario development**. From this, scenarios were developed in small expert groups, the objectives of which were laid out in the possibility space of a constructively designed probable development scenario. The scenarios developed were revised and validated with the help of a two-stage peer review process by the other think tank members and then by participants in a workshop at the University:Future Festival. The scenarios were deliberately developed as factual and, if anything, positive spaces of possibility; the aim was therefore to make our **approach to goals and values in the future** comprehensible with a productive use of AI tools. When developing the scenarios, we orientated ourselves on a very near future. Some of the framework conditions of our scenarios have already been realised in current developments over the course of the one-year working group.

The scenarios served a dual purpose in our think tank work and also in this document. Firstly, as part of the scenarios, we played through the possibilities of engaging with the values and goals described above in order to think through their relevance for the AI-integrating university and to develop recommendations for approaching AI. The scenarios formed the **condensed discourse basis** for the discussion with experts, with whom we validated and enriched the method and results of our think tank. Secondly, in the think tank, through the scenario technique we trialled a methodical approach to developing possible future perspectives from empirical research and reflection on the present. We have thus practised the ability that is described by the term **futures literacy**[8] – not to be confused with the set of supposed future skills currently much discussed in the context of AI, which is considered neces-

7 See, for example: Rosson, Mary Beth; Carrol John M., Scenario-based design, in: Sears, Andrew; Jacko, Julie A. (eds.), The Human-Computer Interaction Handbook, New York 22008, 1041-1060; De Jouvenel, Hugues, A brief methodological guide to scenario building, in: Technical Forecasting and Social Change 65 (2000) No. 1, 37-48; Nathan, Lisa P. et al, Value scenarios. A technique for envisioning systemic effects of new technologies, in: CHI '07 Extended Abstracts on Human Factors in Computing Systems, San Jose 2007, 2585-2590, available at: https://dl.acm.org/doi/10.1145/1240866.1241046 (retrieved on 17 February 2025).

8 Cf. Miller, Riel, Futures Literacy. A hybrid strategic scenario method, in: Futures 39 (2007) No. 4, 341-662.

sary to succeed in a future working and living environment.[9] The methodological approach is based on the assumption that every assessment of developments always contains an implicit idea of possible consequences. We have thus explicated these unspoken ideas, which are often either overly euphoric or dystopian in tone when it comes to technical innovation, and made them usable as a framework of orientation for the joint discussion.[10] A well-founded vision of the future, because this **collaboratively developed orientation framework** is precisely that, is essential for decisions in the present, as it makes people aware of the scope and limits of action. We have decided to apply this dual purpose to this report and invite you to perceive the following scenarios firstly as illustrations of abstract goals and values in possible futures. Secondly, we would like to use our examples to encourage people to use collaboratively developed visions of the future as a starting point for positioning themselves in relation to current developments. Stakeholders might use our scenarios or create their own scenarios.

3.1 Scenario I: (Academic) integrity

The following scenario describes a typical situation that may currently be taking place at many universities. In essence, it is about various aspects of academic integrity and individual as well as **collective responsibility in higher education**, which are particularly evident when it comes to assessment. The scenario does not provide any ready-made 'solutions' for dealing with AI in the context of university assessments; rather, it invites discussion. Accordingly, after the story we contribute our own dialogue, which developed from thinking about three key questions that the expert group (working group) gave itself – on the assumption that these key questions reflect the current need for discussion at universities.

Initial situation

Sarah Kaya is a professor of business informatics at her university and every semester she organises a seminar on robotics. In the past semester, she found that several students wrote large **parts of their assignments with ChatGPT**. She does not simply want to prohibit the use of generative AI, but to encourage a productive and critical approach to it and find a way to organise the seminar in such a way that students develop skills for a constructive approach to AI in their studies.

Constructive vision

When thinking about how to deal with the risks of ChatGPT in assessment situations, Sarah realises that she first needs to re-examine her previous routines with respect to the seminar: **What exactly do I want to achieve with the assignment?** Is the assignment suitable for acquiring and demonstrating the skills that are central to the robotics course? Would an

9 See, for example, Ehlers, Ulf-Daniel, Future Skills. Lernen der Zukunft – Hochschule der Zukunft, Wiesbaden 2020.

10 Cf. Miller, Riel, Transforming the future. Anticipation in the 21st century, Paris 2018.

alternative form of proof of achievement be fundamentally better? The students could use generative AI in preparation, but then present the results of their work in person and without technical assistance. She herself could then give direct feedback.

In the cafeteria, Sarah meets her colleague Yago Wagner, a research assistant in political science. Sara uses the chance encounter to talk about her idea for the new type of assessment. The debate quickly becomes heated: Yago thinks Sarah's solution of giving up assignments is short-sighted. Is this not merely a reaction to external technical developments? Sarah defends her decision, but in the end is happy about the controversial exchange, because this helps her to realise even more clearly why, in her case, an oral presentation of student work as a piece of assessed work can be justified even better than a written assignment. At the same time, however, she realises in conversation with Yago that her approach to the GPT problem in student assignments cannot be a model for all subjects.

Yago also has something to report on a topic that is currently on many people's minds. He is planning a teaching experiment: The students themselves choose whether they write the assignment without generative AI or use generative AI and then document and comment on its use. The Faculty of Social Sciences has already provided guidelines on the latter. The class should be preceded by a discussion with the students about the importance each form of writing competence can have: for their future career, for their personal development, for the subject. On the one hand, Sarah is sceptical: doesn't the first option let the problem back into the seminar room through the back door? Yago counters: Is Sarah's mistrust of the students getting the better of her? On the other hand, Sarah is impressed: with Yago's approach, dealing with ChatGPT becomes more reflective. The coffee break ends all too quickly, but the discussion lingers in both of their minds.

In the current semester, Sarah is putting her idea into practice. In her faculty, she achieves an **amendment to the examination regulations**, which now include the option of an oral presentation. This will also increase the scope for didactic examination decisions in the future. The university has also provided a data-secure infrastructure for the use of generative AI in time for the semester. This ensures that no person in the seminar is at a disadvantage if they allow and now also support the use of suitable systems in the semester process.

Yago also conducts his teaching experiment, but has a big problem to deal with first: students and lecturers in his faculty have raised the objection that the alternative forms of assessment are not comparable and are therefore questionable in terms of examination law. Yago manages not to let his frustration gain the upper hand in the face of these conflicting values. Together with two lawyers from the legal department, a solution is found specifically for teaching experiments. Even the university leadership is enthusiastic about this solution: it is publicly in favour of facing the challenges of generative AI more courageously and with an experimental attitude in the future – including in teaching.

Consideration of the scenario based on key questions

Are individual values or goals of higher education being reinterpreted against the background of technical innovations?

What does our scenario say? In our scenario, the goals of higher education, such as those formulated by the German Council of Science and Humanities (2015) as (specialised) scientific knowledge, preparation for the labour market and personal development,[11] (implicitly) remain in place against the backdrop of AI. Sarah, a business informatics specialist, is primarily focussing on the requirements placed on graduates with a view to future professional activities, which today certainly include knowledge and skills relating to AI. Political scientist Yago's argument is aimed at making students think and take responsibility for their own decisions, which is primarily aimed at personal development. Both recognise that AI will not disappear from academia and will be part of academic disciplines in the future, and consequently rule out banning AI for their assessments. Instead, they strive for forms of assessment that ensure in different ways that **AI** is **not misused**. Accordingly, there are a whole series of statements or allusions in the scenario that indicate that existing values will be retained: in addition to the integrity that gives the scenario its title, above all responsibility, transparency, authorship, self-determination, participation, equal opportunities and openness or a willingness to experiment.

How should the key question be understood? Strictly speaking, key question 1 can be read in two ways: as an empirical question in the sense that **surveys or content analyses** could be used to examine whether and how existing goals and values at universities are shifting or changing, e.g. becoming less important, gaining priority, being dropped, or new ones being added, etc., or as a philosophical question as to whether and for what reasons current goals and values at universities should shift or change. Fundamentally, both interpretations and the associated questions seem important to us: Are individual goals and values already being reinterpreted and is this likely to continue or cease in the future? Would it be desirable to reinterpret individual goals and values now and in the future or not, and what reasons can be given for this?

An empirical description of the current situation – as an answer to the first question – provides a guide as to how people at universities currently deal or intend to deal with AI. Knowing this is important, among other things, when designing initiatives or measures for the use of AI, which (as a rule) requires acceptance and a certain degree of compatibility with the existing system. From the current situation, it is possible to draw conclusions about what is likely to happen in the future. In the context of our discussion of AI in the working group, however, the empirical question does not arise for two reasons: firstly, because we do not have the (time and financial) resources to do so, and secondly, because we have decided not to be driven (further) by technological normative facticity, but to **regain the active role** and negotiate goals and values ourselves in order to then think about how AI can be used as a tool.

A philosophical **determination of the target situation** opens up the space to perceive the future as a possible and mouldable one, and amounts to illuminating what is desirable for humans. It is necessary to determine whether (and why) we have premises with regard to goals and values that are not affected by technical innovations. In the course of our discussion on AI in the working group, including the development of this scenario, we have already identified values and embedded them in the story. Within the scenario, therefore, the question of

11 Cf. Wissenschaftsrat, Recommendations, 9.

whether goals and values are (or should be) reinterpreted does not initially arise, but at most theses become visible on the question of their **relevance and feasibility under AI conditions** (which leads us to the second key question). In other words, with this scenario, we have provisionally answered the philosophical question from our perspective, but we do not see this answer as an absolute, but rather as a possible one; at the same time, it is intended as an impulse to discuss it in the university community and to weigh up different perspectives from which one could also arrive at other answers.

What do we conclude from this? Ultimately, after considering our scenario with regard to the first key question, we are in favour of using the scenario as a thought experiment[12] and varying it in various aspects (e.g. initial problem, AI potential, teachers' perceptions, etc.). In this form, the scenario invites us to reflect on whether there are reasons for or against the reinterpretation of values in the course of AI development and use, and if so, what these reasons are.

Are the conditions for realising the goals and values of higher education changing in the context of AI technologies?

What does our scenario say? Our scenario explicitly addresses two types of conditions that are important for realising existing goals and values even in the AI era: technical infrastructures that enable all students to use AI in a way that is justifiable under data protection law, as well as legal provisions, including examination law, that enable teachers to make flexible didactic decisions and conduct teaching experiments such as giving students the freedom to choose whether to use AI in assessments. Implicitly, a further realisation condition is that teachers have didactic competence and creativity, because ultimately, all questions regarding the use of AI in higher education at the micro level are of course didactic in nature. In this sense, the scenario partly assumes and partly adapts the conditions for realising the goals and values of higher education so that these can also be achieved with AI. The scenario also shows that AI often acts as a **driver for changes** that would make sense in principle but have not yet been implemented, probably due to a lack of pressure from negative effects of non-implementation. The appraisal would probably be different if we had a scenario in which the goals and values of higher education changed. Let us suppose that we were to (largely) abandon our previous value of authorship in favour of a hybrid creation of texts in which human and machine form a network of relationships that can no longer be untangled. In this case, in our scenario Sarah and Yago would make different (assessment-related) didactic decisions or promote other student skills and take them into account in assessment design. They would establish a new value – e.g. hybrid authorship – and would require varying degrees of **change in technical, legal and didactic frameworks** that deviate more clearly from the current ones.

12 Cf. Reinmann, Gabi, Gedankenexperimente als bildungstheoretisches Instrument in der Forschung zu Künstlicher Intelligenz im Hochschulkontext, in: Impact Free 56 (2014). Available at https://gabi-reinmann.de/wp-content/uploads/2024/09/Impact_Free_58.pdf (retrieved on 21 January 2025), and Niesel, Dennis; Jelonnek, Stefan; Wilder, Nicolaus (in press), Gedankenexperimente als Methode pädagogischen Denkens – oder: Über die Notwendigkeit des Möglichen, Pädagogische Rundschau 2025.

How should the key question be understood? Key question 2 can also be read in two ways: as an empirical question in the sense that one could examine via surveys or content analyses what the **current realisation conditions for the use of AI** at universities are and to what extent they have changed in recent years, or as a philosophical question as to whether and for what reasons we should change these conditions and in what way. However, the latter presupposes that we have already come to the conclusion that we need to change individual goals and values. In both cases, we believe that the challenge lies in the fact that 'realisation condition' is an ambiguous term. We have already broken this down into technical, legal and didactic conditions. These three aspects can also be thought of as axes of a coordinate system. These open up a space in which it would have to be weighed up what guidelines and degrees of freedom in technology and (examination) law, as well as in the development of didactic knowledge and skills, are necessary or favourable (or obstructive) in order to achieve certain goals and values of higher education, and where.

What do we conclude from this? With regard to key question 2, we also tend to develop a thought experiment from the scenario, which could be played out in different variations. The coordinate system outlined above with the **three axes of technology, law and didactics** could be used for this purpose. What would be the consequences for teachers and students if a university were to impose strict guidelines and restrict freedom along these axes? What consequences (e.g. for the self-image of teachers and students) are conceivable if the guidelines were broad and the degree of freedom large? What if there were no guidelines at all? In addition, various combinations of guidelines and degrees of freedom of a technical, legal and didactic nature could be played out.

At this point, our own hypothesis (to be tested intellectually) is that the only sensible way for both teachers and students to deal with increasing complexity and growing dynamism in connection with AI would be to allow more autonomy as opposed to setting more guidelines.[13] However, this does not mean that the university leadership should withdraw; on the contrary, when it comes to the realisation conditions in connection with disruptive (technological) innovations such as AI, university leaders have a special responsibility, namely responsibility for orientation. In other words, the university leadership must position itself with regard to goals and values and align the **design of realisation conditions** accordingly. If a university wants to prioritise the integration of AI in all teaching, research and administrative processes, this requires different realisation conditions than in the case of universities for which technology-independent autonomy has priority. Arguments and good reasons can be found for both positions – which must be transparent for both current and aspiring students and teachers – and the entire continuum in between.

University leaders are therefore not only responsible for providing technical infrastructures, defining the legal framework and offering higher education didactic programmes, but also for offering orientation and encouraging positioning within the resulting scope of possibilities. This might take the form of artefacts, events, role models, etc., which provide **comprehensible guidelines** and also promote autonomy without overtaxing the individual or leaving

13 Wilder, Nicolaus; Weßels, Doris, KI und das Ende des Einheitslehrplans? Eine kritische Analyse, in: Weiterbildung. Zeitschrift für Grundlagen, Praxis und Trends 35 (2025) No. 6.

them without orientation. The optimistic expectation is that if this responsibility for orientation is realised, a shared attitude and identity can be developed to counteract the fears and overburdening of members of the university.

How can the goals and values of higher education be strengthened through the responsible use of AI?

What does our scenario say? Our scenario is already written in such a way that it outlines two different initiatives in which – assuming the current goals and values of university assessments – a responsible use of AI in the context of university assessments is strengthened. Business informatics scientist Sarah motivates her students to work with AI, but in such a way that they can independently (orally) represent the resulting product. By redesigning her assessment, she is aligning it with changing skills requirements and at the same time solving the observed problem of **non-transparent and uncontrolled use of AI**. Political scientist Yago encourages students to make a self-determined decision regarding the use of AI, but in both cases endeavours to ensure that writing skills are acquired because he considers them indispensable in his subject and for personal development. With the option of deciding for yourself whether or not to use AI, coupled with the requirement to document and reflect on the use of AI, he demands personal responsibility and transparency. He also solves the problem of non-transparent and uncontrolled AI deployment, but does so in a different way to Sarah. With both measures, the current goals and values of higher education are strengthened. If we had a scenario in which the goals and values of higher education were to change – for example in the direction of hybrid authorship – other measures would have to be taken in assessment design; at the same time, legal requirements would have to be adapted and the necessary technical infrastructures ensured. In this case, it can be assumed that the assessment culture would change profoundly.

How should the key question be understood? In our view, key question 3 should also be critically scrutinised (in retrospect) to determine whether goals and values can actually be strengthened in the same way as attitudes, for example. If this were possible, it would first be necessary to clarify which goals and values should be strengthened and among whom, and how this could be accomplished without becoming dogmatic. Answers to the first key question of whether or to what extent the values and goals of higher education are (or should be) reinterpreted as a result of AI should therefore be considered first. Our scenario in the context of university assessments refrains from this critical enquiry and to a certain extent bows to the currently omnipresent phrase **'responsible use of AI'.** Strictly speaking, this formulation assumes a premise that is currently rarely challenged, namely that AI is the greatest opportunity and danger for humanity, so we have to deal with it responsibly and then everything will be fine. As easy as this promise of salvation is to say, it is difficult to put into concrete terms. What exactly does it mean to use AI responsibly? What are the necessary conditions for the responsible use of AI? Can responsibility be taught? Who accepts responsibility? In retrospect, we would rephrase the question of how the goals and values of higher education can be strengthened through the responsible use of AI as follows: If one has compelling reasons for certain goals and values of higher education, what could be done to build, constructively develop and/or preserve them despite AI, with AI or because of AI?

What do we conclude from this? Here, too, we come to the same conclusion: fruitful discus-

sions could be stimulated above all with a targeted **thought-experimental variation of the scenario** in relation to the last key question in combination with the other two key questions. The closer the scenario remains to the current goals and values of higher education, the more incremental (and presumably easier to implement) the conclusions should be, both for the design of realisation conditions and for **measures to strengthen the responsible use of AI**. The more values and even goals of higher education undergo a transformation (for example to include hybrid authorship), the more we will have to leave familiar paths and presumably also change teaching, learning and assessment cultures – if the change in values and goals is desired and can be justified. With regard to the question of responsibility, however, we believe that a further key question should be added. In the working group, we have so far been working on the question of whether, and if so, to what extent we should change or further develop our teaching and assessment behaviour, including the associated goals and values, as a result of the existence of AI. AI is therefore taken as the starting point and the (responsible) use of the AI systems launched on the market is discussed. However, dealing with AI against the background of objectives such as promoting a critical-reflective attitude can only be realised in the immediate teaching-learning process. From this point of view, teaching should be the starting point and the question should be asked: how can universities' engagement with AI in teaching contribute to the value-orientated design of AI systems?

Our scenario focuses largely on the possibility of AI influencing the design of assessments and, implicitly, teaching (as a consequence of new forms of assessment) by encouraging critical questioning of previous routines. The scenario only indirectly suggests the reverse opportunity that the university's engagement with AI, for example on data protection (as a value), leads to at least the implementation of AI systems being aligned with current academic values (e.g. the data protection-compliant setup of large language models). However, the scenario could be expanded if we wanted to consider more precisely how thinking about our goals and values in higher education could stabilise or further develop the (generally congruent) social value framework, which at the same time should again be the framework for the development and use of AI systems.

Outlook: from integrity to responsibility

Our scenario indicates the extent to which AI is already influencing specific questions of assessment design and, as a result, teaching design. The scenario 'only' addresses the de facto or assumed delegation of assessment-relevant writing tasks by students to generative AI, which is incompatible with the (current) value of academic integrity. However, it is also possible to think through what happens when teachers hand over tasks and thus responsibility to AI in their didactic activities: whether it is the correction of assessed work, the compilation and preparation of learning and assessment content, or interaction with students in the process of acquiring or practising knowledge and skills (via AI-supported tutorial systems). What was largely left out of the consideration of our scenario under the three key questions of the working group, in this context, are questions of **roles in the network of relationships between teachers, students and AI** as well as the resulting changes in the requirements for assuming responsibility – an aspect that goes beyond our scenario and will therefore be briefly discussed in conclusion.

If we are increasingly using AI systems to achieve efficiency gains in the design and implementation of teaching, for example, or to offer students customised learning support for their learning path, we need to ask ourselves: what effect does this have on social relationships and on the roles of teachers and students? What role does AI play in this network of relationships? Which content-related and social parts of the teaching-learning relationship are handled by AI systems and which by the teacher, and for what reasons? Self-tests in our team suggest that **current chatbots** are **'anthropomorphised' in these relationships** and give us the feeling of being in a 'real' social exchange. If these experiences apply across the board, what are the consequences? How far is the path to a human-like robot that realises anthropomorphism? What influence does this have on our ideas about the relationship between human and machine and what does this mean for the university and social value base? And finally, who in this triad (teachers, students, AI) is actually (still) responsible for what?

The transfer and/or diffusion of responsibility in an educational context is not an entirely new phenomenon that has only been triggered by AI. If, for example, teachers define their role primarily as coaches and learning guides in order to strengthen the universally valued goal of **students' personal responsibility**, there is a real risk that they will 'buy their way out' of their own didactic responsibility by transferring it to students (for supposedly good reasons).[14] One might argue that students are adults and have to bear the responsibility anyway, but this attitude seems too simplistic because it ignores the question of what responsibility teachers have. AI-supported tutorial systems, as well as chatbots that generate teaching content at the touch of a button or review dissertations, have huge potential to further drive the delegation of responsibility – just as students could all too quickly be tempted to delegate their responsibility for written artefacts to generative AI.

Against this background, the roles and associated responsibilities in higher education should be discussed and then defined transparently. Of course, there are also values and goals behind such discussions and decisions, which should be explicitly stated. Depending on what these roles and responsibilities look like in the network of relationships between teachers, students and AI, the necessary conditions must also be reflected upon. Finally, the **providers of AI systems** should be included here if we take the above comments on the value-orientated design of AI systems seriously.

Recommendations for action

Numerous recommendations for action could be developed from the mental play-through and continuation of the scenario – depending on the focus of the content and the individual's own values and attitudes towards artificial intelligence (AI). The following recommendations are a highly condensed result of reflections on the authors' scenario. They are based on the fact that the academic values discussed at the beginning also (still) apply in the context of assessments, i.e. among other things: integrity, responsibility, transparency, authorship, self-determination, participation, equal opportunities, openness and willingness to experiment.

14 Ladenthin, Volker, Allgemeine Pädagogik, Baden-Baden 2022.

(1) University leadership teams assume responsibility for orientation: They position themselves on the topic of AI and (a) explain which goals and values should also apply to the design of assessments under 'AI conditions', (b) align technical, (examination-related) legal and didactic conditions in such a way that assessments are also possible with the use of AI, and (c) open up scope for action and decision-making for teachers and students in the context of assessments, or to put it another way: they promote the autonomy of action that teachers need in order to be able to adequately monitor the achievement of learning objectives.

(2) Teachers take responsibility for their actions: They proactively exploit the scope for action granted to them and (a) take into account both changing external competence requirements and academic requirements (as a result of AI) when making decisions in assessment design, (b) give students, where possible, freedom of choice with regard to the use of AI in assessment situations, which is jointly reflected upon in advance, (c) refrain from using AI in assessments if good reasons can be given for doing so, and (d) also recognise and use AI as a driver of changes in assessment routines and assessment culture that are necessary anyway.

(3) Responsibility is not lost in the student-teacher-AI relationship: Teachers and students (a) reflect on their roles in the network of relationships with AI, (b) clarify who is responsible for what in the assessment process and (c) include the entire assessment cycle: from the design of assessed tasks (teacher-AI interaction) to the completion of the assessment (student-AI interaction) to evaluation and feedback (teacher-student-AI interaction), without anthropomorphising the machine.

(4) Meta-recommendation: The extremely fast pace of technological developments in the field of AI poses major challenges, especially for those who have a sense of responsibility for others and at the same time want to actively shape the future and not just be a reactive pawn in the hands of outside interests. Our traditional procedures for dealing with these challenges at universities seem to have reached their limits here. Waiting for empirical evidence, for example – which has been decisive for the formulation of recommendations for action, at least in recent decades – simply takes too long. Although there are now initial results on the effects of using ChatGPT, for example, there are two version numbers between these statements and the present; the possibilities and limits of AI use have shifted considerably since then. The statements that are considered scientific, for example on 'hallucinations' of Large Language Models (LLMs), are thus only evident for the past, but not for the present. As a result, we are chronically lagging behind technological developments and are caught in a potentially paralysing cycle: if the only approach is the perpetual measurement of what is, this makes it impossible to shape the future. We see one way of breaking out of this cycle and directing the perspective of scientific reflection towards the future on the basis of given evidence – and this paper is an expression of this – in the thought-experimental design and play-through of possible scenarios. On the basis of empirical findings (relating to the past), projections of the future can then be drawn up, and on this basis questions of value and objectives – fundamental for recommendations for action and not accessible to empiricism per se – can be reflected upon, examined and, if necessary, renegotiated. New developments can be directly integrated and can be analysed in terms of their impact. Our methodological recommendation is therefore the reversal or drastic revitalisation of the

"strange shrinkage of utopian consciousness" already lamented by Adorno in 1964, which is necessary in order to shape the future in a reasoned, reflective and responsible manner.

Gabi Reinmann, Nicolaus Wilder & Antje Michel

3.2 Scenario II: Autonomy

Initial situation

The team of a part-time, digital certificate course at a university would like to use a chatbot and an AI-based evaluation tool to give students in online seminars quick feedback on their progress in the learning process and to evaluate it. The system is also intended to show how effectively and cost-efficiently the university in question is working. However, some teaching staff also have reservations about the use of such technologies, as they could be associated with considerable risks for students. For example, the system could potentially lead to **negative expectation pressure and stress among** students, and there are various issues relating to the protection of users' privacy.

Constructive vision

A short time later, the university leadership decides that it wants to invest in an **AI-based feedback and evaluation tool**. It opts for a system called AIvaluate, which is already in use at a number of universities in Europe and the USA. The system significantly changes the learning process of students and the teaching practice of teachers. The students are initially delighted with the innovative functions of the tool, which can be used to track learning successes to date and identify problem areas that they still need to address. Teachers, in turn, now have a much more detailed overview of students' strengths and weaknesses and can use this information to better support the individual learning process.

After a while, however, the euphoria towards AIvaluate fades. The frequently sent (push) messages with notifications from the system chatbot make students feel increasingly stressed. The wording of the messages also makes it increasingly clear to them how all-encompassing the monitoring by the system and therefore also by the teachers is. Theoretically, the information provided to teachers by the system can be used to deduce **details about students' everyday lives**. It tracks when and how long they are logged into the system, when they take a break and from where they connect to the webinars. In assessments, it is increasingly noticeable that individual students rely exclusively on AIvaluate's recommendation system, which is not always infallible, when it comes to their learning needs, instead of self-critically and independently identifying gaps in their learning.

Much of this comes to light through a **data leak** in which a student hacked into the system, created a profile as a lecturer and leaked the resulting information to a newspaper. Parallel to a controversial debate in the press, resistance is forming among students at the university at which AIvaluate was used in the certificate programmes. The students are calling for the contract with AIvaluate to be cancelled immediately and for the system to be completely banned from everyday university life. Some teachers are also increasingly showing

solidarity with the students. In order to avoid a massive loss of reputation and the prospect of legal action, the university leadership agrees to the demand after a while and terminates the business relationship with AIvaluate. In future, technical systems are to be examined and, if necessary, authorised by a joint committee of student, teacher and administrative representatives.

A year later, a new system is introduced, AIssist, which is more in line with the needs of those affected and takes into account **ethical considerations for the protection of privacy** and the promotion of student autonomy. The new system deliberately excludes monitoring functionalities (privacy by design) and focuses more on collaboration between students and feedback for teachers and university leaders (calculation of a general 'well-being factor' and anonymised assessments of the quality of the courses on offer). A chatbot is also provided, but it functions more as a planning and organisational assistant and only provides discreet recommendations for individual skills development at the explicit request of the students. The chatbot can also be switched on and off depending on the situation and use of the system is entirely voluntary. With regard to the learning progress of individual students, the system does not provide supposedly detailed information, as it has been found that AI systems convey a false sense of security in this area, but instead encourages students to **self-critically reflect on their own learning progress**. However, as AIssist integrates almost seamlessly into teaching and learning practice and simplifies many previously cumbersome tasks, the system is nevertheless used by almost all students on the certificate programmes and soon afterwards is offered and used on all other courses.

Discussion

The scenario described above can be discussed with regard to several values of higher education, including autonomy and self-determination in the area of personal development, of course, but also maturity and critical thinking skills in the area of (specialised) science.

If we assume that higher education aims to develop autonomous, critical thinkers who can control their own learning processes, assistance systems such as AIvaluate and AIssist offer learners **individualised support** through their guidance and feedback functionality and thus potentially increase autonomy, for example by reducing students' dependence on external help. Reflective use of information provided by automated learning status analyses can help students to develop a feeling for learning needs – particularly in terms of identifying previously unconscious gaps in the learning process. One danger is that by imposing too many requirements or comprehensively monitoring students, they in turn restrict autonomy. Part of an expanded **concept of autonomy in the age of AI** could therefore be to learn how to use assistance systems effectively for oneself without unlearning self-management in the learning process. A system that supposedly breaks down the user's gaps and learning needs in detail harbours the risk that users are no longer able to recognise their own work and learning needs independently.

Data protection is a prerequisite for autonomy. The introduction of systems such as AIvaluate shows that **data protection** is jeopardised by surveillance and data storage in particular. Privacy by design – i.e. embedding data protection principles in the technical design and development of systems – can therefore make an important contribution to guaranteeing

autonomy.

Education presupposes that a person learns for themselves; individualised support for learning processes is therefore considered an ideal. AI systems can make a contribution here by reacting individually to learning progress and needs and providing feedback. However, it is essential for the development of **maturity and critical thinking skills** that students learn the ability to independently recognise the next steps in their learning process. AI systems should provide support, but not take over complete control of the learning steps and restrict students in their responsibility for their own learning and thus in their education.

Other values that are affected by the scenario are transparency and fairness: the algorithms behind AI-based tools used in higher education must be transparent and fair. The lack of transparency and the data leak in the AIvaluate scenario show that there are considerable risks here. Universities should ensure that their AI systems are based on ethical principles in order to design trustworthy AI usage scenarios.

Recommendations for action

The following recommendations for responsible and constructive use of AI-based feedback and evaluation tools at universities arise from the scenario described:

(1) University leadership teams should define clear standards and guidelines for the use of AI-based systems and, in particular, (a) guarantee data protection and privacy, for example through transparent system selection that complies with data protection regulations (privacy by design), (b) make the use of these systems dependent on the consent of the students and (c) have the introduction and use of such systems critically monitored and evaluated by committees. They should also (d) enable teaching staff to deal competently and critically with AI-based systems through targeted measures, programmes and further training and actively promote their willingness to develop these skills.

(2) Teachers take an active role in integrating AI-based tools into their didactic practice by: (a) clearly defining together with students what information may be collected by AI, (b) communicating transparently how the collected data is used and interpreted and what conclusions can or cannot be drawn from it about the learning process, and (c) always using AI-supported feedback in a complementary and supportive manner, but never handing over responsibility for assessing learning progress or success to the AI.

(3) Protecting students from unwanted surveillance and stress-inducing monitoring mechanisms must be a top priority. To this end, it is necessary (a) to strictly limit the scope of data collection by AI systems to a necessary minimum, (b) to always process data anonymously in order to prevent conclusions from being drawn about individual behaviour, and (c) to transparently disclose to students what data is collected and processed, for what purpose and for how long.

(4) The autonomy and responsibility of students should be strengthened, not weakened,

by the use of AI-based systems. This is achieved through measures such as (a) the avoidance of supposedly omniscient, deterministic statements on learning progress and instead suggestions for self-critical reflection on one's own learning processes, (b) the promotion of conscious, self-determined use of the systems, in which students can set and pursue their own goals, and (c) clear communication that the appraisals provided by AI systems do not represent irrefutable truths, but should provide impulses for independent reflection and self-determined learning.

(5) University leaders and teaching staff should actively utilise the experience gained from the use of AI systems to initiate institutional learning processes. This includes (a) regularly carrying out systematic reflections and analyses of previous AI projects, (b) incorporating findings into future decisions on the use of AI and guidelines and (c) establishing an open error culture that promotes innovation and constructive-critical reflection when dealing with AI technologies.

Ulrike Tippe & Martin Wan

3.3 Scenario III: Research

Initial situation

Dr Jona Müller heads a research group on natural language processing (NLP) technologies at an institute for applied computer science. Among other things, her group is developing AI algorithms to help recognise **hate speech and incitement to violence in online communication data**.

In the past, Jona has worked on various occasions with large publishers and internet companies that use such algorithms to filter content on (social) media platforms. For the past year, she has also been cooperating with German police authorities for the first time. Technically speaking, the procedures to be developed are similar to those on which the research group has previously worked, but the new partners from the police change the framework conditions for the research to a considerable extent. Some time ago, an article appeared in a respected weekly newspaper that took a critical look at surveillance technologies in the age of AI. Her project was mentioned in particular as an example of **problematic cooperation between researchers and law enforcement agencies**. Since then, some of her colleagues from the institute have been critical of her. An important project team member has resigned, stating that he cannot reconcile his continued involvement with his conscience. It has so far been impossible to find a competent new employee for the project and she has the feeling that this could have something to do with the application context and poor image of the project as well as the lack of appropriately qualified people.

Jona is fundamentally aware of the responsibility that comes with such technologies. She is also beginning to systematically scrutinise the ethical implications of her research. She realises, for example, that an unthinking introduction of AI could potentially lead to **human rights violations**, for example through data sets that have a bias and thus lead to discriminatory results. In addition, there are no clear legal guidelines regulating the use of such

technologies in law enforcement. Jona feels increasingly overwhelmed and doesn't really know how to deal confidently with the criticisms levelled at her outside of technology development. The project ends a year later with mediocre results and an uneasy feeling in her stomach.

Constructive vision

So Jona actively looks for ways to better understand her unanswered questions. She is increasingly involved in workshops and networks in which **issues of algorithmic fairness**, transparency and the ethical and legal framework conditions of AI and surveillance technology are discussed. She enters into dialogue with ethicists, social scientists and lawyers, as well as representatives of human rights organisations and industry.

Together with such actors, she once again successfully applies for funding, this time for the development of **technologies for the prevention and prosecution of cybercrime**. The project is subject to a vote by the university's ethics committee, as is standard for medical projects today.

In contrast to the previous project, the team is now working within a setting of integrative research right from the start. This means that ethicists, social scientists, lawyers, **representatives of human rights organisations** and industry work together on the development of the technology and that technical components are developed iteratively, taking ethical, social and legal issues into account.

In one of the project consortium's workshops, Jona talks to Dr Fiona Schmidt, a media scientist. Fiona researches the social acceptance of digital technologies and science communication. She proposes to jointly develop an internet portal that discusses problematic aspects of AI technology in the context of surveillance and provides opportunities for citizen feedback and dialogue. The two manage to raise money for the development of the platform and the science communication project is now running in parallel with the technology development project. The internet portal is developing into a central hub for information and **discussion of perspectives on digital surveillance**. The collaborative content and online discussions inform the development of AI technology and accompanying measures such as the design of suitable governance structures that support the ethical and responsible use of the technology. The broad-based approach to AI technology design cannot mitigate all the possible downsides of digital surveillance. But this is not the main focus for Jona.

She particularly appreciates that the interactions with interdisciplinary partners and the interested public bring new questions to her that inspire her research and link her more to the practical level of AI use. She is also happy to be able to reconcile her work with her conscience and possibly contribute positive impulses for **more responsible technology design** to the scientific and public debate.

Recommendations for action

In sensitive areas of research and development, interdisciplinary teams, systematic risk analyses, the involvement of relevant stakeholders and active scientific communication

are important. These measures help to strengthen the trust of various social groups in the integrity of research and the responsible development and use of AI technologies. At the same time, they promote social acceptance of the technologies by creating transparency and opening up opportunities for participation.

(1) Assess ethical implications concretely, contextually and at an early stage ('ethics by design'): As a basic technology, AI can be used in many different ways, resulting in countless design options for the systems, the human-machine interaction and the context of use, which determine whether it is desirable or problematic to use it. This includes topics such as the protection of privacy, algorithmic fairness, transparency and access, human control, risks of dual use and manipulation as well as (ecological, social and economic) sustainability. In the individual specialist disciplines, ethical implications should be assessed at an early stage and, if possible, in relation to specific fields of application (ideally to an intended application) in order to minimise misuse and promote responsible use.

(2) Promote inter- and transdisciplinary cooperation: The challenges and new tasks associated with technology cannot be tackled by one discipline alone, but affect all scientific disciplines and require interdisciplinary cooperation.

At the same time, non-scientific and various social actors are becoming particularly relevant when it comes to technology, i.e. transdisciplinary cooperation is moving into focus here. This concerns cooperation with, for example, art and design, actors from the cultural sector, religious communities, politics, NGOs, associations and foundations.

Technological changes affect all areas of life, transform society and thus represent a task for society as a whole; conversely, science needs normative guidelines from society, for example on desired application scenarios or the nature of training data (with due consideration of bias). In transdisciplinary work, the object of research and research questions are developed jointly by academia and society and academia focuses its attention on current challenges in society. When academia and society learn together and develop proposals for the future, research methods and communication change. It is therefore important to try out new participatory formats and to find/learn a common language/basis for understanding (i.e. to clarify concepts, terms etc.).

(3) Risk analysis and stakeholder involvement in sensitive areas: The involvement of stakeholders in research and development processes offers both opportunities and challenges. External actors such as state institutions, companies or civil society actors can provide practical perspectives and resources, which can promote the social acceptance and ethical acceptability of technology design and application. At the same time, an area of tension arises, as the involvement of stakeholders can lead to problematic influence or even financial and infrastructural dependencies. Here, social participation and academic freedom must be weighed against each other.

Stakeholders change research by helping to shape topics and methods. Cooperations with large technology companies can promote innovative approaches, but harbour the risk of adapting the research agenda to external interests. Similar dynamics arise with state partners whose political objectives can influence research priorities.

Procedural approaches such as ethics by design[15], value-sensitive design and integrated research can help to exploit the advantages of such collaborations and minimise the risk of undesirable developments.

(4) Engage in science communication – communicate knowledge responsibly and adequately to those affected (social and political stakeholders, and possibly also the general public) and enable participation: Rapid technological developments lead to social uncertainty, a need for orientation, and polarisation between technological anxiety and technological euphoria. This uncertainty poses challenges for individuals, institutions and democracy, especially when people see their identity, agency or professional future threatened. The media sometimes exacerbate these tensions through exaggerated reporting, which is why scientifically sound and differentiated media work and science communication are essential.

Researchers have a special responsibility not only to generate knowledge, but also to communicate it appropriately and pluralistically. It is important not to leave the power of technological interpretation to large companies or individual actors alone, but to create opportunities for dialogue that enable participation. Participatory interfaces can offer society opportunities for feedback and co-design, thereby promoting transparency and critical thinking. It is important to emphasise that science communication tasks should not be carried out by researchers alone, but that supporting structures are available for dialogue and knowledge transfer.

Simon Hirsbrunner, Aljoscha Burchardt & Anna Puzio

15 Ethics by design means that ethical considerations are systematically integrated into the development process of technologies from the outset, instead of only being taken into account afterwards. Value-sensitive design pursues a similar approach by incorporating social values – such as fairness, transparency or data protection – into the design of technical systems so that they are developed not only functionally, but also in a value-orientated way.

4 Perspectives for action for universities in the context of AI: opportunities for skills development

The overarching question underlying all the scenarios described is: How should universities develop in the context of artificial intelligence? How must the skills of university members be developed in order to strengthen the goals of **(specialised) science, personal development and preparation for the labour market** as well as the associated values? Who are the actors involved, how can their area of responsibility be described and what resources are required?

The scenarios already suggest answers to these questions, but in this section they will be explored at a more general level. This is intended to provide universities with **support for their development**, which is becoming necessary in view of AI technologies. The design of degree programmes and teaching for skills development should be primarily oriented towards the goals and values of the university and should take technological developments into account.

4.1 Overarching recommendations for action

Identifying AI as a strategic topic

In order for universities to actively shape the far-reaching changes brought about by artificial intelligence, a targeted strategic approach is required. A key step is the **conscious identification of AI as a development topic at institutional level**. Universities should adopt a structured approach to ensure that teachers, researchers and students alike have the opportunity to engage with AI and are empowered to use it in a reflective manner.

The aim must be to raise awareness of the potential and challenges of AI. University-wide awareness-raising and training formats can contribute to this. In addition, **departments and organisational units** should be encouraged to develop their own guidelines for dealing with

AI that address specific, or subject-specific, challenges and objectives. It is important that universities choose an open, exploratory approach that leaves **room for different levels of integration** – from the creation of AI-free spaces to the comprehensive implementation of AI in academic processes. It is crucial that this process is organised autonomously and with a sound basis in terms of content, instead of being determined prematurely by external political and financial framework conditions.

The successful implementation of such concepts can be favoured by universities creating **spaces for interdisciplinary dialogue** that operate outside the usual committee and administrative structures. Formats such as university-wide discussion forums or themed innovation events could be an effective means of achieving this.

Successful university strategies for dealing with AI should be **further developed** in an **iterative process and regularly reviewed**. A conscious focus on university-specific needs, an open design methodology and the promotion of a discursive culture can help to realise a sustainable and reflective integration of AI at universities. It is crucial to strengthen the **self-determination of all university members** in dealing with AI in order not only to optimally utilise potential, but also to counter risks appropriately.

Promoting a general understanding of AI

All university members should develop a general understanding of artificial intelligence in order to ensure the responsible and didactically meaningful use of AI tools. The focus should not only be on the practical use of AI, but fundamentally on critical reflection on the underlying technology. This refers in particular to the **understanding of generative AI as statistical software**: language models – including so-called 'reasoning' or 'deep research' models – are still primarily based on statistical methods of pattern recognition. They generate texts by calculating word order probabilities from their training data and, if necessary, from additional external sources at runtime (referred to as Retrieval-Augmented Generation, RAG), in the case of advanced models combined with approaches of symbolic, i.e. rule-based AI.[16]

This is important because the ELIZA effect described by Joseph Weizenbaum,[17] the danger of self-deception due to the illusion that AI possesses genuine understanding, is particularly high in modern AI models. Because in most cases the generated output appears to make sense semantically and in terms of content, we are inclined to trust the system and **assume that it has human-like understanding**. Nevertheless, AI systems are and will remain software that works according to the input-processing-output principle. The fact that the processing of AI systems is highly complex and no longer manually comprehensible due to the sheer mass of weights and parameters does not mean that what happens in AI systems

16 Cf. Se, Ksenia, How Do Agents Plan and Reason?, available at: https://huggingface.co/blog/Kseniase/reasonplan (retrieved on 11 March 2025).

17 Cf. Wan, Martin, Kann KI bewusst sein? Anmerkungen zur Diskussion um LaMDA (July 2022), available at: https://digiethics.org/2022/07/22/kann-ki-bewusst-sein-anmerkungen-zur-diskussion-um-lamda/ (retrieved on 27 February 2025).

is inexplicable in principle. However, the lack of transparency due to the sheer size and the associated impenetrability of the weights of AI models must be part of the problem awareness of their users.

This also goes hand in hand with an **awareness of the sources of error** in AI-generated outputs. As statistical models, AI systems do not 'understand' meanings and contexts, but replicate content solely on the basis of statistical probabilities. Instead of summaries, language models therefore often only generate abridged texts that omit essential information.[18] For academic work, however, an understanding of contexts across different sources as well as **source criticism** is essential. While language models sometimes invent sources that look correct but do not exist,[19] current 'reasoning' and 'deep research' models often fail to differentiate between reputable and dubious sources or do not prove with the source given what they claim to prove.[20] Even with expected future improvements to the models, for example through access to a greater number of more reputable sources, the ability to apply critical judgement to the outputs and the sources given remains as important as ever.

Understanding AI models as statistical software is therefore a prerequisite for the assessment not only of specific AI outputs, but more broadly of appropriate and inappropriate or potentially dangerous areas of application for AI. Those who understand that AI does not 'understand' can better recognise the dangers of risky applications of AI, for example in fully automated decision-making processes. **AI exploration skills** can also be helpful here, i.e. the testing of AI models and the ability to push them to their limits – for example through skilful prompting. Regular use of the relevant models can also help to train creative AI skills, i.e. the use of AI to support human creativity. It would be desirable to have comprehensive, reliable and data-efficient access to common AI models at German universities.

Practising source and ideological criticism

The traditional academic expertise of source criticism is becoming even more important

18 Cf. Wierda, Gerben, When ChatGPT summarises, it actually does nothing of the kind, available at: https://ea.rna.nl/2024/05/27/when-chatgpt-summarises-it-actually-does-nothing-of-the-kind/ (retrieved on 28 February 2025).

19 Cf. Wan, Martin, Warum ChatGPT nicht das Ende des akademischen Schreibens bedeutet, available at: https://digiethics.org/2023/01/03/warum-chatgpt-nicht-das-ende-des-akademischen-schreibens-bedeutet/ (retrieved on 28 Feb 2025) and Ders, Warum ChatGPT (auch weiterhin) nicht das Ende des akademischen Schreibens bedeutet, available at: https://digiethics.org/2024/01/23/warum-chatgpt-auch-weiterhin-nicht-das-ende-des-akademischen-schreibens-bedeutet/ (retrieved on 28 February 2025).

20 Cf. Jones, Nicola, OpenAI's 'deep research' tool: is it useful for scientists? Available at: https://www.nature.com/articles/d41586-025-00377-9 (retrieved on 28 February 2025) and Major, Eddie, RAGs, Reasoning and Deep Research: What's new in AI and what might it mean for teaching in 2025? Available at: https://www.adelaide.edu.au/learning/news/list/2025/02/21/rags-reasoning-and-deep-research-whats-new-in-ai-and-what-might-it-mean-for (retrieved on 28 February 2025).

thanks to AI. Commercial AI providers in particular show little transparency with regard to their training data. This lack of transparency continues with the content generated by AI models. In this context, **source criticism** also includes an awareness of model training using copyrighted content and its subsequent (partial) reproduction. 'Reasoning' and 'deep research' models automatically decide which sources they use to generate their texts, so users must be able to make judgements about their quality.

Awareness of possible biases in the training data also requires the ability to **apply ideological criticism**. Depending on the provider, training corpus and fine-tuning, models exhibit different biases and distortions.[21] There should also be an awareness of the socially and ecologically problematic aspects of large AI models (exploitation of click workers in the Global South, resource consumption of the necessary hardware) in order to evaluate the use of AI from an ethical perspective.

The area of ideological criticism also includes **awareness of the PR of AI companies compared with actual progress in AI research**. Marketing terms such as AGI (Artificial General Intelligence) or 'emergence' promote the mystification of AI and the perception of AI models as knowledge and decision-making machines with human-like capabilities.[22] At the same time, it has not yet been convincingly demonstrated that AI is more than just sophisticated software[23], nor what conditions need to be met for software to develop human-like abilities. There is no philosophical or scientific consensus on the exact functioning of the human

21 The Chinese model DeepSeek, even in the uncensored version, answers from a perspective favourable to the Chinese government, cf. Lee Myers, Steven, DeepSeek's Answers Include Chinese Propaganda, Researchers Say, in: New York Times (online), available at: https://www.nytimes.com/2025/01/31/technology/deepseek-chinese-propaganda.html (retrieved on 28 February 2025). Elon Musk's Grok 3 betrays itself when asked about the biggest disseminators of fake news, the answer is said to be difficult because the model is allegedly instructed to keep Musk and Trump out of the answer, cf. Hanfeld, Michael, Die KI von Elon Musk verplappert sich, in: FAZ (online), available at: https://www.faz.net/aktuell/feuilleton/grok-3-laesst-musk-und-trump-bei-frage-nach-fake-news-weg-110319792.html (retrieved on 28 February 2025).

22 Cf. also Weizenbaum, Joseph: "Most of the harm that the computer could potentially cause depends less on what the computer can or cannot actually do and more on the characteristics that the public ascribes to the computer. The non-specialist has no choice but to attribute to the computer the characteristics that are conveyed to him by the propaganda of the computer community, which is reinforced by the press. The computer scientist therefore has the enormous responsibility to be modest in his claims." (idem., Albtraum Computer. Ist das menschliche Gehirn nur eine Maschine aus Fleisch? In: ZEIT 27 (1972) No. 3 of 21 January 1972, 43.)

23 Cf. Butlin, Patrick et al, Consciousness in Artificial Intelligence: Insights from the Science of Consciousness, available at: https://doi.org/10.48550/arXiv.2308.08708 (retrieved on 28 February 2028).

brain[24], on the question of consciousness or the answer to the qualia problem.[25] The question of whether AI will one day have human-like abilities cannot therefore be answered purely in technological terms, but requires a philosophical positioning within the mind-brain debate.[26]

Developing awareness of unintentional deskilling

As with almost all tools and technical aids, one risk of using AI is deskilling, i.e. the **potential loss of human expertise** through its improper and unconsidered use, with individual and collective consequences[27] which initial studies have already pointed out.[28] This is problematic because (a) AI systems can fail, (b) humans should still be able to monitor AI systems in the future and must therefore be able to perform the relevant tasks themselves, and (c) the outsourcing of tasks and social interactions to AI could lead to an **increasing loss of control**.[29] The phenomenon is primarily discussed in the context of assessments (potential attempts to cheat through the use of AI), scientific papers (outsourcing of writing and thinking processes to AI) and research (lower research quality, less diversity and declining scientific

24 Cf. Monya, Hannah et al, Das Manifest. Elf führende Neurowissenschaftler über Gegenwart und Zukunft der Hirnforschung, in: Gehirn & Geist 3 (2004) No. 6, 30-37.

25 The qualia problem refers to the question of the relationship between subjective-phenomenal perception and mental states, cf. also Chalmers, David J., Facing Up to the Problem of Consciousness, in: Journal of Consciousness Studies 2 (1995) No. 3, 200-219.

26 Both the former Google engineer Blake Lemoine, who assumed that the LaMDA language model was conscious, and the futurist Ray Kurzweil answer the question of why a machine should be believed with regard to a claimed consciousness with mere beliefs: "benefit of the doubt" (Lemoine) and "leap of faith" (Kurzweil), see Wan, Martin, Kann KI bewusst sein?

27 Cf. German Ethics Council, Mensch und Maschine – Herausforderungen durch Künstliche Intelligenz, Opinion of 20 March 2023, available at https://www.ethikrat.org/fileadmin/Publikationen/Stellungnahmen/deutsch/stellungnahme-mensch-und-maschine.pdf (retrieved on 28 February 2025), 353ff.

28 Cf. Lee, Hao-Ping et al, The Impact of Generative AI on Critical Thinking: Self-Reported Reductions in Cognitive Effort and Confidence Effects From a Survey of Knowledge Workers, available at: https://www.microsoft.com/en-us/research/wp-content/uploads/2025/01/lee_2025_ai_critical_thinking_survey.pdf (retrieved on 28 February 2025) and Gerlich, Michael, AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking, in: Societies 15 (2025) No. 6, available at: https://doi.org/10.3390/soc15010006 (retrieved on 28 February 2025).

29 Cf. Reinmann, Gabi, Deskilling durch Künstliche Intelligenz? Potenzielle Kompetenzverluste als Herausforderung für die Hochschuldidaktik (= HFD Discussion Paper No. 25 of October 2023), available at: https://hochschulforumdigitalisierung.de/sites/default/files/dateien/HFD_DP_25_Deskilling.pdf (retrieved on 28 February 2025), 7.

integrity).[30] The unintentional and ultimately undesirable loss of skills referred to here must be distinguished from long-term shifts in skills profiles, for which there are good reasons.

Occasional attempts such as that of the Prague University of Economics and Business to abolish bachelor's theses in the face of generative AI have attracted media attention.[31] Is such a shift in skills profiles already taking place here? Academic writing, as an individual process of scholarly engagement with a topic, continues to be an important part of skills acquisition for students and the achievement of academic qualifications. Written assignments and qualification theses should demonstrate the **ability to carry out independent academic work** – where an academic question can be dealt with sufficiently and originally in a written presentation to academic standards; the documentation and justification of the research design, the chosen methods and the individual process of insight come before the results. In this respect, the Prague initiative can be seen as an example of possibly unintentional deskilling, unless alternative forms are found that enable the skills acquisition described above in a similarly effective way.

Rules for the use of AI in studies, teaching and assessments on the part of the university can create a binding framework for the individual actions of teachers and students. However, such a framework must be sufficiently discussed and explained. The **self-determination of all stakeholders at the university** in dealing with generative AI must be strengthened in order to counteract any potential loss of expertise. In addition to a general understanding of the technology, this also involves an attitude of not allowing oneself to be controlled and steered by it. It is also necessary to discuss which **AI-independent basic skills** should be taught. This cannot be determined on an interdisciplinary basis and is a genuine scientific didactic challenge. **Human-centred interaction between human and machine** reinforces the basic idea that technical features and human capabilities can complement each other – coupled with an awareness of what people may and should be permitted to unlearn in the process. In this way, awareness of unintentional deskilling can be created without sacrificing the benefits of AI.[32]

Building a basic legal understanding

The first requirements of the EU AI Act have been in force since 2 February 2025. Since then, universities have been obliged to ensure that all stakeholders can "assess legally and tech-

30 Cf. ibid., 8.

31 Cf. Bager, Jo, ChatGPT & Co: Uni in Prag schafft Bachelorarbeiten ab, in: Heise Online, available at: https://www.heise.de/news/ChatGPT-Co-Uni-schafft-Bachelorarbeiten-ab-9546851.html (retrieved on 7 March 2025).

32 Cf. Reinmann, Deskilling, 12 f.

nically what it means to use AI".[33] Questions that universities must be able to answer are: "**What is an AI system within the meaning of the AI Act?** Who does the AI law apply to? What special considerations apply to universities as public bodies when using AI? When is AI permitted but highly risky? How can high-risk applications be recognised? Under what specific conditions may high-risk AI be used? When is AI prohibited?"[34]

The German Rectors' Conference will provide support services and guidelines for this purpose.

Avoiding an atmosphere of mistrust

In the context of generative AI, the topic of "'exam fraud' is often one-sidedly emphasised. Exam fraud is not a new phenomenon and is not necessarily becoming more frequent due to new technologies. Universities have always had to rely on assignments being written independently and without the help of other people or ghostwriters. **A generalised culture of mistrust towards students does not do justice to the majority of those who use AI honestly and creatively.** Instead, a constructive integration of digital tools should be encouraged, which strengthens the existing trust between teachers and learners and improves the quality of learning.

In addition, **assessment formats** should be regularly reviewed and, if necessary, adapted. Skills-orientated tasks that stimulate a deeper engagement with the learning content, and work that demonstrates reflection and transfer, can only be inadequately performed by generative AI alone. A combination of trust, clear communication of assessment requirements and targeted didactic integration of AI helps to sustainably improve the quality of learning and minimise the risk of cheating.

Developing visions to be able to make reflected decisions

Reflecting on our own work as a working group towards the end made it increasingly clear to us how important it is to have the ability to detach oneself – at least temporarily and at appropriate times – from what is factual and to think about alternative possibilities in a reasoned manner when dealing with AI in the university context. Here we refer to the development of future scenarios and outlines for higher education with AI. The approach we have chosen here, namely to approach alternative possibilities by developing constructive scenarios, has proved to be extremely fruitful for us – also in discussions with experts – and a method of **creating orientation and** (re)establishing **agency and the power to act** where one all too often feels powerless in view of the force and speed of AI development. Being able to develop educational scenarios through to social outlines – and this is not about quixotic pipe dreams, but about **constructive, directly usable or thought-sharpening visions** – could

33 Schwartmann, Rolf, cited in: AA.VV., Erste Vorgaben zum Einsatz von KI greifen, in: Forschung und Lehre (online), available at: https://www.forschung-und-lehre.de/recht/erste-vorgaben-zum-einsatz-von-ki-greifen-6878 (retrieved on 28 February 2025).

34 Ibid.

be a hitherto neglected skill for dealing with AI in a suitable way in order to overcome what Mittelstraß has already criticised about the education system and society as a whole in the context of digitalisation: namely that progress[35] no longer belongs to us, but we belong to progress. However, we seem to have lost this ability in large parts of society and especially in academia. As early as 1964, Adorno described this as the "strange shrinkage of utopian consciousness". The current multitude of guidelines and recommendations for action on the use of AI[36] – mainly reactive, presumably only effective in the short term, sometimes even factually questionable and then disorientating (instead of orientating), but above all unimaginative – could be an indication that supports Adorno's diagnosis today.

Scenario work, as we have prioritised for this paper, is *one* way of approaching the task of orientation scientifically and thus methodically in times of crisis. Another possibility is the transition from the construction of factual scenarios to counterfactual thought experiments.[37] This approach also seems promising, although it is even more methodologically demanding due to its counterfactuality. Restoring the ability to think and argue in terms of possibilities seems to us to be one of the central challenges in the education system in order to be able to orientate oneself and shape society, and not merely react to the ideas envisaged by commercial or ideological actors in the face of all possible crises. This is the only way to develop **differentiated and context-dependent concepts for dealing with AI** – beyond naïve euphoria and dystopian fatalism.

Thinking and arguing in terms of possibilities moves away from generalised, ideological 'if-or-if-not' questions towards 'how' questions that need to be discussed against the background of goals and values, as we have attempted to do in this paper. We invite you to use our scenarios as a starting point for **your own examination of the possibilities of AI** in the relevant configurations or, affirmatively or critically inspired, to design your own scenarios and use them as an opportunity for discussion. At the same time, interacting opposite directions, paradoxical side effects, in other words risks of all kinds, must always be taken into account and negotiated, as we have also done in various ways in this paper. In order to do justice to the complexity of AI as a circular technological-social subject, differentiated interdisciplinary research is ultimately required, which, in addition to an empirical description of concrete relationships, can also design possibilities and futures.

---

35 Cf. Mittelstraß, Jürgen, Bildung in einer Wissensgesellschaft, in: heiEDUCATION Journal 3 (2019), 21-36, 29, available at: https://doi.org/10.17885/heiup.heied.2019.3.23942 (retrieved on 11 March 2025).

36 Cf. also Weßels, Doris, Schluss mit absurden KI-Regeln, in: ZEIT 80 (2025) No. 3 of 16 January 2025, 31, available at: https://www.zeit.de/2025/03/kuenstliche-intelligenz-studium-hochschulen-regeln (retrieved on 11 March 2025).

37 Cf. Gehring, Petra, Rechtspolitische Bemessung möglicher gesellschaftlicher Gefahren digitaler Technologien? Zwei Gedankenexperimente mit anschließender Erwägung, in: Schreiber, Gerhard; Ohly, Lukas (eds.), KI:Text. Diskurse über KI-Textgeneratoren, Berlin 2024, 355-359, available at: https://doi.org/10.1515/9783111351490-022 (retrieved on 11 March 2025), and Reinmann, Gabi, Hüter, Kümmerer, Vormund? Eine Universität der Avatare: Ein Gedankenexperiment, in: Impact Free No. 61 (2025), available at: https://gabi-reinmann.de/wp-content/uploads/2025/01/Impact_Free_61.pdf (retrieved on 11 March 2025).

5 Literature

AA.VV., Erste Vorgaben zum Einsatz von KI greifen, in: Forschung und Lehre (online), available at: https://www.forschung-und-lehre.de/recht/erste-vorgaben-zum-einsatz-von-ki-greifen-6878 (retrieved on 28 February 2025).

Bager, Jo, ChatGPT & Co.: Uni in Prag schafft Bachelorarbeiten ab, in: Heise Online, available at: https://www.heise.de/news/ChatGPT-Co-Uni-schafft-Bachelorarbeiten-ab-9546851.html (retrieved on 7 March 2025).

Butlin, Patrick et al, Consciousness in Artificial Intelligence: Insights from the Science of Consciousness, available at: https://doi.org/10.48550/arXiv.2308.08708 (retrieved on 28 February 2028).

Chalmers, David J., Facing Up to the Problem of Consciousness, in: Journal of Consciousness Studies 2 (1995) No. 3, 200-219.

Deci, Edward L.; Ryan, Richard M., Self-determination theory: A macrotheory of human motivation, development, and health, in: Canadian Psychology 49 (2008) No. 3, 182-185

German Research Foundation, Guidelines for Safeguarding Good Research Practice. Code of Conduct, Bonn 2022 (corrected version 1.1), retrieved on 21 January 2025 at: https://zenodo.org/records/6472827/files/kodex_leitlinien_gwp_dfg.1.1.pdf.

German Ethics Council, Mensch und Maschine – Herausforderungen durch Künstliche Intelligenz, opinion of 20 March 2023, available at https://www.ethikrat.org/fileadmin/Publikationen/Stellungnahmen/deutsch/stellungnahme-mensch-und-maschine.pdf (retrieved on 28 February 2025).

Ehlers, Ulf-Daniel, Future Skills. Lernen der Zukunft – Hochschule der Zukunft, Wiesbaden 2020.

Gehring, Petra, Rechtspolitische Bemessung möglicher gesellschaftlicher Gefahren digitaler Technologien? Zwei Gedankenexperimente mit anschließender Erwägung, in: Schreiber, Gerhard; Ohly, Lukas (eds.), KI:Text. Diskurse über KI-Textgeneratoren, Berlin 2024, 355-359, available at: https://doi.org/10.1515/9783111351490-022 (retrieved on 11 March 2025).

Gerlich, Michael, AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking, in: Societies 15 (2025) No. 6, available at: https://doi.org/10.3390/soc15010006 (retrieved on 28 February 2025).

Hanfeld, Michael, Die KI von Elon Musk verplappert sich, in: FAZ (online), available at: https://www.faz.net/aktuell/feuilleton/grok-3-laesst-musk-und-trump-bei-frage-nach-fake-news-weg-110319792.html (retrieved on 28 February 2025).

Huber, Ludwig, Hochschuldidaktik als Theorie der Bildung und Ausbildung, in: idem. (ed.), Ausbildung und Sozialisation in der Hochschule (= Enzyklopädie Erziehungswissenschaften 10), Stuttgart 1983, 127-129.

Jones, Nicola, OpenAI's 'deep research' tool: is it useful for scientists? Available at: https://www.nature.com/articles/d41586-025-00377-9 (retrieved on 28 February 2025).

Ladenthin, Volker, Allgemeine Pädagogik, Baden-Baden 2022.

Lee, Hao-Ping et al, The Impact of Generative AI on Critical Thinking: Self-Reported Reductions in Cognitive Effort and Confidence Effects From a Survey of Knowledge Workers, available at: https://www.microsoft.com/en-us/research/wp-content/uploads/2025/01/lee_2025_ai_critical_thinking_survey.pdf (retrieved on 28 February 2025).

Lee Myers, Steven, DeepSeek's Answers Include Chinese Propaganda, Researchers Say, in: New York Times (online), available at: https://www.nytimes.com/2025/01/31/technology/deepseek-chinese-propaganda.html (retrieved on 28 February 2025).

Major, Eddie, RAGs, Reasoning and Deep Research: What's new in AI and what might it mean for teaching in 2025? Available at: https://www.adelaide.edu.au/learning/news/list/2025/02/21/rags-reasoning-and-deep-research-whats-new-in-ai-and-what-might-it-mean-for (retrieved on 28 February 2025).

Miller, Riel, Futures Literacy. A hybrid strategic scenario method, in: Futures 39 (2007) No. 4, 341-662.

Mittelstraß, Jürgen, Bildung in einer Wissensgesellschaft, in: heiEDUCATION Journal 3 (2019), 21-36, 29, available at: https://doi.org/10.17885/heiup.heied.2019.3.23942 (retrieved on 11 March 2025).

Monya, Hannah et al, The Manifesto. Elf führende Neurowissenschaftler über Gegenwart und Zukunft der Hirnforschung, in: Gehirn & Geist 3 (2004) No. 6, 30-37.

Nathan, Lisa P. et al, Value scenarios. A technique for envisioning systemic effects of new technologies, in: CHI '07 Extended Abstracts on Human Factors in Computing Systems, San Jose 2007, 2585-2590, available at: https://dl.acm.org/doi/10.1145/1240866.1241046 (retrieved on 17 February 2025).

Niesel, Dennis; Jelonnek, Stefan; Wilder, Nicolaus (in print), Gedankenexperimente als Methode pädagogischen Denkens – oder: Über die Notwendigkeit des Möglichen, Pädagogische Rundschau 2025.

Reinmann, Gabi, Deskilling durch Künstliche Intelligenz? Potenzielle Kompetenzverluste als Herausforderung für die Hochschuldidaktik (= HFD Discussion Paper No. 25 of Oct. 2023), available at: https://hochschulforumdigitalisierung.de/sites/default/files/dateien/HFD_DP_25_Deskilling.pdf (retrieved on 28 February 2025), 7.

Idem., Gedankenexperimente als bildungstheoretisches Instrument in der Forschung zu Künstlicher Intelligenz im Hochschulkontext, in: Impact Free 56 (2014). available at https://gabi-reinmann.de/wp-content/uploads/2024/09/Impact_Free_58.pdf (retrieved on 21 January 2025).

Idem., Hüter, Kümmerer, Vormund? Eine Universität der Avatare: Ein Gedankenexperiment, in: Impact Free 61 (2025), available at: https://gabi-reinmann.de/wp-content/uploads/2025/01/Impact_Free_61.pdf (retrieved on 11 March 2025).

Idem.; Watanabe, Alice, KI in der universitären Lehre, in: Schreiber, Gerhard; Ohly, Lukas (eds.), KI: Text. Diskurse über KI-Textgeneratoren, Berlin 2024, 29-46.

Rosson, Mary Beth; Carrol John M., Scenario-based design, in: Sears, Andrew; Jacko, Julie A. (eds.), The Human-Computer Interaction Handbook, New York 22008, 1041-1060; De Jouvenel, Hugues, A brief methodological guide to scenario building, in: Technical Forecasting and Social Change 65 (2000) No. 1, 37-48.

Se, Ksenia, How Do Agents Plan and Reason?, available at: https://huggingface.co/blog/Kseniase/reasonplan (retrieved on 11 March 2025).

Wan, Martin, Kann KI bewusst sein? Anmerkungen zur Diskussion um LaMDA (July 2022), available at: https://digiethics.org/2022/07/22/kann-ki-bewusst-sein-anmerkungen-zur-diskussion-um-lamda/ (retrieved on 27 February 2025).

Idem., Warum ChatGPT nicht das Ende des akademischen Schreibens bedeutet, available at: https://digiethics.org/2023/01/03/warum-chatgpt-nicht-das-ende-des-akademischen-schreibens-bedeutet/ (retrieved on 28 February 2025).

Idem., Warum ChatGPT (auch weiterhin) nicht das Ende des akademischen Schreibens bedeutet, available at: https://digiethics.org/2024/01/23/warum-chatgpt-auch-weiterhin-nicht-das-ende-des-akademischen-schreibens-bedeutet/ (retrieved on 28 February 2025).

Weizenbaum, Joseph, Albtraum Computer. Ist das menschliche Gehirn nur eine Maschine aus Fleisch? In: ZEIT 27 (1972) No. 3 of 21 January 1972, 43.

Weßels, Doris, Schluss mit absurden KI-Regeln, in: ZEIT 80 (2025) No. 3 of 16 January 2025, 31, available at: https://www.zeit.de/2025/03/kuenstliche-intelligenz-studium-hochschulen-regeln (retrieved on 11 March 2025).

Wierda, Gerben, When ChatGPT summarises, it actually does nothing of the kind, available at: https://ea.rna.nl/2024/05/27/when-chatgpt-summarises-it-actually-does-nothing-of-the-kind/ (retrieved on 28 February 2025).

Wilder, Nicolaus; Weßels, Doris, KI und das Ende des Einheitslehrplans? Eine kritische Analyse, in: Weiterbildung. Zeitschrift für Grundlagen, Praxis und Trends 35 (2025) No. 6.

German Council of Science and Humanities, Recommendations on the relationship between higher education and the labour market. Second part of the recommendation on the qualification of skilled workers (Drs. 4925-15), Bielefeld 2015, retrieved on 21 January 2025 at: https://www.wissenschaftsrat.de/download/archiv/4925-15.pdf.

About this publication

HFD working papers reflect the opinion of the respective authors.
The HFD therefore does not endorse the statements made in this paper.

ISSN (Online) 2365-7081; 11. Year

Citation reference
Working Group "Artificial Intelligence: Essential competences at universities" (2025). Artificial intelligence: Principles for action in university teaching. Working paper no. 86. Berlin: Higher Education Forum on Digitalisation.

Publisher
Office of the Higher Education Forum on Digitisation at Stifterverband für die Deutsche Wissenschaft e.V.
Capital City Office • Pariser Platz 6 • 10117 Berlin • T +49 (0)30 322982-520
info@hochschulforumdigitalisierung.de

Editor
WG "Artificial Intelligence: Essential competences at universities"

Publishing house
Edition Stifterverband – Verwaltungsgesellschaft für Wissenschaftspflege mbH
Baedekerstraße 1 • 45128 Essen • T +49 (0)201 8401-0 • mail@stifterverband.de

Layout
Typesetting: Julia Rosche
Template: TAU GmbH • Köpenicker Straße 154a • 10997 Berlin

The Higher Education Forum on Digitalisation is a joint project of Stifterverband, the CHE Centre for Higher Education and the German Rectors' Conference. It is funded by the Federal Ministry of Education and Research.

www.hochschulforumdigitalisierung.de/en

STIFTERVERBAND
CHE Centrum für Hochschulentwicklung
HRK Hochschulrektorenkonferenz
GEFÖRDERT VOM Bundesministerium für Bildung und Forschung