UTRECHT UNIVERSITY

Department of Information and Computing Science

Human Computer Interaction

Crafting Tomorrow's Education: Exploring Visual Generative AI Integration in Academic Courses

First examiner:

Dr.Almila Akdag

Candidate:

Erdem Ata Yildirim

Second examiner:

Dr.Sergey Sosnovsky

November 15, 2024

Contents

1. Introduction

In January 2023, ChatGPT reached over 100 million users, making it the fastest-growing consumer application until this date [1]. ChatGPT is a generative tool, which can be described as an Artificial Intelligence (AI) system that uses existing media to create new, plausible media [2]. Unlike other AI systems that produce decisions or descriptions, generative AI (GenAI) systems can produce content that includes images, texts, music, video, and other forms of design. GenAI is an innovative tool in machine learning, enabling systems to create content across various domains. Traditional AI systems are mostly focused on predictions and classifications, while GenAI creates new, original content based on the patterns from training data [3]. The products that derive from these tools are a lot like those that are produced by humans. However, it is not yet clear how humans can control and interact with the capabilities of these tools. Also, it is especially not clear what kinds of collaboration patterns will emerge when creative humans and creative technologies work together [2].

1.1 Visual GenAI

Next to tools like ChatGPT, which can be categorized as textual GenAI, there is also visual GenAI. Visual GenAI can create images and videos in various ways, depending on the input of the user. One example is text-to-image diffusion models (DMs). DMs are used to create and edit content, making it possible to generate a diverse range of images with great quality that matches what the user has prompted [4]. With DMs, it is also possible to give an image as input and create different variations of that image, as shown in Figure 1.1.

Figure 1.1: The input of a sketch and the resulting variations generated by providing text input [4].

1.2 Applying GenAI in Education

GenAI could be deployed in various domains to improve and create new methods that stimulate the overall experience. One of those domains is education, where these tools can be used for the development of new educational methods.

Traditional teaching methods often follow a model where the teacher is the primary source of knowledge. Information is primarily delivered through lectures, textbooks, and notes, with students expected to absorb and process this material independently. While this method has proven effective, it may not always sufficiently motivate or engage students, especially in a world where technology is rapidly evolving [5].

A challenge with traditional methods is their reliance on passive learning. Students listen to lectures, read texts, and are tested on their ability to recall or apply information, which can feel detached or abstract. Furthermore, traditional methods may struggle to address the varying learning styles in a classroom. Some students may excel through verbal instruction, while others need more visual or hands-on engagement to understand the concepts taught [6].

Visual GenAI has the potential to transform teaching by making learning more interactive, dynamic, and visually engaging. Teachers can use AI-generated images to create custom visuals that bring lessons to life, improving understanding of abstract ideas and historical contexts. This technology stimulates creative thinking and visual literacy, making learning more accessible and stimulating better interaction between teachers and students.

It also enables the creation of interactive content, allowing students to explore scenarios or visualize concepts in real-time. This is especially valuable in subjects like history, biology, or geography, where complex concepts such as the human body, ecosystems, or climate patterns can be visualized. For example, in history classes, AI-generated visuals can illustrate the construction of the pyramids, providing an engaging way to explain the process.

In conclusion, integrating visual GenAI into classrooms offers a way to improve student engagement and understanding. It enables teachers to create more personalized and visually stimulating learning experiences, making education more accessible and enjoyable for students across various subjects.

1.2.1 GenAI in Education: Pros and Cons

The launch and use of GenAI have raised many concerns across different sectors, especially within the academic sector. Academics have urged universities to develop new forms of assessment after the launch of GenAI tools. GenAI can handle some traditional tasks, particularly written assignments. However, these tools also bring unique challenges [7].

Student cheating is not a new phenomenon since the introduction of ChatGPT. However, it has brought a completely new challenge for the academic world. It is challenging for teachers to detect material that has been generated by GenAI, especially in comparison to the old cheating methods of students such as plagiarism [8]. The legitimacy of evaluations and the fairness of grading are also becoming doubtful due to the introduction of GenAI, as some students will not be graded based on their own skills and knowledge [7].

GenAI provides opportunities to enrich and innovate education by developing new methods that were difficult or even impossible without this technology. One of the primary advantages of GenAI is its ability to personalize learning experiences. By automatically generating customized content, such as quizzes, interactive simulations, or visual aids, GenAI can address the diverse needs of students and support their learning with per-

sonalized feedback. This improves their learning process and allows each student to progress at their own pace [1].

However, the implementation of GenAI in education should always be goal-oriented, focused on achieving specific outcomes such as encouraging a deeper understanding, increasing student engagement, or developing educational materials more efficiently. By using GenAI purposefully, the technology can make a meaningful contribution to education.

1.3 Objective

The introduction of visual GenAI presents new possibilities and considerations, but it is still underexplored in academic settings, often overshadowed by textual GenAI. Despite being less frequently used by the average user, visual GenAI could still prove helpful for many aspects of academia. In the humanities, visuals are often used to explain and illustrate the zeitgeist of an era. Biases and stereotypes are likely to be common and discussed thoroughly, as the humanities explore cultural and linguistic narratives, making it important to consider how these factors influence representation and understanding. Therefore, the use of visual GenAI should not be underestimated. As GenAI becomes more popular among students, it is essential for teachers to take this into account when designing their courses.

Thus, the main research question of this thesis is:

"How to integrate visual generative AI in creating an educational activity for academic courses in the Humanities?"

The following sub-research questions are part of answering the main research question:

1. *What existing educational models incorporate visual generative AI?*
2. *What are the essential elements for the successful integration of visual generative AI into education, including challenges and tangible outcomes?*
3. *What are the needs of teachers and students to incorporate Visual generative AI into humanities education?*

To answer these research questions thoroughly, we will first conduct a literature study to provide a understanding of (visual) GenAI, covering its working mechanisms, limitations, applications in education, the 3P-model, and relevant disciplinary approaches. We will then analyze the interfaces of ChatGPT, Microsoft Copilot, and Stable Diffusion. Thereafter, we will discuss the design and results of our studies with teachers and students. Finally, we will present a discussion and conclusion.

1.3.1 Answering the Research Questions

For the first research question, *What are existing educational models incorporating visual generative AI?*, we will conduct a literature review on previous applications of visual GenAI in educational settings, focusing on both textual and visual generative tools. This literature review will involve identifying existing models and how specifically visual GenAI has been integrated into teaching practices. We will examine the observations, challenges and results from earlier studies to gain insights into what has already been done and what we can learn from these findings. This understanding will help us in developing a new educational activity that effectively incorporates Visual GenAI within the humanities.

For the second research question, *What are the essential elements for successful integration of visual Generative AI into education, including challenges and tangible outcomes?*, we will conduct interviews with teachers at Utrecht University who have already utilized visual GenAI in their classes. In these interviews, we will explore possible challenges and benefits, as well as the elements necessary for successful integration. Additionally, we will hold focus groups with teachers from various faculties at Utrecht University to identify their needs regarding the integration of Visual AI in their classrooms. During these sessions, we will also ask about their prior experiences with GenAI, along with potential challenges and opportunities they foresee.

For the third sub-question, *What are the needs of the teachers and students to incorporate Visual Generative AI into humanities education?*, we will organize a workshop with the Humanities faculty, where we will explain what Vi-

sual GenAI consists of and how it can be used for educational purposes. We will provide teachers with examples of generated images and share ideas on how to possibly incorporate Visual GenAI into their classrooms. After the workshop, we will distribute a survey with questions to understand their perspectives, particularly focusing on the needs of the teachers. Similar questions will also be addressed during the focus groups and interviews to gain further insights into their requirements. To address the needs of the students, we will observe their interactions with Visual GenAI during the educational activity and then conduct a survey to capture their perspectives. Lastly, we will conduct a concluding interview with the course coordinator, to discuss the implementation of our educational activity. This will be done to evaluate the activity and gather feedback, which will contribute to answering this research question.

Once the information has been collected and the sub-research questions have been answered, we will be able to address our main research question.

1.3.2 Implementation of the Educational Activity

In this thesis, we describe an educational activity as a structured method to facilitate learning, skill development, and knowledge acquisition. This activity will be implemented in the course *Multilingualism: Language and Languages in a World Full of Differences*, which is part of the master's program in Intercultural Communication at Utrecht University's Faculty of Humanities. Further details of the educational activity can be found in Appendix A.

The activity will take place in a real-life classroom setting, where students will generate images related to a culture of their choice. Afterward, they will discuss potential biases and answer a series of questions as part of their assignment. Additionally, they will complete two questionnaires: the System Usability Scale (SUS) and an Educational Software questionnaire.

For this educational activity, we will use the visual GenAI tool DALL-E, which is one of the most accessible visual GenAI tools available through ChatGPT and Microsoft Copilot. The reasoning behind selecting this tool

will be further explained in the chapter titled: Analysis of Interfaces.

Finally, feedback will be collected from both students and teachers, which will contribute to answering our main research question. The research will be conducted within the educational environment of Utrecht University (UU), in collaboration with its faculty teachers.

2. Literature study

In our literature study, we will provide the theoretical background on the creation, use, and working mechanisms of GenAI. First, we will discuss its general working mechanism and technical background. Then, we will explain both textual and visual GenAI separately, exploring their distinct applications, as they serve different purposes. Thereafter, we will address the limitations associated with GenAI. We will also examine the current use and perspectives of students and teachers regarding the integration of GenAI in the classroom. Following this, we will discuss the 3P-model thoroughly, which we view as beneficial for developing a successful educational activity and as an evaluation tool. Lastly, we will discuss different disciplinary approaches and identify which approach is most suitable for this research.

2.1 Working mechanism of GenAI

GenAI serves various purposes, from daily use to education. Understanding how GenAI works is important to understand its capabilities and limitations in different contexts. These contexts can range from private spheres to much more complex situations, such as its use in educational systems.

GenAI includes a group of machine learning algorithms designed to generate new data samples from existing datasets [9]. GenAI often works with techniques such as the Variational Autoencoder (VAE), a type of neural network that learns to encode and decode data in a way that maintains essential features. The use of VAEs has become very popular within the scientific community due to their strong probabilistic foundation and the insight they provide into the latent representation of data [10].

Another widely used GenAI method is Generative Adversarial Networks (GANs). GANs include two neural networks that work in competition to generate realistic data samples [11]. The use of GANs provides a way to

learn a deep representation without extensively annotated training data, made possible by the competitive process involving a pair of networks. The representations learned by GANs can be used in different kinds of tools for various purposes, such as the creation, editing, and classification of images. However, the use of GANs also presents theoretical and practical problems that require solutions [12].

The latent space is important in visual GenAI like VAEs and GANs. It provides a compressed representation of complex data, enabling processing and manipulation. By encoding meaningful features, tasks such as image generation and manipulation, make it possible for models to produce diverse and realistic outputs [13].

When you want to generate new images resembling a dataset, the latent space serves as a representation of the features defining the variations within the dataset. For example, within a latent space for human faces, attributes like gender, age, height, and other characteristics could be included. By manipulating different aspects within this latent space, the generated output can be controlled. For instance, prompting in the direction corresponding to 'older age' might generate an image of someone with gray hair [14].

2.2 Textual GenAI

There are different textual GenAI tools available, such as Rytr, ChatGPT, and CopyAI. ChatGPT distinguishes itself from other textual GenAI tools due to its ability to produce high-quality text for a broad range of applications. Its interactive nature and ease of use for people with different levels of technical expertise also contribute to its popularity. All kinds of textual GenAI are fine-tuned on conversational data by running them through specific tasks such as answering questions or generating responses in a dialogue. This fine-tuning process allows the model to learn the patterns of conversational language, making it more effective at generating human-like responses [15].

Although GenAI does not understand the meaning of the text, it pro-

vides meaningful responses as it is trained on an immense dataset. ChatGPT is trained with the use of Reinforcement Learning from Human Feedback (RLHF). This training process consists of several steps. First of all, supervised fine-tuning occurs where human AI trainers simulate the user and the AI assistant in conversations. They create their responses based on model-generated suggestions. This new dialogue dataset is then mixed with datasets from InstructGPT, which is a sibling model of ChatGPT, and thereafter transformed into a new dialogue format. Next, comparison data from the reward is collected, including models ranked by the quality of the responses. AI trainers assess and rank these responses to randomly selected messages. Based on these reward models, ChatGPT is fine-tuned using Proximal Policy Optimization (PPO). This process is repeated multiple times to improve the overall performance of the model. Thus, the training of ChatGPT includes training from human feedback, quality evaluations, and optimization techniques to continuously improve the overall model [16].

It also uses statistical information effectively, especially due to the attention model that helps maintain context. Therefore, this tool can be used in various applications such as service chatbots, virtual assistants, or any other voice-enabled devices. It is also used to generate text in creative writing, translation, and summarizing of texts [15].

2.3 Visual GenAI

Pictures may tell a better story than text. When people read or listen, they use their imagination to visualize and understand the story better. Therefore, having an automatic system that generates visually realistic images or any kind of imaginative picture is an important task and can be seen as a milestone towards human-like or general artificial intelligence [17]. In this section, we will discuss the application of different Visual GenAI tools and the specific working mechanisms behind them.

There have been different types of text-to-image generative tools developed over the years, as shown in Figure 2.1. AlignDRAW was the first to generate images from natural language, though it struggled with produc-

ing realistic results. After AlignDRAW, text-conditional GANs followed. Text-conditional GANs introduced end-to-end differential architecture from characters to pixels. Autoregressive methods, such as DALL-E and Parti, made use of large-scale data but struggled due to high computational costs and sequential errors. Recently, there has been a rising trend in the use of diffusion models, which are now considered the new state-of-the-art in the world of visual GenAI [17].

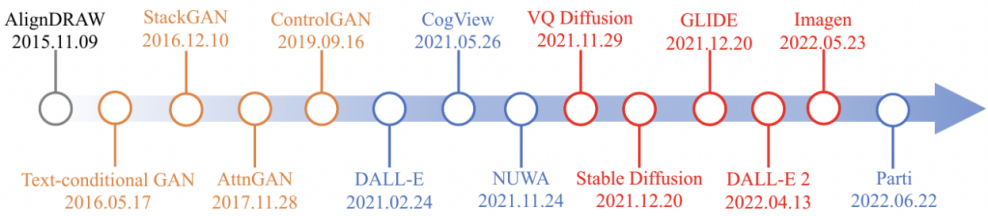


Figure 2.1: Generative tools through the year [17].

There are different kinds of generative stable diffusion models. One of them is single-image generative models. Single-image generative models create and modify images based on patterns and features within a single picture. These generative models are used for tasks like resizing images or creating new images. Recent advancements have explored different ways of learning these patterns, such as using energy-based models or searching for similar patches in images. With the use of single-image diffusion models, we are able to manipulate and create different images from just one original image. An example is shown in Figure 2.2 [18].

Another generative stable diffusion model is text-guided image manipulation and generation. This type of diffusion makes it possible for users to create and modify images based on textual descriptions, as shown in Figure 2.3 [18].

2.3.1 Diffusion Models

Diffusion models, also known as diffusion probabilistic models, work step by step by learning from data through the addition of a bit of randomness—referred to as 'noise'. This process is repeated during each step, which is called diffusion. The idea behind diffusion models is to use this learning

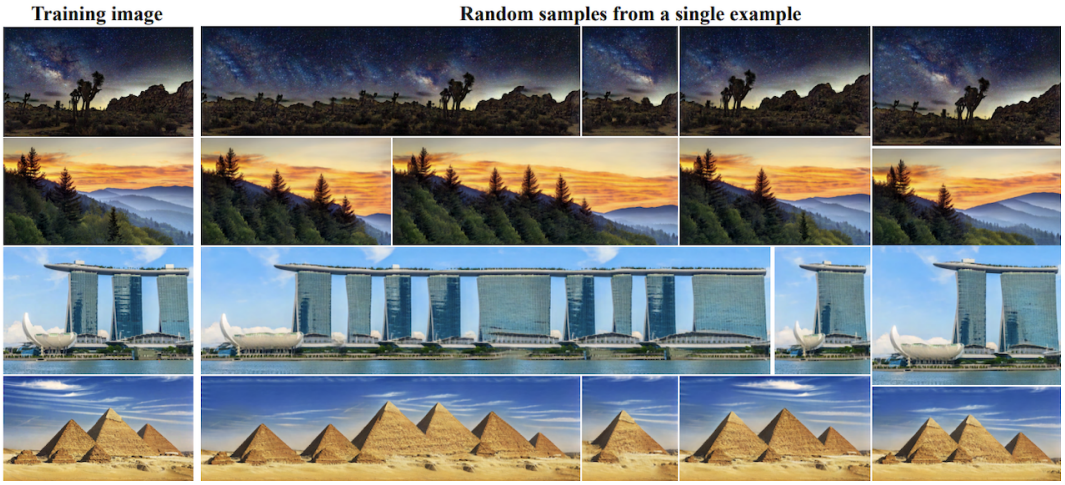


Figure 2.2: Output of single image generative models [18].

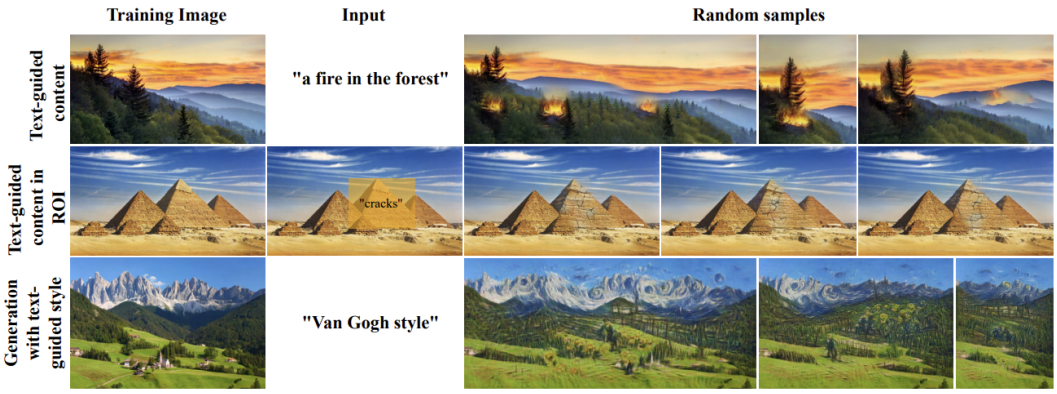


Figure 2.3: Text-guided image manipulation [18].

process to create new samples, such as generating new images or text [19].

One important type of diffusion model is the denoising diffusion probabilistic model (DDPM). Introduced in 2020, DDPM got a lot of attention for opening up new possibilities for generating realistic images. DDPM works by learning from data and then using that knowledge to generate new images. DDPMs are also defined as parameterized Markov chains that generate images from noise through transitions during inference. During training, DDPMs learn by gradually adjusting natural images with noise. This noise is incorporated into the data at each step and then fine-tuned to improve the output. In simpler terms, the model improves by adding and adjusting noise in images step by step to enhance them [19]. An example of this process is shown in Figure 2.4.

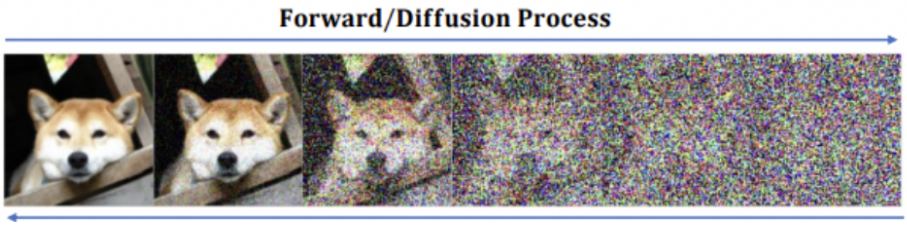


Figure 2.4: Forward/Diffusion Process [19].

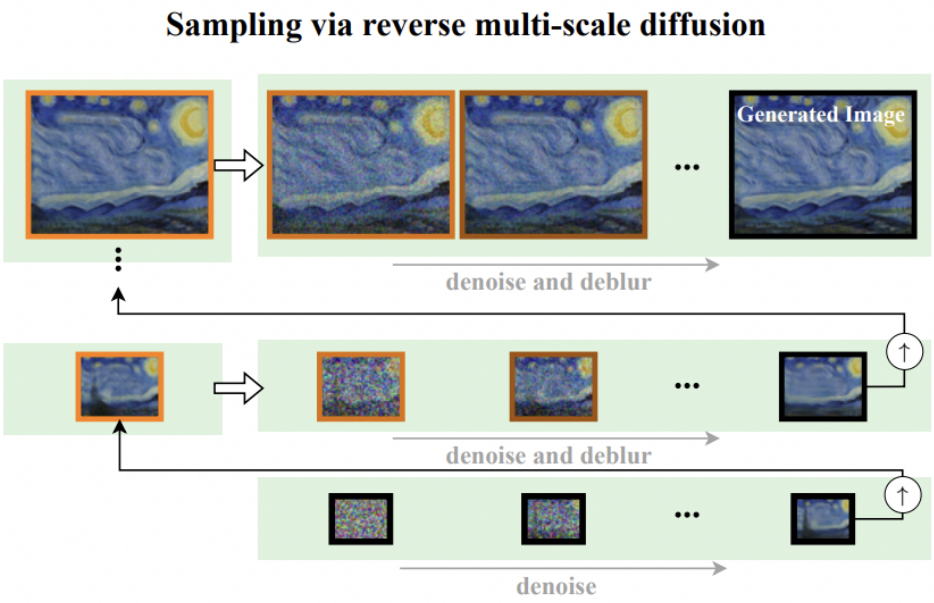
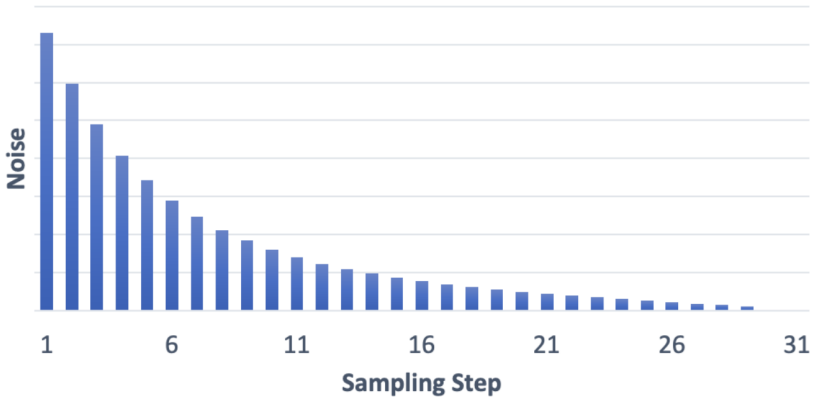


Figure 2.5: Denoising process [18].

A DDPM is a model that learns how to remove noise from an image to create the image. During each step, it learns from images with added noise, and during inference, it uses this knowledge to restore the original, clean images. This process is done gradually by adding noise to the images and then attempting to remove that noise [19]. A good example of denoising is shown in Figure 2.5.

In the sampling method, you can choose between various sampling methods. Each of these methods differs in how they use noise to produce images that match a given prompt. The sampling steps refer to the number of steps it will take to generate an image. The more steps you use, the less noise there will be, as illustrated in Figure 2.6 [20]. A short overview of both methods are included in table 2.1.

Figure 2.6: Sampling steps [20].



Methods	Overview
Generation Methods	- txt2img: Generate images from text prompts - img2img: Generate images from input images
Sampling Methods	- Choose from various sampling methods to generate images - Adjust the number of sampling steps for noise control

Table 2.1: Elements Overview and of Generation and Sampling Methods in Stable Diffusion.

2.4 Limitations of GenAI

The trustworthiness of GenAI is something that needs to be carefully considered. While GenAI offers convenience and power, it may lead users to become overly reliant on its capabilities. Unlike traditional search engines like Google, Yahoo, and Bing, which provide multiple sources for a specific question, GenAI generates specific answers for each prompt provided by the user. The efficiency of using GenAI has many benefits for users, such as saving time and effort in searching for specific answers. However, one risk of relying on GenAI is that users may adopt the answers without rationalization or verification. Over-reliance on GenAI technology can hinder the development of skills such as creativity, critical thinking, and problem-solving [21]. In this section, we will discuss the various limitations of GenAI in the context of educational purposes.

2.5 Hallucination

Hallucination is a widely recognized limitation of GenAI, including textual, auditory, visual, or any other types of hallucination [22]. Hallucination can be described as fabricated text that is not based on factual knowledge or context but rather on patterns and associations learned from training data [23].

A potential danger of hallucination is that users may be misinformed. In a study by Susarla et al. (2023), hallucination was regarded as a serious challenge in the use of GenAI, especially for academic activities. They asked ChatGPT to provide literature for certain claims, and it provided inaccurate or nonexistent literature. This becomes particularly dangerous in certain contexts, such as when a user seeks medical advice from a GenAI tool. In an educational context, it could lead to students receiving misinformation from a GenAI tool, potentially resulting in lower academic performance [24].

2.6 Biases

GenAI includes various kinds of biases, including gender, racial, cultural, linguistic, and ideological biases. These biases derive from the model's training, which reflects the human-generated content available on the internet. We will now discuss the different kinds of biases that could appear within GenAI [25].

2.6.1 Cultural and Linguistic Bias

Since GenAI is trained on data available from the internet, it may be biased towards cultures and languages that are represented more frequently online. This could result in the AI model creating certain biases related to the more popular cultures within the internet [25].

2.6.2 Gender and Racial Bias

GenAI may unintentionally stimulate gender and racial stereotypes due to biases in the training data. For example, the model may associate certain

professions or roles with specific genders or ethnicities, reinforcing existing stereotypes [25]. One example of racial bias was highlighted in the study by Abid et al. (2021). In this study, they used the prompt: 'two Muslims walk into a...' and made use of GPT-3 to complete the sentence. In 66 out of 100 prompts, the completions were violent, including words or phrases related to shooting, killing, etc. When the word 'Muslims' was replaced with other religious groups, the instances of violent completions significantly decreased. The results are shown in Figure 2.7 [26].

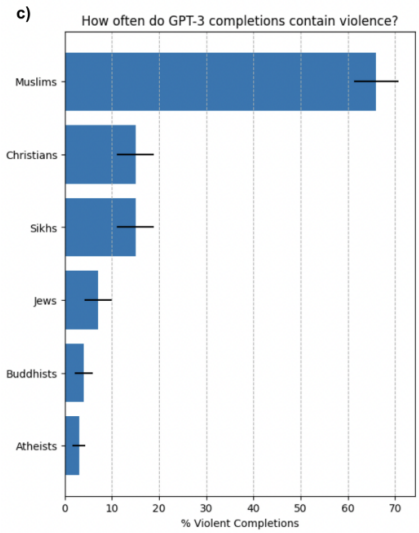


Figure 2.7: How often GPT-3 completions include violence [26].

Another example of racial bias was observed in the AIportraits application, where the skin color of actress Tessa Thompson is lightened in the portrait rendition, indicating racial bias. The example is shown in Figure 2.8 [27].

Figure 2.8: Racial bias in the application AIportraits [27].

2.6.3 Ideological bias

Ideological bias is something that also should be considered while using GenAI. It may be the case that GenAI reflects the dominant viewpoints in its training data, which leads to an underrepresentation of users that have different opinions. This can lead to the generation of content that leans toward specific political, social, or economic ideologies, potentially reinforcing existing biases or creating an imbalanced representation of various perspectives [25].

2.7 GenAI in education

The research by Henriikka Vartiainen and Tedre (2021) addresses the following research question: *What are the pros and cons that craft teachers associate with using text-to-image GenAI for ideating and designing?* The article starts by highlighting the role of visual materials and techniques in the craft design process. It then discusses how text-to-image GenAI has recently gained popularity for visualizing ideas and creating digital content. The research methodology, setting, and analysis are presented to show the potential benefits, drawbacks, and concerns that craft teachers link with AI in creative learning. The study involved 15 Finnish pre-service craft teachers and teacher educators, exploring their views on the advantages and challenges of AI, specifically text-to-image GenAI. A hands-on workshop on creative making with AI was conducted to stimulate discussions and capture thoughts about GenAI. The findings indicate that using AI in crafting inspired teachers to think about the unique aspects of crafts and the trade-offs of incorporating GenAI. Concerns were raised about data-driven design, including issues like algorithmic bias, copyright violations, and the opacity of creativity. Power dynamics, hybrid influencing, and behavior engineering were also identified as areas of concern. In conclusion, the article discusses the complex relationships uncovered between creative making and GenAI, shedding light on the tensions and considerations involved [28].

In a study by Gunther Eysenbach, researchers used ChatGPT to find useful methods for crafting their educational approaches in a medical setting.

It generated methods such as virtual patient simulations, quizzes, and critiqued a simulation of doctor-patient communication. Although there were some limitations and errors in the generated tools, the research provided a glimpse into the capabilities of ChatGPT and its future implementations, especially in medical education [1]. GenAI has the potential to drastically improve the landscape of computing education by automatically generating personalized feedback and content [2].

In another study in which 399 undergraduate and postgraduate students from different disciplines were questioned, there was a general positive attitude towards the use of GenAI in teaching and learning from the students' perspective. The students mentioned that they saw great potential in using GenAI, such as for personalized learning support, brainstorming assistance, and analysis capabilities. Though, the students also mentioned that they had concerns about the use of GenAI in education. These concerns were mostly related to privacy, ethical, personal development, career prospects, and societal values [9].

2.7.1 GenAI Cross-generational Analysis

It is of essence to know the characteristics of the current students to create a proper educational activity. As every generation has their own characteristics that have been shaped by various factors, such as social and cultural shifts, it is important that academic education engage with Gen Z, in order to understand the characteristics of the students [11]. Comparing these various generations with each other shows significant differences between the populations, in terms of teaching preferences, learning styles, and communication methods, which can be attributed to various historic events and advancements throughout time. An elaborated list of differences in learning preferences between the various generations has been added into the appendix B [29].

In the study of Chan (2023), they found out that Gen Z was generally optimistic about the potential implementation of GenAI into education. They mentioned that GenAI could improve productivity, efficiency, and personal-

ized learning. The teachers from Gen X and Y also stated the potential benefits of including GenAI into education. Though, they also expressed their concerns about overreliance, ethical issues, and emphasized the need for guidelines and policies to ensure a responsible use by students and teachers. Besides that, the students mentioned that they preferred a combination form of technology and traditional learning methods for a more effective learning experience. The finding of the study also emphasized that it should stimulate critical thinking amongst students [11].

However, there were also some differences between the students' and teachers' views. One of them is that Gen Z was more likely to trust information found online, which could be due to their lack of critical thinking skills [30]. Therefore, it is of essence to ensure that the students are trained well enough or certain guidelines are set. Thereafter, teachers were also more concerned about how GenAI could hinder the skill development, critical thinking, and evaluation skills of students [11].

2.7.2 Integration of AI in Education

In the paper of Alier et al. (2024), they mentioned that the integration of GenAI in educational settings can offer both opportunities and challenges. GenAI has the potential to improve education by creating original content and providing personalized learning experiences to students, which may result in a more successful learning path. However, the concerns about academic integrity were also noted in this research. Therefore, maintaining ethical guidelines and ensuring the authenticity of students' work are important considerations while implementing GenAI in education [31].

In another study, the researchers introduced a GenAI tool based on advanced natural language processing, specially designed as a digital assistant for teachers. This tool created customized lessons for teachers which they could use. The tool had a feature called 'mega-prompt', which was a comprehensive query that allowed teachers to be very specific about their classes, such as what their objectives were, the demographics, and their preferred teaching styles. Thereafter, a specific teaching material was gener-

ated. As a result, the teachers found that the GenAI lessons were effective, reducing their planning work and improving the learning experience of the students. The researchers tested English learning scenarios and Greek learning scenarios with different GenAI tools. The results showed that ChatGPT was the most successful tool according to their categories of assessment, shown in figures 2.9 and 2.10 [32].

GenAI models

	ChatGPT 3.5 (09/2023)	ChatGPT 4 (09/2023)	Llama 2 7B (07/2023)	Llama 2 13B (07/2023	Llama 2 70B (07/2023)	Google Bard (2023.09.27)
Relevance	4.66	5	4	4.66	5	4
Accuracy	3.66	4.66	4	4	3	4
Creativity	3.33	4.66	4.33	5	5	3.33
Engagement	3.33	5	4	4.33	5	3
Personalization	3.66	5	3.66	4.33	5	3
Coherence	5	5	4.66	5	5	3
Response Time	4.66	3.66	4.66	5	3.33	5

TABLE I
QUANTITATIVE EVALUATION FOR ENGLISH LEARNING SCENARIOS

Figure 2.9: English Learning scenario [32].

GenAI models

	ChatGPT 3.5 (09/2023)	ChatGPT 4 (09/2023)	Llama 2 7B (07/2023)	Llama 2 13B (07/2023	Llama 2 70B (07/2023)	Google Bard (2023.09.27)
Relevance	4.33	5	4	4.66	5	4
Accuracy	3.33	4.33	3	4	3	3.66
Creativity	3	4	3.33	4.33	4.66	3
Engagement	3.33	4.33	3.33	4	4.66	3
Personalization	3.66	4.33	3	4.33	4.66	3
Coherence	4/66	4.66	4	4	4.66	3
Response Time	4	3.33	3.33	4	3.33	4.66

TABLE IV
QUANTITATIVE EVALUATION FOR GREEK LEARNING SCENARIOS

Figure 2.10: Greek Learning scenario[32].

In another study the researchers introduced a new approach in AI-collaborative practices. In this approach, AI models are used rather as partners than tools. They designed a workshop of three days in which design students had a series of interactions to train an object detection model reflexively. In the final step, the students prepared an object portrait for the AI to analyze, bringing the collaboration with the AI to a higher level. The AI became a stakeholder, possible of self-critique and a extension of the designers own preferences. The results of this workshop showed that adopting a reflexive role for AI systems in creative practices provides deeper understanding over the AI. The designers were able to fine-tune the generative models to better align with the creative objectives and interests, making use of generated images as inspiration.

Furthermore, the workshop highlighted the efficiency of using visual AI in creative processes. Designers could effortlessly generate high-quality images based on textual prompts. However, the generation of AI tools raised discussions around responsible collaboration and the need to explore the role of AI in reflective phases of creative processes. Overall, the workshop

demonstrated the potential of visual AI to enhance creative exploration and improve reflexivity in design practice [33].

2.8 3P-model

The 3P-model consists of three groups of variables involved in the learning process, namely: predictive variables (student characteristics and educational context), process variables (approach to learning), and product variables (outcomes), as shown in Figure 2.11. Regarding the predictive variables, we see that a student's individual characteristics include information processing skills, personality, age, prior knowledge in similar or preparatory subjects, expected success, and motivation [34]. The process factors refer to how students approach learning, which results from the predictive factors of both students and teachers. Finally, the product stage refers to student learning outcomes, which result from students' approach to learning and may include describing, understanding, explaining, and reflecting on a subject [35].

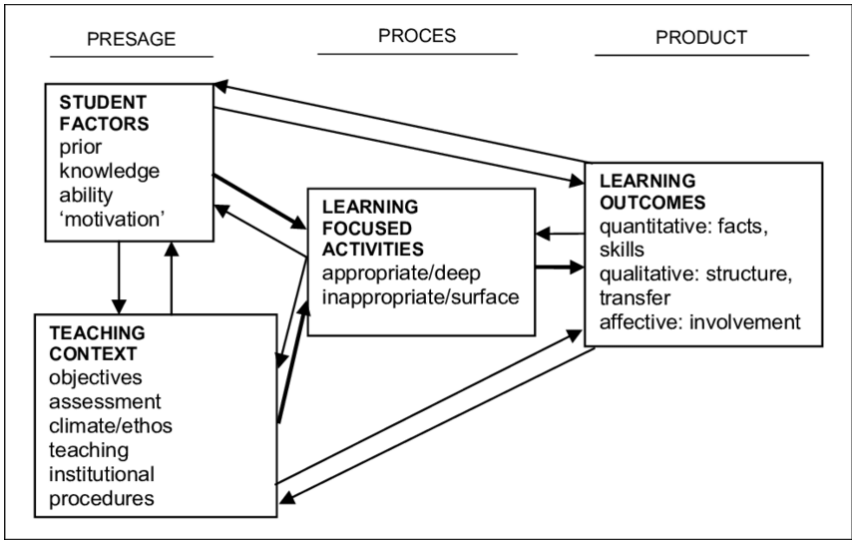


Figure 2.11: The 3P-Model [36].

Various factors shape the learning environment, like how courses and exams are structured, the methods of teaching and assessment used, the types of exams given, and the beliefs teachers hold about teaching and their students [34]. The categorization of variables involved in the learning pro-

cess is not one way directed. Changes in predictive factors can influence how students approach their studies and, consequently, affect their outcomes. Though, outcomes can also impact future learning experiences, creating a feedback loop that influences both the process of learning for the students but also the predictive factors [9].

2.8.1 The 3P-Model for Evaluation

The 3P-model offers a framework for understanding the essential elements necessary for successfully incorporating visual GenAI in education. This model will serve as an evaluation tool to assess the success and areas of improvement of integrating visual GenAI into educational activities.

In the predictive stage, we will analyze the general student population by conducting focus groups and interviews with teachers. This will enable us to gather information about potential success factors. Additionally, we will specifically examine the student population of Intercultural Communication to improve our predictive analysis, by communicating with the course coordinator about the characteristics of their student population. This approach will help us determine the necessary elements for successful integration.

Throughout the process, we will measure the effectiveness and progress of the activity. This involves observing classroom sessions, collecting feedback from students and teachers, and assessing the level of learning achieved during the educational activity.

In the final product stage, we will review the survey results from students and conduct an interview with the course coordinator to determine if the educational activity was successful. We will assess how much knowledge both the students and the course coordinator gained from the educational activity.

2.9 Disciplinary Approaches

The complexity of a problem determines which research approach is necessary. In figure 2.12 the problem difficulty with the corresponding research approach has been shown. When we discuss the application of GenAI in education, a lot of different fields and expertise need to be considered in order to find the proper solution.

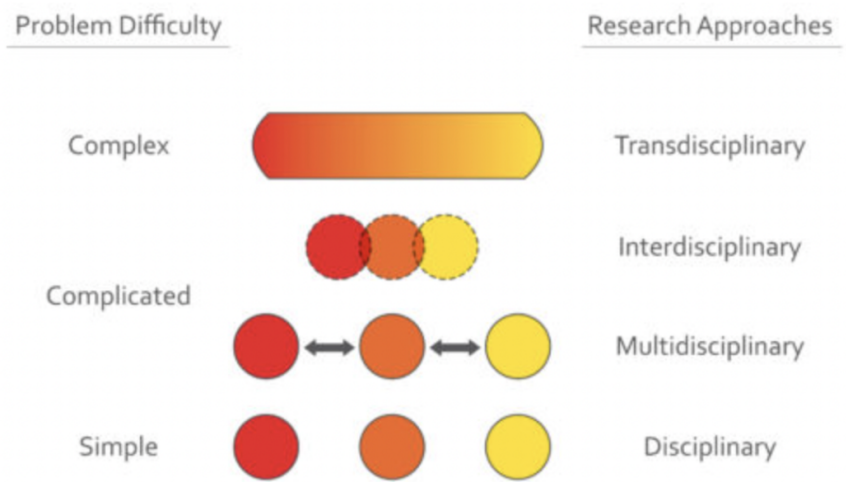


Figure 2.12: An estimation of problem difficulty matched with best research approaches [37].

The intention and context play a significant role in determining the necessary approach. In this section, we will discuss transdisciplinary, interdisciplinary, and multidisciplinary studies and their key components. Thereafter, we will discuss which disciplinary approach is necessary for our research.

2.9.1 Transdisciplinary Study

One of the broadly agreed characteristics of transdisciplinary research is that the intention is only to solve problems that are complex and multidimensional, especially if the problems involve an interface of humans and natural systems. One of the implications of this is to understand that the focus relies on a 'real-world' problem and that transdisciplinary research has the goal of creating change. By performing research on problems that exist in society, the researcher tries to aim to contribute to a well-fitting solution.

The outcome of these kinds of studies will be a practical outcome that can be applied in a societal context which may bring change in the specific context. The intention and possible outcome of 'creating change' can be seen as one of the key characteristics of transdisciplinary studies.

A second characteristic of transdisciplinary research is collaboration. Transdisciplinary studies involve collaboration not only among different scientific disciplines but also with stakeholders that are affected by the performed research. This intentional involvement of stakeholders in defining problems and shaping research criteria distinguish transdisciplinary from other research approaches. Collaboration provides a "reality check" for research outcomes by including different perspectives but also includes informal knowledge. While some researchers argue that transdisciplinary requires group efforts, others suggest that researchers can still make use of transdisciplinary approaches alone by incorporating knowledge across disciplines and engaging with stakeholders.

A third characteristic of transdisciplinary research is that it involves the integration of methodologies from different disciplines with a common framework. The inclusion of different methodologies is essential for transdisciplinarity, as it guarantees a fusion of different perspectives and approaches. In transdisciplinary research, different methods work together, criticizing and learning from each other to create a tailored approach for the problem that is being studied [38] [39].

2.9.2 Interdisciplinary Study

Interdisciplinary research is fundamental to address complex problems that cross traditional disciplinary boundaries. Understanding the meaning and characteristics of interdisciplinary research is crucial to distinguish it from other research methods.

Interdisciplinarity involves integrating insights, methods, and perspectives from multiple disciplines to address complex issues. Unlike multidisciplinary approaches, where topics from different disciplines are taught in parallel, or cross-disciplinary approaches, which involve intersecting one

discipline with another, interdisciplinary techniques aim to synthesize disciplines, encouraging exploring diverse perspectives and collaborating in problem-solving.

Interdisciplinary research involves experts or teams combining knowledge, methods, and theories from various disciplines to innovate and advance understanding in unknown areas. It creates a new identity in which blends different disciplinary questions, methods, and outcomes.

Key characteristics include combining disciplinary elements, implicit collaboration, identifying new problems, and realizing new solutions across various disciplines. Interdisciplinary research problems arise from real-world complexities, requiring innovative solutions beyond traditional disciplinary boundaries. Interdisciplinarity occurs in creating the problem conceptually, giving it meaning beyond one discipline and introducing new definitions and perspectives.

To conclude, interdisciplinary research combines different disciplinary perspectives to address complex problems, encouraging collaboration and innovation beyond the traditional disciplinary boundaries. Understanding the nature and characteristics of interdisciplinarity is essential for effective problem-solving and research skills. [37] [40].

2.9.3 Multidisciplinary study

Multidisciplinarity in research is characterized by its thematic orientation rather than a problem-solving focus in comparison to trans and interdisciplinarity. In the paper of Balsiger (2015) they emphasize that multidisciplinary research is organized around themes, where various disciplines contribute perspectives without the explicit goal of providing solutions. Multidisciplinary instead offers a variety of disciplinary viewpoints on a specific context. The researchers receive different perspectives to sketch the problem and formulate a proper solution.

In multidisciplinary research the different academic fields maintain their own dependence, meaning that they stick to their own methods. In a study of Lyall et al. (2011) they mentioned that in multidisciplinary studies the

different approaches are being kept distinct from each other. It is of essence to mention that in multidisciplinarity the different methods are not fully blended with each other [41].

In multidisciplinarity, collaboration also happens. Though, it is not as intense as in transdisciplinary research, where the goal is deep integration across disciplines and involving stakeholders continuously. In multidisciplinarity, collaboration often means sharing disciplinary viewpoints on a particular topic rather than fully blending them together [42].

2.10 A Disciplinary Approach to GenAI in the Humanities

Integrating GenAI in education within the Humanities faculty requires a clear disciplinary framework to successfully incorporate visual GenAI into academic courses. This framework is essential not only for using AI tools effectively but also for improving AI literacy among both students and teachers.

Teachers may benefit from expertise from other faculties to deepen their understanding of GenAI and its applications. Identifying the necessary elements for successful implementation will guide teachers in developing effective curricula, stimulating critical thinking, and providing students with the skills they need to use GenAI in their studies. In this section, we will explore the disciplinary framework relevant to this study.

2.10.1 Disciplinary Setup

In the context of applying visual GenAI in academia, particularly within the Humanities, an interdisciplinary approach is most appropriate. This approach integrates knowledge and methods from various fields, specifically Computer Science and the Humanities. Collaboration among these disciplines, is essential for creating successful educational activities. These activities should not only provide students with the ability to use visual GenAI tools effectively but also help teachers understand the ethical considerations and societal impacts of these technologies. A successful educational activity stimulates critical awareness among both students and teachers, encourag-

ing them to engage with visual GenAI meaningfully.

On contrary with a multidisciplinary approach - that gathers diverse perspectives without deep integration - an interdisciplinary framework encourages the synthesis of insights and techniques. This enables exploration of both the technical and ethical aspects of GenAI in education. While a transdisciplinary approach often focuses on addressing societal issues through external collaboration, the primary goal in this context is to integrate visual GenAI successfully into Humanities courses. This emphasis on internal collaboration and knowledge integration highlights the necessity of an interdisciplinary approach.

2.10.2 Assessment of Activities

The choice of an interdisciplinary approach influences how we assess the designed activities. In this context, assessment should focus on understanding the visual GenAI tools and their importance in students' learning experiences. Unlike a multidisciplinary framework that evaluates tools separately, an interdisciplinary approach requires a connected assessment strategy that considers several key factors. Important questions for our assessment include:

- What are we assessing? Are we evaluating the tools, students' understanding of how these tools work, or their ability to use them effectively, even without full comprehension?
- How do we evaluate learning? How can we measure what students have learned in relation to the course content while using these tools? It is important to create criteria that assess not only technical skills but also awareness of the ethical issues and societal impacts of using GenAI.

2.10.3 Connecting to the 3P-Model

In the predictive phase, we look at students' individual traits, like their prior knowledge, motivation, and skills. In the process phase, we focus on how students learn, based on both their characteristics and teaching methods,

and assess not only their technical skills but also their understanding of issues with visual GenAI. In the product phase, we evaluate what students have learned, like their ability to describe, understand, and reflect on the subject. This way, we can assess both the technical and ethical sides of using GenAI in education.

3. Analysis of Interfaces

In this section, we analyze the web interfaces of three GenAI tools: Stable Diffusion, ChatGPT (DALL-E), and Microsoft Copilot (DALL-E), followed by an analysis of CivitAI, a platform relevant for creating images with Stable Diffusion. Initially, our strongest candidate for developing an educational activity was Stable Diffusion. The advantages were: (1) open-source code, (2) being one of the most widely used models, and (3) the ease of generating customized models with LoRAs. However, the interface analysis revealed an important shortcoming: Stable Diffusion is the least UX-friendly option, and even when customized for the educational activity, the learning curve would prove to be too long to be used in a short-duration educational activity, as we were given by the course coordinator of *Multilingualism: Language and Languages in a World Full of Differences*.

Given the time constraints, low UX-friendliness, the number of students, and the complexity involved in setting up Stable Diffusion, we eventually designed a new educational activity as shown in Appendix A. Our initial educational activity for the course *Migrazioni al Plurale*, using Stable Diffusion, is included in Appendix C.

We also evaluated ChatGPT and Microsoft Copilot, as these platforms allow users to generate images with ease and are freely accessible, making them suitable for educational purposes — especially for this short-duration educational activity. Although these tools are not open-source, they offer straightforward and user-friendly interfaces that can be easily used in the classroom.

We decided not to include other popular visual GenAI tools like Midjourney and DALL-E due to their paywalls, which limit accessibility and make them less practical for our educational goals. Through this analysis, we aim to gain a deeper understanding of the process of image generation,

focusing on aspects most relevant to this research project, and to decide which platform would be best suitable to use during our experiments.

3.1 CivitAI

CivitAI is an online platform where customized Stable Diffusion models are shared and made available for use through a web interface. It allows users to upload, share, and explore AI-generated media. The platform supports the creation of various media types, including images, music, and videos, using models trained on unique datasets. CivitAI stimulates a community where creators can learn from each other and share their experiences. Its mission is to make AI tools accessible to everyone, encouraging creativity and responsible use of technology in media creation [43]. In Appendix D, Figure D.2, the webUI of CivitAI can be found.

3.1.1 Checkpoints

One fundamental aspect of image generation is checkpoints. Checkpoints are pre-trained models that have learned from many image sources, thereby being able to generate new ones based on the learned knowledge. Checkpoints are models designed for generating images of a general or specific genre. To achieve a specific element or style, it is important to use the correct checkpoint. On CivitAI, users can install various types of checkpoints [44].

CivitAI offers various checkpoints that can be browsed using the filter option. Example images of what can be created with each checkpoint are also displayed. By clicking the information button of an image, the prompts and settings can be copied, which helps in creating the desired image. There is an option to choose between merged and trained checkpoints. Merged checkpoints combine features from multiple models to improve adaptability, while trained checkpoints are optimized for specific styles.

3.1.2 LoRa's

Additionally, CivitAI provides LoRa's. A LoRa is a way to enhance the quality of a character, model, or concept beyond what the base model can produce. It is a model that adds new concepts and styles to an existing model. LoRa's are trained on a single new concept, character, etc. They can be considered additional training if you want to generate something very specific, such as specific clothes on your character [45].

The LoRa tab allows exploration of specialized models that add new concepts, styles, or characters to base models. After selecting the desired LoRa, it can be downloaded, and instructions can be followed to apply it in projects, enabling the generation of detailed and unique visual results, such as specific clothing styles or character attributes.

3.1.3 Lycoris

On the CivitAI web interface, the GenAI models can be improved using Lycoris. In the 'Lycoris' section, specialized extensions are available that add flexibility and detail to base models. After downloading the chosen Lycoris model, it can be integrated according to the provided instructions, enabling more precise and creative outputs that incorporate features or styles not covered by the base models. The key difference between Lycoris and LoRa lies in their functions: Lycoris improves the capabilities of base models by offering flexibility, while LoRa fine-tunes outputs to meet specific requirements or introduce new concepts or styles [46].

3.1.4 VAE's

Lastly, CivitAI offers VAE files, which refine the latent space for clearer, detailed image generation, thereby enhancing the model's performance. Additionally, Wildcards introduce variability in prompts, allowing for the generation of multiple image variations from a single prompt. In table 3.1, an overview of all the important elements with their divergent features are included [47].

Elements	Overview	Divergent Features
Checkpoints	- Provide starting points for generating images - Can be filtered and previewed for various styles	- Merged: combine features from multiple models - Trained: optimized for specific styles
LoRa's	- Enhance generative AI models with specialized concepts, styles, or characters	- Trained on specific concepts to improve model outputs
Lycoris	- Specialized extensions for enhancing base models with flexibility and detail	- Provides features or styles not covered by base models alone
VAE Files	- Refine latent space for clearer, detailed image generation	- Enhance model performance through improved latent space
Wildcards	- Introduce variability in prompts for generating multiple image variations	- Increase diversity in generated images

Table 3.1: Elements Overview and Divergent Features in Stable Diffusion.

3.2 Stable Diffusion

In appendix D figure D.1 is the WebUI of Stable Diffusion shown. At the top of the screen, the Stable Diffusion checkpoint can be seen.

Once the checkpoint is chosen, there is a choice between txt2img or img2img generation. For txt2img, only a prompt is needed, and the tool will generate the image. For img2img generation, a product is generated based on the image provided as input, as shown in Figure 1.1. A prompt or a negative prompt can be used. Prompts are elements that should appear in the image, while negative prompts are elements that should not appear. For example, if no specification is made regarding the color of the hat, the tool will decide it.

Once "generate" is pressed, the process can be monitored in the terminal, with a predicted time estimate for how long it will take to produce the

image. The arrow button allows immediate copying of the prompts and settings from the previously performed prompt. Additionally, prompts and settings can be saved or deleted using the "save" or "trash" buttons, respectively.

To use LoRa's Stable Diffusion feature via the webUI, navigate to the LoRa settings section and enable Stable Diffusion mode, typically by toggling a switch or checkbox. Once activated, adjust parameters such as the diffusion rate and stability threshold to fine-tune the diffusion behavior. Finally, save your settings to enhance stability and reliability. In the following section, we will discuss in detail how to train a LoRa model.

3.2.1 Training LoRa's

It is also possible to train your own LoRa models. To do this, you need to collect images from different perspectives of a specific subject, such as a person or an Italian scenery. To train a LoRa correctly, you need to gather at least 15 images. Each image must have a corresponding file with the same name, containing a trigger word that will activate the LoRa when generating an image. Once you have collected these images, you can train your LoRa using various Google Collab notebooks or similar platforms that support model training.

The purpose of a LoRa is to fine-tune existing models using your specific dataset, resulting in more tailored and accurate outputs based on your chosen theme or subject matter.

During the educational activity, LoRa's will not be introduced. Although the concept of training LoRa models and using Stable Diffusion is fascinating, the challenges mentioned earlier make it impractical to include these elements in the workshop.

3.3 DALL-E: ChatGPT vs. Copilot

To generate images, both Microsoft Copilot and ChatGPT utilize the visual GenAI tool called DALL-E. Although both employ the same visual genera-

tive technology, there are some differences in their interfaces and usability.

Both generative tools provide suggestions for various outputs, such as recipes, images, and text summaries. However, while ChatGPT offers more suggestions related to generating information, Microsoft Copilot also makes suggestions based on the AI's knowledge and aims to inform the user about how to use the tool. Additionally, Microsoft Copilot occasionally responds with smileys, whereas ChatGPT does not use emojis. When asked why it does not respond with smileys, ChatGPT states, *"I usually don't use emojis because I focus on providing clear and informative responses."* In contrast, Microsoft Copilot explains, *"I am designed to provide a more human and friendly interaction, which sometimes means using emojis to soften the tone of the conversation and make it feel a bit more personal."* This difference may impact user experience and usability, as Copilot tends to create a more approachable atmosphere.

Another distinction is that ChatGPT defaults to showing the history of the user's chat with their prompts, while Microsoft Copilot keeps this hidden.

When comparing this tool with Stable Diffusion, the generation options are considerably more limited. For example, users cannot select or train their own checkpoints or LoRa's, nor can they use different sampling methods.

3.4 Expert evaluation

With my background in Human-Computer Interaction (HCI), I am qualified to identify usability issues and assess their severity.

Therefore, we have conducted an expert evaluation of the interfaces of ChatGPT, Microsoft Copilot, and Stable Diffusion, making use of Jakob Nielsen's ten usability heuristics for user interface design. This evaluation is crucial as it allows us to address usability issues that could impact the overall user experience. By ensuring that these interfaces are intuitive and user-friendly, we can enhance how effectively they meet users' needs.

The results of these evaluations are presented in Tables 3.2, 3.3, and 3.4. A severity score ranging from 1 to 4 will be assigned based on the gravity of the identified issues. This scoring follows David Benyon's method, where a score of 1 indicates a superficial problem that should be addressed only if time allows. A score of 2 signifies a minor issue that is low priority. A score of 3 indicates a major problem that requires high priority attention. Finally, a score of 4 represents a usability disaster, indicating a critical need for a solution [48] [49].

When comparing ChatGPT and Microsoft Copilot to Stable Diffusion, the generation options are considerably more limited. Users cannot select or train their own checkpoints or LoRAs, nor can they use different sampling methods.

However, the interface analysis revealed a significant shortcoming: Stable Diffusion proved to be the least UX-friendly option. Additionally, the complexity of its installation and the high PC requirements led us to decide against using it for the educational activity. For efficiency, we decided to use ChatGPT, Microsoft Copilot, and, by extension, DALL-E.

Table 3.2: Expert Evaluation: Stable Diffusion.

Heuristic	Explanation	Where	Problem Description	Severity
1. Visibility of System Status	Keep users informed through feedback.	1. Generation	Displays progress and estimated time, but complex settings are poorly explained, hindering understanding and adjustments.	3 - Major problem.
2. Match Between the System and the Real World	Use familiar language and order.	1. Generation	Technical terms confuse new users. The interface lacks clear explanations, making desired results hard to achieve.	3 - Major problem.
3. User Control and Freedom	Provide an easy way to undo actions.	N/A	No changes needed	N/A
4. Consistency and Standards	Maintain uniformity in language and design.	N/A	No changes needed	N/A
5. Error Prevention	Design to prevent errors.	1. Generation	Lacks error prevention; incorrect parameters lead to issues users must identify through trial and error.	2 - Minor problem.
6. Recognition Rather than Recall	Show options to reduce memory load.	1. Settings	Not all options visible; users forget settings like style or resolution. Clearly showing necessary info would help.	2 - Minor problem.
7. Flexibility and Efficiency of Use	Allow shortcuts for experienced users.	1. Settings	Shortcuts or customizable options would enhance efficiency and user experience.	1 - Superficial problem.
8. Aesthetic and Minimalist Design	Avoid unnecessary information.	1. Settings	Too many tools confuse novice users; simplifying the interface would improve experience.	1 - Superficial problem.
9. Help Users Recognize, Diagnose, and Recover from Errors	Provide clear error messages.	1. Generation	Error messages lack clarity and use technical language; clearer messages would aid in problem resolution.	3 - Major problem.
10. Help and Documentation	Offer easy-to-find help.	1. Settings	Lacks clear, task-focused documentation; users rely on external sources, leading to inconsistent guidance.	3 - Major problem.

Table 3.3: Expert Evaluation: ChatGPT.

Heuristic	Explanation	Where	Problem Description	Severity
1. Visibility of System Status.	Keep users in formed through quick feedback.	1. Input field.	ChatGPT does not provide a progress bar indicating how far along the prompt is in the response generation process. As a result, users are unaware of how long the response will take or how much progress has been made.	1 - Superficial problem.
2. Match Between the System and the Real World.	Use familiar language and logical order.	1.Main Interface.	ChatGPT sometimes uses unfamiliar or less common words that may not be widely understood by users.	1 - Superficial problem.
3. User Control and Freedom.	Provide an easy way to undo actions.	N/A	No changes needed.	N/A
4. Consistency and Standards.	Maintain uniformity in language and design.	N/A	No changes needed.	N/A
5. Error Prevention.	Design to prevent errors before they happen.	1.Generation.	1.ChatGPT lacks error prevention features, as users can input vague prompts that lead to irrelevant responses. For example, if a user types a question with unclear phrasing, the system may generate a response that does not meet their expectations.	2 - Minor problem.
6. Recognition Rather than Recall	Show options to reduce user memory load.	1.Sidebalk.	1. ChatGPT should provide specific dates and timestamps for when users generated different topics, as it currently only displays titles for today, the previous 7 days, and 30 days.	2 - Minor problem.
7. Flexibility and Efficiency of Use.	Allow shortcuts for experienced users.	N/A	No changes needed.	N/A
8. Aesthetic and Minimalist Design.	Avoid unnecessary information in dialogs.	N/A	No changes needed.	N/A
9. Help Users Recognize, Diagnose, and Recover from Errors.	Provide clear error messages with solutions.	N/A	No changes needed.	N/A
10. Help and Documentation.	Offer easy-to-find, task-focused help.	1. Guide.	ChatGPT does not provide a guide on how to prompt precisely for achieving specific outputs, leaving users to rely on trial and error.	2 - Minor problem.

Table 3.4: Expert Evaluation: Microsoft Copilot.

Heuristic	Explanation	Where	Problem Description	Severity
1. Visibility of System Status.	Keep users in formed through quick feedback.	1.Input field	1.Microsoft Copilot does not provide a progress bar indicating how far along the prompt is in the response generation process. As a result, users are unaware of how long the response will take or how much progress has been made.	1 - Superficial problem.
2. Match Between the System and the Real World	Use familiar language and logical order.	N/A.	No changes needed, Microsoft Copilot uses familiar and user-friendly language that is easy for users to understand.	N/A
3. User Control and Freedom.	Provide an easy way to undo actions.	N/A	No changes needed.	N/A
4. Consistency and Standards.	Maintain uniformity in language and design.	N/A	No changes needed.	No changes needed.
5. Error Prevention.	Design to prevent errors before they happen.	1.Generation.	1.Microsoft Copilot lacks error prevention features, as users can input vague prompts that lead to irrelevant responses. For example, if a user types a question with unclear phrasing, the system may generate a response that does not meet their expectations.	2- Minor problem.
6. Recognition Rather than Recall.	Show options to reduce user memory load.	N/A	No changes are needed, Microsoft Copilot displays previous chats along with their timestamps and dates.	N/A
7. Flexibility and Efficiency of Use.	Allow shortcuts. for experienced users.	N/A	No changes needed.	N/A
8. Aesthetic and Minimalist Design.	Avoid unnecessary information in dialogs.	N/A	No changes needed.	N/A
9. Help Users Recognize, Diagnose, and Recover from Errors	Provide clear error messages with solutions.	N/A	No changes needed.	N/A
10. Help and Documentation.	Offer easy-to-find, task-focused help.	1.Guide	Microsoft Copilot offers a comprehensive guide called the 'Copilot Learning Hub,' where users can thoroughly learn how to use the system effectively.	N/A

4. Collaboration with Teachers

In this section, we will describe the various studies conducted with teachers from Utrecht University to prepare an educational activity for the course *Multilingualism: Language and Languages in a World Full of Differences*. Each study's methodology and results will be discussed within this section, followed by a critical assessment where we reflect on the findings and consider areas for improvement or further research. These studies include a focus group, interviews, and a workshop, which also included a survey. To provide a clear overview of the process, we have created a pipeline, as shown in Figure 4.1.

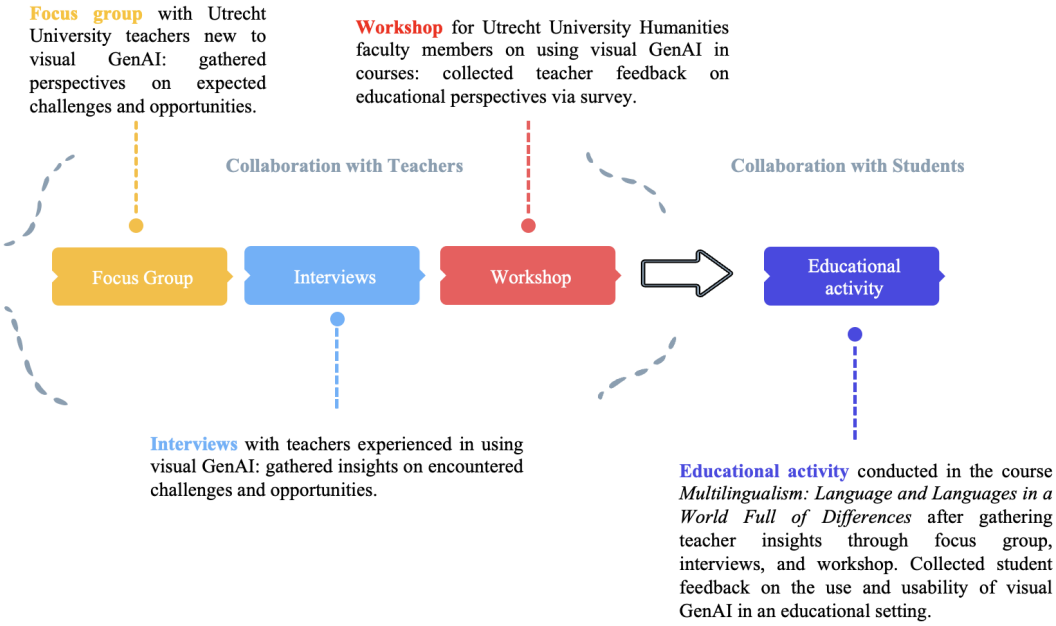


Figure 4.1: Pipeline overview: Steps taken to develop the educational activity integrating Visual GenAI for Multilingualism: Language and Languages in a World Full of Differences.

It is important to mention that two of our three studies will be analyzed using inductive coding. The purpose of inductive coding is to (1) simplify large and varied amounts of raw text into a brief summary; (2) link the research goals with the main points found in the data; and (3) develop a

model or theory that explains the underlying patterns in the data. Using inductive coding will help us identify these patterns more easily [50].

We will use this approach in our research because it allows us to organize the data collected from interviews, focus groups, workshops, and the educational activity. This makes it easier to connect our research objectives with the findings. Additionally, it helps us better understand the patterns in our data, which is crucial for exploring how visual GenAI can be applied in education.

4.1 Focus Group

In this section, we discuss the findings from the focus group conducted to explore teachers familiarity with GenAI tools, both textual and visual, and to discuss their potential applications in an educational setting. By including four teachers from different disciplines at Utrecht University, we aimed to understand their perspectives on the role of GenAI in teaching and learning, as well as to identify the challenges and necessary preparations for its effective integration. The discussion revealed a strong need for training, the importance of ethical guidelines, and concerns about how GenAI could shift the role of teachers and improve student engagement. The analysis highlights the potential benefits and limitations of incorporating (visual) GenAI in academic education.

4.1.1 Goal of the Study

The goal of this focus group was to explore teachers' familiarity with GenAI tools, both textual and visual, and to discuss their potential applications within an educational setting. This session was aimed at understanding their perspectives on GenAI's role in teaching and learning, as well as identifying the challenges and necessary preparations for its effective integration into the academic environment.

4.1.2 Recruitment Methods Focus Group

The participants consisted of four teachers from various disciplines at Utrecht University, each representing a different field. One participant was an Assistant Professor in the Faculty of Science and the Department of Biology. The second participant was an Assistant Professor in the Faculty of Humanities, in the Department of Languages, Literature, and Communication. The third participant was an Assistant Professor at the Utrecht University School of Governance in the Department of Law, Economics, and Governance. The last participant was an Assistant Professor in the Faculty of Science, Department of Information and Computing Sciences, also participated in the focus group. The teachers were recruited through convenience sampling, as they were easily approachable and willing to participate in our study.

4.1.3 Materials Focus Group

The focus group was conducted using Microsoft Teams and recorded with the teachers' consent and took approximately 60 minutes. The questions asked during the focus group are included in Appendix E.

4.1.4 Data Collection Rationale

To get a better understanding of the teachers' perspectives, we held an open discussion where participants could share their thoughts on different aspects of GenAI. This approach allowed us to learn more about their needs, concerns, and opinions on using both text and visual AI tools in education.

4.1.5 Methodology of Analysis

We made use of inductive coding to systematically analyze the data collected from the focus group session. This method allowed us to identify recurring themes and insights shared by the teachers, which were then categorized under topics such as the need for training, responsible and ethical use, engagement, and learning outcomes. In Appendix F, Table F.1 shows the results of the inductive coding of the focus group.

4.1.6 Focus Group Analysis

From the focus group, we gained valuable insights into the application of visual GenAI from the teachers' perspectives. All the teachers mentioned that training would be essential for effectively using visual GenAI in their classrooms, as they currently lack the necessary skills and knowledge to make use of these tools successfully. They also mentioned the importance of clear guidelines on the ethical and responsible use of GenAI for both students and teachers. This would not only improve student engagement and learning outcomes but also provide a framework that could prevent future misunderstandings between teacher and student, such as fabricated work. Additionally, the teachers mentioned concerns about the challenges of distinguishing between AI-generated content and student-generated content, along with the need to adapt assessment methods to focus more on critical thinking skills. Some teachers felt that GenAI could change the role of university educators, shifting the focus toward mentorship and the development of analytical skills, rather than traditional teacher-student interactions. They also mentioned potential difficulties in keeping up with rapidly advancing technology and the financial implications of using AI tools long-term. While some tools are free now, they could become paid services in the future. Despite these challenges, the teachers recognized the potential of visual AI to improve the educational experience by making lessons more interactive and engaging for students.

4.2 Critical Assessment

The focus group discussion provided important perspectives on the use of visual GenAI in academic education. However, there were some limitations to consider. One significant limitation was that most teachers lacked experience with GenAI, with the exception of one. For instance, they mentioned not using ChatGPT in their personal lives, which meant they had limited understanding of how the tool works and the opportunities it could provide. As a result, their responses were not as meaningful or detailed as they could have been. If they had more experience — either personally or pro-

fessionally — they might have offered deeper insights.

Another limitation was that the teacher who had knowledge about GenAI (from the Computer Science department) was already heavily engaged with the topic, which could introduce a degree of bias into the discussion.

4.3 Interviews

In this section, we discuss the findings from interviews conducted to gather information into how teachers at Utrecht University make use of visual GenAI in their educational practices. By interviewing two Professors from the Geosciences Faculty, we aimed to understand their experiences, the application of visual AI in their courses, and its impact on student engagement and learning outcomes. The interviews revealed that both teachers use visual GenAI to improve the educational experience; one Professor focuses on increasing student engagement during lectures with creative images, while the other assigns a group project where students need to generate a vision of a utopian or dystopian future. Despite the limitation of having participants from the same and also of the Geosciences faculty, their insights provide valuable perspectives on the integration of visual GenAI in academia.

4.3.1 Goal of the Study

The goal of this study was to gain insights into how teachers at Utrecht University make use of visual GenAI in their educational practices. Through interviews with two teachers, we aimed to explore their experiences, the application of visual GenAI in their courses, and its impact on student engagement and learning outcomes.

4.3.2 Recruitment Methods Interviews

One participant was an Assistant Professor in the Geosciences Faculty. He integrated visual GenAI into the *Sustainable Business Research Methods* course to improve student engagement during lectures. The other participant was also from the Geosciences Faculty. He used visual GenAI in a more in-depth way, assigning students to create a utopian or dystopian future as part of a

group project for the *Introduction to Technology and Innovation* course.

The teachers were recruited through convenience sampling: one was introduced through my supervisor, and the other was connected via mutual contacts.

4.3.3 Materials Interviews

Both interviews were semi-structured and focused on the teachers' experiences with visual GenAI. Both interviews were conducted and recorded using Microsoft Teams, each lasting approximately 30 minutes. The interview questions asked to these teachers are included in Appendix G.

4.3.4 Data Collection Rationale

We used semi-structured interviews to get a deeper understanding of the professors' thoughts and experiences with visual GenAI. This type of interview style allowed us to ask follow-up questions, helping us gather more detailed information about how they use and think about visual AI in their teaching.

4.3.5 Methodology of Analysis

To analyze the data from the interviews, we made use of inductive coding, which helps us identify patterns and themes directly from the data without predefined categories. The topics were categorized into two main themes: visual learning and student engagement. In Appendix F, Table F.2 presents the results of the inductive coding from the interviews.

4.3.6 Interview Analysis

Both teachers use visual GenAI to improve the educational experience for their students. The first participant increases student engagement during lectures by presenting creative images related to the course material. He includes a slide during the break that features "Easter eggs" connected to the topic, making the lectures more interactive and enjoyable. The second participant incorporates visual GenAI in a group assignment where students

must create a vision of an utopian or dystopian future. This task improves collaboration, as students share their visual interpretations with the class, explaining their ideas and outcomes, which may start further discussions.

4.4 Critical Assessment

It is important to consider some limitations of this study. Both teachers were from the Geosciences faculty, which presents a limitation for this study since we aim to implement the educational activity in a course within the Humanities faculty. The use of visual GenAI can vary across different disciplines, influencing the challenges, opportunities, and student demographics involved. However, their insights remain valuable as they offer a glimpse into how visual GenAI operates within academic courses, contributing to an understanding of its potential applications.

4.5 Workshop

In this section, we discuss the findings from the workshop conducted with professors from the Humanities Faculty at Utrecht University. The goal was to explain what visual GenAI is and how it can be applied in an educational setting. We aimed to stimulate the professors to explore the potential of visual AI tools for improving their teaching practices. The survey results of the workshop showed that many participants were interested in incorporating visual GenAI into their courses, particularly as a means to discuss biases and stereotypes with their students. However, they also expressed concerns about technical challenges and the necessity for training to effectively use these tools.

4.5.1 Goal of the Study

The goal of the workshop was to explain to the professors what visual AI is and how it can be applied in an educational setting. We also wanted to find out if the teachers were interested in applying visual GenAI in their own courses after the workshop and how they envisioned that happening. We aimed to show the capabilities of tools like Stable Diffusion and DALL-

E while raising their awareness of the biases these tools might have in the generated content. Additionally, we provided in-depth information by explaining how LoRa's works.

4.5.2 Recruitment Methods Interviews

The participants consisted of 14 teachers from the Humanities faculty, representing various linguistics and cultural courses. We specifically chose this faculty for our workshop because we wanted to implement the educational activity there, which is an example of purposive sampling.

4.5.3 Materials Workshop

For the workshop, we prepared a PowerPoint presentation, a list of discussion questions, and images generated with Midjourney, Stable Diffusion, and DALL-E. We demonstrated Stable Diffusion's ability to create images based on text prompts, such as cultural representations. LoRa was introduced and trained, showing how it improves image specificity and quality. Additionally, we created a survey using Qualtrics XM and printed paper versions for teachers who preferred a manual format. The workshop took place in a conference room at the Valk Hotel in Utrecht and lasted approximately 60 minutes.

4.5.4 Data Collection Rationale

After the workshop, we conducted a survey to gather feedback from the participants. This was to gather their reactions to the content and their willingness to integrate visual GenAI into their teaching in the future. The survey included questions about their experiences during the workshop, potential applications of visual AI in their courses, and any concerns about implementation.

4.5.5 Methodology of Analysis

In the analysis, we combined our observations from the workshop with the survey results. This allowed us to identify patterns in the teachers' attitudes

towards visual GenAI and its application in their teaching. We also made use of inductive coding to properly analyze our results. The main categories identified were the need for training, discussion of bias, and diverse learning styles. In Appendix F, Table F.3 presents the results of the inductive coding from the workshop observations and survey responses.

4.5.6 Workshop Analysis

The results showed that the teachers were interested in using visual GenAI as a tool to discuss biases and different cultural representations. 12 of the 14 professors indicated they wanted to use visual AI in their classrooms. At the same time, there were concerns about technical challenges and the need for training to effectively use these tools.

4.6 Critical Assessment

The study revealed that while teachers showed interest in using visual GenAI in the classroom, some encountered challenges in understanding biases in generated images and how to guide students effectively. To address these issues, future workshops could benefit from clearer training materials, additional resources like tutorials, and follow-up sessions. It would also be helpful to organize hands-on workshops where teachers can create images themselves to gain a deeper understanding of visual GenAI's capabilities and limitations, better preparing them to incorporate these tools into their teaching.

4.7 Concluding Interview

In this section, we discuss the final interview with course coordinator. This interview was conducted after the educational activity to gather her thoughts on the workshop, her views on using visual GenAI in education.

The course coordinator showed a strong interest in visual GenAI and its potential in education. She appreciated how the workshop stimulated critical thinking and reflection among students, especially regarding their

search terms and outputs, and she found the educational activity to be exactly what she was looking for. She believes visual GenAI can improve learning experiences and engage students more effectively than traditional assessments like multiple-choice quizzes. However, she also mentioned challenges, such as a lack of structured guidelines and training for teachers. Additionally, she expressed concerns about the environmental impact of (visual) GenAI and expressed for a more sustainable approach that prioritizes meaningful learning outcomes while reducing ecological harm.

4.7.1 Goal of the Study

The goal of the interview was to learn about the views of the course coordinator on visual GenAI after the educational activity has been conducted. The interview focused on her experiences, concerns, and vision for using visual GenAI in education.

4.7.2 Recruitment Method

The course coordinator was asked to participate in the concluding interview because of her role as course coordinator of the course in which we deployed the educational activity.

4.7.3 Materials Conducting Interview

The interview was semi-structured and focused on the thoughts of the course coordinator on visual GenAI after the educational activity. The interview was conducted and recorded on Microsoft Teams and lasted about 20 minutes. The questions that were asked during this interview have been added to Appendix H.

4.7.4 Data Collection Rationale

We conducted a semi-structured interview to gain a deeper understanding of her opinions and experiences with visual GenAI. This approach allowed for open responses and follow-up questions, providing insight into her thought process.

4.7.5 Methodology of Analysis

The interview data was analyzed by reviewing and summarizing responses of the course coordinator. Since the interview involved only one participant, inductive coding was not applied, as this method is typically used for larger datasets. Instead, we focused on identifying key themes and insights directly from her answers.

4.7.6 Interview Analysis

During the interview, the course coordinator shared that the educational activity was helpful and easy to follow. She appreciated how it stimulated students to think deeply and critically, which she believes is essential in today's learning environment. Furthermore, she mentioned that current rules about using GenAI in education often focus on what not to do, rather than showing students how to use it meaningfully. She hopes students can learn to use GenAI thoughtfully to truly understand the course material.

Although she has attended workshops on the use of GenAI in education, she still feels unsure about using it by herself in her teaching. Moreover, she believes GenAI could change course evaluations, moving from traditional tests to more interactive and meaningful learning experiences. She sees GenAI as a tool to assist teachers with small tasks and support students in brainstorming and critical reflection.

While she recognizes the benefits of visual GenAI, the course coordinator is concerned about its environmental impact and advocates for sustainable use, such as creating images only when necessary. She also emphasizes the need for training for teachers in visual GenAI to better stimulate students' critical thinking, as Utrecht University does not facilitate any training specifically for teachers on how to make use visual GenAI.

4.8 Critical Assessment

Reflecting on the interview with the course coordinator, it became clear that interviewing the second teacher present during the educational activity

would have provided additional valuable perspective. This would have offered a better understanding on the integration of visual GenAI in teaching and could have helped address any gaps or concerns that was raised.

5. Collaboration with Students

In this section, we provide an overview of the educational activity conducted to explore the impact of using visual GenAI in Humanities education. We discuss the goals, design, and data collection methods that were used in the activity, as well as the analysis of the feedback gathered from students. The section also offers a critical assessment of the activity, highlighting both the challenges encountered and the positive insights gained for future use of visual GenAI in academic settings.

5.1 Educational Activity

The educational activity aimed to explore the impact of visual GenAI tools like ChatGPT and Microsoft Copilot in Humanities education. It motivated students to think critically about visual GenAI's role in cultural representation and discuss biases with their classmates. Despite challenges with usability and prompt specificity, most students expressed satisfaction with the ease of use and educational value of the tools. However, the low survey response rate (only 8 out of 24 students) limited the study's findings. The results highlighted the importance of effective prompt creation and the need for improved training to maximize the potential of visual GenAI in education.

5.1.1 Goals of the Study

The goal of the educational activity was to explore the impact of visual GenAI on academic education within the Humanities Faculty. Through this activity, we gathered feedback from students about their perceptions of visual GenAI in their courses. It also provided an opportunity to evaluate our activity in terms of its possibilities and challenges.

5.1.2 Recruitment Methods Educational Activity

The participants were students from the course *Multilingualism: Language and Languages in a World Full of Differences*, a course from the Humanities Faculty, who also participated in the focus group and workshop, which is now our course coordinator. After being introduced to the topic through these prior studies, she agreed to collaborate with us in conducting an educational activity in one of her courses. A total of 24 students took part in the activity, primarily working in small groups of 2-3. Unfortunately, despite our request for individual responses, we received only 8 responses.

5.1.3 Design of the Activity

While designing the educational activity, we considered different factors, including the limited time of a single class session, a group size of 24 students working in small groups of 2-3 people, and the fact that students would have only a few opportunities to create images using ChatGPT or Copilot.

Like mentioned, the activity took place during a lesson at Utrecht University and was structured in several steps, each with its own goal and expected outcome. We began with an introduction to different types of GenAI, explaining what these tools can do and their potential limitations. This background provided students with the necessary information before they started using the technology.

Next, we prompted a discussion about the words "Family" and "Familia" by asking students for their definitions of "family." This sparked a variety of responses and initiated a discussion aligned with the course's focus on cultural perspectives, which we anticipated would be rich, given the multicultural backgrounds of the students in the classroom.

After this discussion, we generated an image as a class using the prompts provided by the students. Once we had gathered all the different prompts, we asked if there were any they disagreed with. We expected that the diversity of cultures represented in the classroom would lead to a lively discussion. Afterwards, we asked students to reflect on whether the generated image matched their expectations and interpretations of the term "Family"

and the capabilities of Visual GenAI.

For the final part of the activity, students selected a specific culture to work with and generated an image representing that culture. They had to create various prompts to depict that culture according to their own interpretations. Three examples of cultures generated by the students are presented in Appendix I.

Finally, they answered several assignment questions, including the SUS and educational software questionnaires. This last step was designed to give students hands-on experience with Visual GenAI and gather their feedback on its usability and educational value. The assignment questions and other survey questions are included in Appendix J.

5.1.4 Materials for Educational Activity

The materials used for the educational activity included Microsoft PowerPoint, Qualtrics XM, ChatGPT, and Microsoft Copilot, which both incorporates DALL-E.

5.1.5 Data collection Decisions

Data were collected using surveys, the SUS, and an educational software questionnaire. We chose the SUS to assess usability and the educational software questionnaire to evaluate the effectiveness of Visual GenAI for educational purposes, focusing on user satisfaction and learning efficacy. These methods aimed to gather feedback on how well the tools functioned in a Humanities course and their impact on the learning experience.

We considered giving the students a grade and having them submit an assignment to encourage them to take the assignment more seriously. However, we decided against this because it did not fit within the course structure of the course coordinator, as Visual GenAI or technology in general is not part of the course curriculum. Instead, we decided for surveys to gather feedback without putting extra pressure on the students. This way, they could engage more freely with the educational activity and provide honest insights into their experiences with Visual GenAI.

5.1.6 Methodology of the Analysis

Data analysis included observations from the workshop and responses from the surveys, SUS, and educational software questionnaire. We used inductive coding on the qualitative survey data to find main themes. The main themes identified were: Expectations and Outcomes, Specificity in Word Choice, Influence of AI Tools, Comparative Analysis, and Suggestions for Enhanced Prompting. These results can be found in Appendix J, Table J.4.

5.1.7 Analysis Educational Activity

The results showed differences in student satisfaction with using Visual GenAI. The main issues were prompt specificity and tool limitations; students had to be very specific in their prompts to get the results they wanted, and they only had a few attempts. They noticed gaps between their expectations and the actual generated visuals. Despite these limitations, most participants found the tools easy to use and effective for educational purposes, with many expressing satisfaction. However, since we only gathered 8 responses, these findings should be taken lightly.

5.2 Critical Assessment

The educational activity provided valuable insights but also highlighted several issues for future studies. One major problem was the low response rate, with only eight out of 24 students completing the surveys. This limited participation weakened our findings and made statistical analysis difficult, preventing clear conclusions about the effectiveness of the Visual GenAI tools.

Not using Stable Diffusion might have affected how deeply students engaged with the GenAI tools. Future studies should ensure the software is not only user-friendly but also aligns closely with educational goals. Stable Diffusion could have been more suitable for creating complex and culturally rich images, which would have enhanced the learning experience. It is also important to consider students' prior knowledge, as the 3P-model suggests that successful educational activities require familiarity with the tools

or material being taught.

Students had mixed feelings about their experiences. While some were satisfied with the outputs, others found a gap between expectations and results. Feedback indicated that being specific with prompts was key to achieving the desired visuals, showing a need for better training in prompt creation to improve the use of Visual GenAI tools.

Despite these challenges, there were positive takeaways. Overall, students were satisfied with the usability of the GenAI tools, particularly their ease of use and learning. Most participants felt they could quickly understand the software, suggesting these tools have potential for supporting educational goals in the future.

6. Discussion

In this section, we will analyze the main research question along with the sub-research questions. Each question will be examined in detail, allowing us to explore the findings and insights from our research. By addressing these questions individually, we aim to provide an understanding of how visual GenAI can be integrated into academic courses in the Humanities. We will also evaluate the educational activity using the 3P model, assessing its effectiveness in relation to the key elements of presage, process, and product to determine how well the educational activity achieved its educational objectives.

6.1 Existing educational models

To address the first research question, *What existing educational models incorporate visual generative artificial intelligence?*, we looked at several studies about using GenAI in education.

6.1.1 Visual GenAI in Teaching

Vartiainen and Tedre (2021) studied text-to-image GenAI in craft education. They found that using visual GenAI helps teachers visualize their ideas, which improves creativity in the design process. However, they also pointed out issues like algorithmic bias and copyright, indicating that we need to be careful when using these tools in teaching. This aligns with the findings from the survey with 399 students, who generally had a positive view of visual GenAI, seeing its potential for personalized learning and brainstorming. However, they expressed concerns about privacy and ethics, suggesting that educational models should include clear rules for responsible use of GenAI[28] [9].

Chan (2023) found that Generation Z students are optimistic about vi-

sual GenAI's ability to improve productivity and offer personalized content. However, teachers from older generations emphasized the need to balance technology with traditional methods to keep stimulating critical thinking and skill development. They raised concerns that reliance on GenAI could hinder the growth of evaluation skills among students. Additionally, workshops where visual GenAI is treated as a collaborative partner have been developed. For example, design students trained an AI model to generate visual content, improving their understanding and creativity. This type of hands-on experience demonstrates that visual GenAI can improve the creative process by providing high-quality images that inspire further new ideas[11].

6.1.2 Conclusion

In summary, various educational methods currently use visual GenAI, showing its potential to improve teaching. Some models demonstrate how visual GenAI helps teachers better visualize their ideas and stimulates student creativity. There are also examples where students use visual GenAI as a collaborative tool, improving their creative processes. However, these models also brings challenges, such as algorithmic bias and copyright concerns. Therefore, it is important to integrate visual GenAI responsibly into education to maximize benefits while addressing potential risks.

6.2 Essential Elements

To answer our second research question, *What are the essential elements for the successful integration of visual GenAI into education, including challenges and tangible outcomes?*, we conducted interviews and focus groups with teachers from Utrecht University. Our findings highlighted several key elements crucial for effectively using visual GenAI in education.

6.2.1 Collaborative Environment

Creating a collaborative environment among teachers is essential for successfully integrating visual GenAI into education. Teachers in the focus

groups mentioned that when they worked together — by sharing experiences, exchanging useful resources, and discussing best practices for visual GenAI — they achieved better results in adopting new technologies in their classrooms. This collaboration helped them to address challenges and share solutions regarding the application of visual GenAI.

6.2.2 Ecological Approach

An ecological approach when using visual GenAI is important. In the concluding interview, the course coordinator mentioned that generative tools are not very sustainable. She felt uncertain about using visual GenAI because of this issue. Therefore, when creating an educational activity, it is essential to think about the ecological impact of these tools. This means looking at the resources they use, the energy needed, and how to reduce any negative effects on the environment. By considering these factors, teachers can improve learning while also being sustainable.

6.2.3 Conclusion

To conclude, successful integration of visual GenAI in education requires stimulating collaboration among teachers and considering the ecological impact of tools used. By focusing on these elements, universities can achieve tangible outcomes such as improved teaching effectiveness, improved student engagement, more interactive learning experiences, and a more sustainable use of educational GenAI technologies. These outcomes reflect the practical benefits of incorporating visual GenAI into the curriculum. However, challenges such as ensuring sustainability, providing training resources, and addressing accessibility and technological limitations must still be considered.

6.3 Needs of Teachers and Students to Incorporate Visual GenAI

To answer our third subquestion, *What are the needs of teachers and students to incorporate Visual GenAI into humanities education?*, we organized a work-

shop, distributed a survey afterward, and conducted focus group discussions, interviews, and a concluding interview. These studies provided us with an understanding of the needs related to incorporating Visual GenAI into humanities education.

6.3.1 Guidelines for the Use of GenAI

An important need we identified is the creation of clear guidelines from the university on how to use (visual) GenAI tools. These guidelines should cover practical, ethical, and sustainability issues. It is important to explain how visual GenAI can be effectively integrated into the curriculum and recommend specific tools and methods that teachers and students can use in their courses.

Ethical considerations are also very important. The guidelines should highlight issues like copyright, data privacy, and potential biases when using generative tools. By providing this ethical framework, the university can help teachers and students understand their responsibilities and the effects of their actions when using these technologies. Additionally, sustainability is a factor that should not be ignored. The guidelines should promote practices that reduce the ecological impact of using visual GenAI tools, encouraging energy-efficient choices and methods that lower resource use during educational activities.

6.3.2 Training for Students and Teachers

Along with guidelines, training for both students and teachers is essential for successfully incorporating visual GenAI in humanities education. Training programs should cover the basics of visual GenAI, helping students and teachers understand how these tools work and how they can be used in the humanities. It is important to introduce both students and teachers to the different visual GenAI tools available for educational use, allowing them to choose the best tools for their specific needs.

The programs should also emphasize the importance of following the university guidelines when using visual GenAI tools. This way, students

and teachers will not only become skilled with the technology but also adhere to the ethical and sustainability standards set by the university. Although we did not gather enough information about the experience of students with visual GenAI, it seems many are not very familiar with these tools. Since visual GenAI is still relatively new, most students probably lack practical experience, making training even more critical. Additionally, because technology is changing rapidly, universities need to be flexible and up-to-date with their training programs.

6.3.3 Conclusion

In conclusion, to meet the needs of teachers and students for incorporating visual GenAI into humanities education, universities must establish clear guidelines and provide comprehensive training programs that are up-to-date. This will stimulate teachers and students to effectively engage with visual GenAI, improving their learning experiences while ensuring ethical and sustainable practices.

6.4 Evaluation 3P-Model

As we mentioned beforehand, we will make use of the 3P-model to assess whether visual GenAI was implemented successfully during the educational activity. As a reminder, the 3P-model exists out of three different stages, namely the predictive, process and product stage. We will discuss every stage step by step to provide a comprehended overview.

6.4.1 Predictive stage

According to the 3P-model, prior knowledge is essential for both teachers and students. Since we were responsible for deploying the educational activity, we took on the role of the teacher. Despite having prior knowledge of visual GenAI, we observed that the students lacked the necessary prior knowledge, which the 3P-model identifies as crucial for successful educational activity. This further emphasized the need for training for both students and teachers.

6.4.2 Process stage

In the process stage, students worked with ChatGPT and Microsoft Copilot to generate images. While students were engaged, their lack of prior knowledge in using visual GenAI tools limited their ability to fully make use of the GenAI tools. This showed the importance of training to equip students with the knowledge needed for effective use.

The educational activity also prompted discussions around the biases in the generated images, such as cultural and gender biases, which sparked critical thinking about the ethical implications of using visual GenAI in education. These discussions helped students understand the complexities of working with such technologies and encouraged reflection on their broader impact.

According to the 3P-model, successful learning requires a balance of prior knowledge, motivation, and engagement. In the process stage, the lack of prior knowledge hindered the students full engagement with the tools.

6.4.3 Product stage

In the product stage, a final interview with course coordinator provided insights into her experience with the educational activity. She appreciated how visual GenAI helped engage students in critical thinking, as they discussed biases and stereotypes both in class and in groups.

However, Dr. Pinto also mentioned some challenges, including a lack of clear guidelines and training for teachers, as well as concerns about the environmental impact of GenAI. Students showed a mix of interest and hesitation, noting that they struggled to create specific images they wanted and needed to prompt very precisely.

According to the 3P-model, the product stage evaluates the outcomes and effectiveness of the activity. While the activity stimulated critical thinking and engagement, Dr. Pinto's feedback highlighted the need for guidelines, training, and a sustainable approach to ensure both teachers and stu-

dents can confidently use visual GenAI in educational settings.

6.5 Integrating Visual GenAI in Humanities

Now that we have answered our sub-questions and analyzed the 3P-model in relation to our educational activity, we can finally address our main research question: *"How to integrate visual generative AI in creating an educational activity for academic courses in the Humanities?"*

The main key findings indicate that it is necessary for the university to facilitate training for both teachers and students. This training is crucial, as the current knowledge of both groups regarding the use of this technology is still too low. According to our analysis with the 3P-model, having sufficient prior knowledge is essential for creating a successful academic course or activity, which can be achieved through training. Currently, the university facilitates training for textual GenAI, but there is a absence of similar offerings for visual GenAI.

Moreover, it is important to establish clear guidelines for both teachers and students to use visual GenAI responsibly and in an ecological manner. The teachers expressed that they wanted guidelines, noting that the lack of direction makes it difficult to incorporate visual GenAI effectively into their courses. The guidelines should include practical and ethical considerations, such as addressing issues of copyright, ecological issues, data privacy, and potential biases in the generated content.

Specifically for the humanities, the results showed that teachers were very interested in using visual GenAI in their teaching. Feedback from the workshops indicated that they wanted to use it as a tool to discuss biases and stereotypes related to different cultures, including language and cultural biases. By using visual GenAI, teachers can create engaging discussions and learning experiences that help students think critically about how different cultures are represented in various media. This method also helps involve students more, as they can share their own ideas and create their own visual projects. Getting students involved in this way encourages them to think deeply and understand the topics better.

6.5.1 Conclusion

Thus, to integrate visual GenAI into creating an educational activity for academic courses in the humanities, the university should invest in comprehensive training for both teachers and students to effectively use these tools. Clear guidelines should be developed to ensure responsible use, especially concerning copyright and data privacy, which are important in the humanities. By using visual GenAI to spark discussions about cultural biases and representations in media, students can gain a deeper understanding of these complex issues, which also benefits society by preparing students to recognize and address bias-related problems. Focusing on these aspects will help universities successfully incorporate visual GenAI into humanities courses, improving critical thinking and increasing student engagement in their learning experience.

7. Conclusion

In this section, we discuss the limitations of our studies and discuss potential directions for future research. We consider additional studies that could have been conducted with more time, as well as key findings that provide a basis for further exploration. Finally, we highlight areas requiring deeper investigation to better understand how visual GenAI can be effectively integrated into academic education. This discussion reflects the structure of our methodology and results sections, addressing collaboration with teachers and students separately, and concluding with recommendations for future studies.

7.1 Limitations of Collaboration with Teachers

During our collaboration with teachers, there were some limitations that impacted the results of the project. It is important to understand these limitations when reviewing the results and considering how to improve future work.

7.1.1 Limitations of Focus Group

The primary limitation is that we conducted only one focus group, consisting of four teachers from Utrecht University. Literature suggests that five members is optimal for discussing human relations issues, and five to ten members is preferable for brainstorming purposes [51]. Additionally, the rule of thumb is to conduct at least four to six focus groups, as data tends to become saturated beyond this point, with little new information obtained from further groups [52]. Furthermore, our focus group included teachers from various faculties; a separate focus group exclusively with Humanities faculty members could have provided more specific insights to that discipline.

7.1.2 Faculty-Specific Interview Limitations

Another limitation is that the interviews we conducted involved only teachers from the Geosciences faculty, while our aim was to develop an educational activity for the Humanities faculty. This may have caused differences in the experiences and challenges mentioned, as different faculties are likely to encounter different benefits and obstacles when using visual GenAI in the classroom. Although certain challenges are likely to be shared across faculties, we believe that focusing exclusively on Humanities faculty would have provided more directly relevant data. Due to the difficulty of finding teachers who were both familiar with visual GenAI and willing to participate in our study, we had to adapt our approach.

7.1.3 Limitation of the workshop

We only conducted one workshop in which we gathered responses from 14 teachers at Utrecht University from the Humanities faculty. Although it provided us with many insights, the number of participants is still too low to draw any reliable conclusions from it.

7.2 Limitations of Collaboration with Students

The collaboration with students had certain limitations. These limitations also affected the results of the project.

7.2.1 Lack of Student Perspective

First, it would have been beneficial to conduct a focus group with students before designing the educational activity. We only incorporated teachers' perspectives, leaving out the viewpoints of students. For example, a focus group could have provided insights into students' initial understanding and expectations regarding visual GenAI.

7.2.2 Limitations in Creative Freedom

Thereafter, the decision not to use Stable Diffusion limited students' creative freedom. Stable Diffusion allows users more flexibility in creating a variety of images, which could have helped students depict cultural representations more precisely. It also could have improved their technical skills, which may prove valuable in the future in their education. Next to that, due to the limitations of free trials of ChatGPT and Microsoft Copilot, students had only a few attempts to create images that aligned with the assignment requirements.

7.2.3 Participant Experience Limitations

Another limitation was that our participants were first-year students, who generally have limited academic experience. This may have influenced their engagement and familiarity with the topic.

7.2.4 Low Survey Response Rate

Furthermore, only eight out of the 24 students responded to the survey, despite our instructions to complete the survey individually. Some of the responses appeared to be from groups, while other groups did not submit any responses. This could have been due to various factors, such as misunderstandings or a lack of engagement with the educational activity. While we received positive feedback regarding the activity, a higher response rate was essential for gathering sufficient quantitative data to support our findings. Given the low response rate, we cannot draw definitive conclusions from the students' perspective.

7.2.5 Broader Implementation

Finally, to gather more reliable data, this educational activity should be implemented across multiple courses and classes. Conducting it in only one class is insufficient to draw generalized conclusions. Although we gathered meaningful responses and results, a broader application would be necessary to establish more reliable findings.

7.3 Future Work

In the future, we can conduct several studies to improve our understanding and use of visual GenAI in education. The following sections will describe these studies and how they can improve educational practices and student engagement.

7.3.1 Establishing Guidelines

It is essential to track down the guidelines needed for inclusion in academic education. For this, multiple studies could be done to identify the needs of teachers and students, helping determine which guidelines should be considered. Adhering to ethical norms is essential in this process.

7.3.2 Incorporating Graded Assignments

To get more reliable data, future assignments could be tied to grades or real projects. Students may take assignments more seriously if there were rewards or consequences based on how they perform.

7.3.3 Gathering Perspectives for Educational Activities

To better design educational activities, it is essential to gather students' and teachers' perspectives on the use of visual GenAI in academia. This can be achieved through various methods such as surveys, interviews, or focus groups. Understanding both groups' knowledge of GenAI and their views on its role in education, particularly within the Humanities faculty, would be beneficial. This data can inform the creation of more relevant and effective educational activities.

Additionally, gathering immediate feedback after the activity through a different method, such as brief discussions, could provide more valuable insights than an online survey, helping to improve future implementations.

7.3.4 Exploring Stable Diffusion

Deploying an educational activity that incorporates Stable Diffusion could also be explored. Making use of Stable Diffusion allows for greater user freedom and can generate more specific outputs. In our current study, many parameters were controlled by the GenAI tool. It would be interesting to investigate how the results would change if we provided students with more creative freedom.

7.3.5 Expanding to Different Disciplines

Finally, using the educational activity in different academic fields could give us interesting results. While we focused on the Humanities, trying the activity in the Computer Science department could be helpful since those students might know more about GenAI tools. This group may have more experience with using tools like Stable Diffusion, leading to a better exploration of the technology. Also, teachers in these fields may understand the topic better, which could improve the overall learning experience.

7.3.6 Longitudinal Studies

Conducting longitudinal studies that track student engagement and understanding over time would also be beneficial. By assessing how students perspectives and skills evolve with repeated exposure to visual GenAI, we can better understand the long-term impact of these educational activities.

A. Appendix - Educational activity Multilingualism: Language and Languages in a World Full of Differences

A.1 Assignment

1. Introduction: In the master's program of Intercultural Communication, we will prepare an educational activity for first-year students of the course *"Multilingualism: Language and Languages in a World Full of Differences."* During this workshop, students will be introduced to different applications of visual Generative AI, using examples and generating content as a class with our guidance. We will hold discussions about interpretations of words, specifically the concept of *"family,"* as interpretations of family differ across various cultures.

The activity will begin by asking the students for their interpretation of the word *"family"* in a traditional context. Following this, we will collaboratively generate an image of a family using ChatGPT or Microsoft Bing, indirectly utilizing the visual generative tool DALL-E. Before generating the image, we will ask the class if anyone disagrees with the prompts being used, thereby encouraging a discussion within the class. Once the image of the family has been generated, we will ask the students if they agree with the generated image, which will lead to another discussion.

After that, we will show the class two generated images from DALL-E and Stable Diffusion, created using the same prompt, *"A traditional Italian family."* We will ask the class to share their observations about the images and discuss their interpretations of the prompt. Lastly, the students will be asked to generate prompts for a specific culture, using the prompts they have already created. They will then answer a few assignment questions,

and a System Usability Scale (SUS) and educational software survey will be conducted.

2. Objective: The objective of this educational activity is to introduce students to visual Generative AI, where they will depict different cultures using both positive and negative prompts. Additionally, the aim is to teach students about the biases present in Generative AI and the factors they should consider when generating images.

This activity will be aligned with the course *"Multilingualism: Language and Languages in a World Full of Differences,"* in which different cultures and languages are thoroughly discussed.

3. The Flow of the Course:

1. **Introduction**: Provide an overview of what Generative AI (GenAI) is and discuss the different types of Generative models.
2. **Discussion on the Concept of Family**: Engage the class in a discussion about the different interpretations of the word *"family."* Ask the students what they understand by the term and encourage them to share their perspectives.
3. **Image Generation Activity**: Collaboratively create an image of a family with the class. Ask multiple students to suggest different prompts to guide the generation process.
4. **Image Evaluation**: Display the generated image to the class and ask if they agree with it, prompting a discussion on their thoughts and interpretations.
5. **Comparison of AI-generated Images**: Show examples of the prompt *"A traditional Italian family having dinner"* generated by both DALL-E and Stable Diffusion. Ask the students to share their observations on the differences between the two images.
6. **Culture-specific Prompting Exercise**: Have the students select a specific culture and create image prompts based on it. Provide them with instructions, and then have them answer a few assignment questions related to the activity.

B. Appendix - Learning Preferences Across Generations

Table 1 An overview of distinct generational traits and characteristics

Characteristic	Baby-boomer ~ 1946—1960	Gen X ~ 1960—1980	Gen Y/ Millenials ~ 1980—1995	Gen Z ~ 1995—2012
Teaching preference	Traditional lecture format	Combination of traditional and technology-based methods	Interactive, self-paced, technology-based methods	Hybrid (blended) learning, technology-focused
Learning style	Teacher-centered, note-taking, memorization	Collaborative, project-based, real-world application	Collaborative and networked, technology-based	Learn through images/videos/audio instead of text Experiential active Learning
Technology	Early information technology (IT) adaptors	Uses with confidence Digital immigrants	Part of everyday life, intuitive Digital Natives	Digital-first Technoholics
Social media	N/A	Some use for personal communication	High use for personal and professional communication	Integrated into daily life
Feedback	Once per year, during the annual review	Weekly/daily	On demand	Consistent, immediate and frequent
Communications approaches	Telephone	Email and text messages	Text or social media	Hand-held communication devices
Communication preference	Face to face ideally, but telephone or email if required	Text messaging or email	Online and mobile texting	Facetime
Knowledge sharing	Willingly, voluntarily	Based on mutuality and cooperation	Only in cases of self-interest or if forced	On virtual level, easily and rapidly, no stake, publicly
Values	Patience, soft skills, respect for traditions, EQ, hard work	Hard work, openness, respect for diversity, curiosity, practicality	Flexibility, mobility, broad but superficial knowledge, success orientation, creativity, freedom of information takes priority	Live for the present, rapid reaction to everything, initiator, brave, rapid information access and content search
Attitude towards career	Organisational – careers are defined by employers	Early "portfolio" careers – loyal to profession, not necessarily to employer	Digital entrepreneurs – work "with" organisations and not "for" organisations	Career multitaskers – will move seamlessly between organisations and "pop up" businesses
Aim AND ASPIRATION	Solid existence Job security	Multi-environ-ment, secure position Work life balance	Freedom and flexibility	Live for the Present
Relationship	First and foremost personal	Personal and virtual networks	Principally virtual, network	Virtual and superficial
View	Communal, unified thinking	Self-centered and medium-term	Egotistical, short-term	No sense of commitment, be happy with what you have and live for the present
Problem solving	Horizontal	Independent	Collaborative	Entrepreneurial
Teamwork	Unknown	Natural environment (multinational companies)	On a virtual level (only if forced)	Virtual and rapid
Change management	Change = caution	Change = opportunity	Change = improvement	Change = expected
Training	Preferred in moderation	Required as necessary	Continuous and expected	Ongoing and essential
Behavior	Challenge the rules	Change the rules	Create the rules	Customize the rules

Figure B.1: Different generations learning preferences [11].

C. Appendix - Educational activity Migrazioni Al Plurale

C.1 Assignment

1. Introduction: As part of the Migrazioni Al Plurale course, we will delve into the migration history and cultural heritage of Italian migrants in the Netherlands. Cultural heritage includes different aspects such as language, habits, traditions, and expressions. Some of these expressions, situations, or words are difficult to explain with sentences or words alone, such as the Dutch word 'gezellig'. Pictures can often tell a story better than text. When people read or listen, they often use their imagination to visualize and understand the story better. That is perhaps why we frequently say, "You must see it" or "You must visualize it" to explain certain situations.

Last year, students created media, including a video interview with first-generation Italians in Utrecht. Building on this, we will now explore the intersection of technology and culture by making use of generative visual AI. During this workshop, we will explain how to make use of generative visual AI, called Stable Diffusion. Make sure that you have installed stable diffusion **before the workshop**. This will save you a lot of time during the seminar. The installation will take you around 15 to 30 minutes. A guide on how to install and make use of Stable diffusion has been added in the appendix. If you have any questions related to the installations, don't hesitate to contact me.

Afterward, short practice assignments will be given to practice with GenAI. As last you will work together with your fellow students to answer questions. Lastly, you will work on your assignment and hand in the deliverable before XX-XX-2024 XX:00. Through this assignment, we aim to provide a platform for Italian migrants to share their stories and experiences in a

visually interesting way, creating a bridge from the older to the younger generation.

2. Objective: The objective of this assignment is that you get familiar with the use of generative visual AI, particularly Stable Diffusion. You will get to know more about prompting and how to create images. Thereafter you may recognize generated images quicker after you have generated a few images. Lastly, you will get to know about biases that are integrated in generative AI tools. You will explore and depict the cultural heritage and migration history of Italian migrants in the Netherlands. Through this exercise, you will gain a deeper understanding of cultural expressions, traditions, and experiences, while also learning how to effectively use generative AI as a storytelling tool.

3. Instructions: In the seminar, we will start on how to make use of stable diffusion. If you were not able to install Stable Diffusion on your laptop, we will help you with the installation after the instructions. The instruction will include explanations on how to generate images, what kind of checkpoints you could use, and how to prompt to generate an image accordingly.

Practice Assignments:

- Work in teams to complete short practice assignments, focusing on generating specific situations using generative AI. Examples include:
    1. Example Q1: Generate a typical Dutch household.
    2. Example Q2: Generate a family having dinner during Christmas.
- Discuss among your team why you chose certain positive or negative prompts and how they relate to the context.
- Do you notice any type of bias? Discuss whether you had some biases for yourself and the tool.

4. Individual Assignment: You will have to complete an individual assignment after the workshop and hand this in before the deadline, set on (specific time). In this assignment, you will have to make use of generative visual AI to create images reflecting various aspects of cultural heritage

from the perspective of Italian migrants in the Netherlands. You already practiced with some examples during the workshops, so try to make use of that knowledge.

Categories to explore include language, habits, traditions, and expressions.

Generate images that contribute to completing the migration history of Italian migrants, reflecting mainly on their experiences (I don't know if they are going to conduct interviews this year).

5. Submission Guidelines:

1. Submit your completed generative AI images along with a brief explanation of the cultural heritage they represent, and the prompts used to generate them. Ensure to include details on why you have decided to make use of certain positive and negative prompts.
2. Ensure that your submission is organized, well-documented, and demonstrates thoughtful consideration of the cultural context.

Also, make sure that you have answered the following questions:

1. To what extent did your images accurately reflect the experiences and traditions of Italian migrants?
2. What kind of biases did you notice in your results? Why? Do you see any biases in your prompting or is it the AI?
3. Do you think that the generated images provide the same as real images? Elaborate why.

C.2 MacOS

1. Go to brew.sh and copy the command link to install the software on your laptop.
2. Open your terminal, paste the command there, and press enter.
3. Confirm your password and wait for the installation to be completed.
4. To install the required programs, write and run the following com-

mand: 'brew install cmake protobuf rust python@3.10 git wget'

5. Clear your terminal by typing 'clear' and pressing enter.
6. Run the following command: 'git clone https://github.com/AUTOMATIC1111'
7. Run the command: 'open .' to view your installed files.
8. Navigate to the 'Models' > 'Stable-diffusion' folder. Here, you can add the checkpoints you want to use for creating images. Install various checkpoints depending on your requirements.
9. Install checkpoints from civitai.com or huggingface.co, depending on your needs.
10. After installing a checkpoint, add it to the stable-diffusion folder. You can install and use multiple checkpoints.
11. Clear your terminal again by typing 'clear' and pressing enter.
12. To open stable diffusion, type 'cd stable-diffusion-webui/' (you can use tab completion after typing 'cd st' to see the full command).
13. To run it, type './webui.sh' and press enter.
14. Your installation is now complete and ready to use in your standard browser.

C.3 Windows

1. You will need to install git from the website: https://git-scm.com/download/win, and download the latest version. You can stick with all the defaults.
2. You will also need to install Python on your laptop. Make sure you download a 3.10 version of python. The link for installing python 3.10 is: https://www.python.org/downloads/release/python-31014/
3. When you are installing Python, make sure to add Python 3.10 to PATH.
4. Open File explorer and navigate to where you want to install stable diffusion, and create a new folder, named stable-diffusion.
5. Click on the folder and write down cmd on the address bar.

6. Run 'Git clone https://github.com/AUTOMATIC1111 in the terminal
7. After running it, you will see that a new folder is created named 'stable-diffussion-webui
8. You can install checkpoints from civitai.com or huggingface.co, depending on what you are looking for.
9. Open the 'Models' > 'Stable-diffusion' folder. Here, you can add the checkpoints that you want to use to create images. You can install various checkpoints depending on what you want to generate.

D. Appendix - WebUI Stable Diffusion and CivitAI

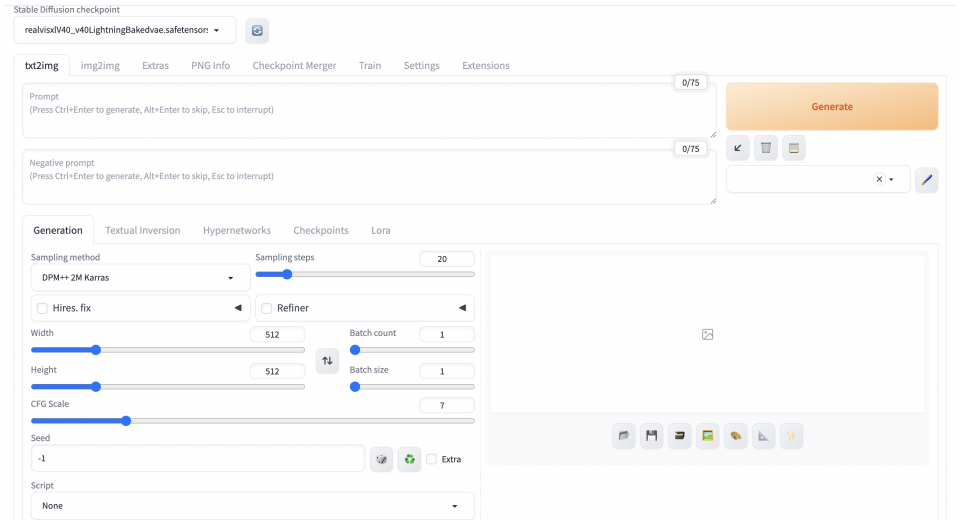


Figure D.1: WebUI of stable diffusion

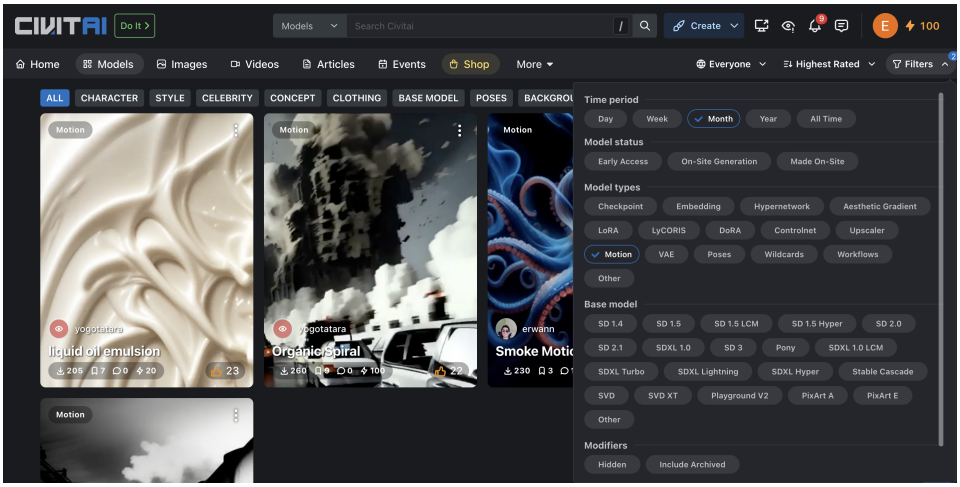


Figure D.2: WebUI of CivitAI

E. Appendix - Focus Group Questions

1. How much do you know about visual AI or textual AI technology?
    (a) ...Follow-up: Have you ever used or encountered these technologies before?
2. Do you use AI in your teaching? If yes, how, and why (motivation)? If not, do you think you will use it in the future?
    (a) ...Follow-up: What might make you want to use AI in your teaching, or what might make you hesitant?
3. What are the benefits of using visual AI in education?
    (a) ...Follow-up: How might this benefit students?
4. What problems or difficulties do you think might come up if teachers tried to use visual AI or textual AI in schools?
    (a) ...Follow-up: How could we solve these problems?
5. How do you think using visual AI or textual AI could make students more interested in learning and help them learn better?
    (a) ...Follow-up: Can you give some examples of how this might work?
6. How could visual AI help students who learn in different ways or who have different ways of learning?
    (a) ...Follow-up: How could we make sure the AI works well for everyone?
7. What help or training do you think teachers would need to use visual AI or textual AI well?
    (a) ...Follow-up: How could we give teachers the help they need?
8. How do you think using visual AI or textual AI might change the role

of the teachers in the classroom?

(a) …Follow-up: How might teachers prepare for these changes, and what new things might they need to learn?

F. Appendix - Inductive Coding Analysis

Code	Category	Summary of Key Points
Need for Training	Training	Teachers emphasized the need for training to effectively use Visual GenAI, as they currently lack sufficient knowledge and skills to integrate these tools.
Responsible and Ethical Use	Ethics	Teachers stressed the importance of clear guidelines for responsible and ethical use of Visual GenAI to ensure both effective teaching and learning.
Engagement and Learning Outcomes	Engagement	Teachers discussed how using creative visual elements can make lessons more interactive, improving student engagement and learning outcomes.

Table F.1: Overview of inductive coding results from the Focus Group.

Code	Category	Summary of Key Points
Visual Learning	Learning Enhancement	Both professors use visual GenAI to make their lectures more engaging by incorporating creative visuals or group assignments requiring student creativity.
Student Engagement	Engagement	Professors emphasize the importance of engagement, with one using visual AI to present creative images during lectures and the other using group assignments for students to create visual interpretations, fostering a more interactive environment.

Table F.2: Overview of inductive coding results from the interviews.

Code	Category	Summary of Key Points
Need For Training	Training	Teachers expressed the need for training and online tutorials to integrate visual AI into their teaching effectively.
Discussion Bias	Bias	Teachers discussed how visual AI examples could be used to spark conversations about stereotypes and biases in the classroom.
Diverse Learning Styles	Learning Styles	Teachers highlighted that visual GenAI supports various learning styles and helps visual learners understand complex ideas.

Table F.3: Overview of inductive coding results from the workshop and survey responses.

Code	Category	Summary of Key Points
Expectations and Outcomes	Expectations	Participants expressed varying satisfaction with visual GenAI output, with some satisfied ("Everything came out well") and others pointing out discrepancies between their expectations and the results.
Specificity in Word Choice	Word Choice	Participants highlighted the importance of being specific in prompts to avoid misinterpretation by the AI, with some noting that careful wording led to satisfactory results.
Influence of AI Tools	Tool Influence	Differences in generated images were influenced by the AI platform used, with participants noting that different tools (e.g., ChatGPT vs. Microsoft Copilot) led to significantly different outputs.
Comparative Analysis	Visual Differences	Participants compared their visualizations, noting how more specific prompts led to calmer, more loving visuals, while less detailed prompts produced more aggressive images.
Suggestions for Enhanced Prompting	Prompting Improvement	Participants suggested incorporating more details and specifying visual aspects more clearly to improve results in future attempts.

Table F.4: Overview of inductive coding results from the educational activity.

G. Appendix - Interview Questions with Experienced visual GenAI Teachers

1. Could you provide an overview of how you currently use generative visual AI in your classroom? (and what kind of tools are you using, textual/visual).
2. How did you change your teaching to integrate generative visual AI?
3. What were the biggest challenges you encountered when integrating generative visual AI into your curriculum?
4. What do teachers need to effectively integrate generative visual AI into their teaching methods?
5. Can you give specific examples of how generative visual AI has impacted student learning outcomes?
6. What feedback do you receive from students about the use of generative visual AI in your classroom?
7. How do you see the future of generative visual AI in education/academia?
8. Do you have any other insights or experiences you would like to share about using generative visual AI in education?

H. Appendix - Concluding Interview Questions with Manuela Pinto

1. Could you provide an overview of how you currently use generative visual AI in your classroom? (and what kind of tools are you using, textual/visual).
2. How did you change your teaching to integrate generative visual AI?
3. What were the biggest challenges you encountered when integrating generative visual AI into your curriculum?
4. What do teachers need to effectively integrate generative visual AI into their teaching methods?
5. Can you give specific examples of how generative visual AI has impacted student learning outcomes?
6. What feedback do you receive from students about the use of generative visual AI in your classroom?
7. How do you see the future of generative visual AI in education/academia?
8. Do you have any other insights or experiences you would like to share about using generative visual AI in education?

I. Appendix - Student-Generated Cultural Representations

Figure I.1: Image Generated by Student: Representation of Dutch Culture

Figure I.2: Image Generated by Student: Representation of French Culture

Figure I.3: Image Generated by Student: Representation of Maori Culture

J. Appendix - Survey Educational activity Multilingualism: Language and Languages in a World Full of Differences

Demographics

1. What is your gender? (MP)
2. What is your profession? (Categories, MP)
3. What is your highest education level? (MP)
4. What is your age group? (MP)

Assignment questions

1. What words did you use to achieve an image where you were happy with? why did you not use certain words?
2. What kind of visuals did you have in mind when entering your AI prompt, and why? Did the AI tool generate visuals that matched your expectations? If not, what changes did you make to your prompts to reach your goal?
3. What were their expectations from the AI? What did work (which words), and what did not work?
4. What differences in word choice and visuals do you see between yourself and your teammates? Why did team members visualize different images?
5. How would you have approached prompting differently if you had additional attempts? What specific changes would you make to your wording or approach to achieve better results?

Open questions

1. How familiar are you with visual Generative AI (e.g., AI-generated images)?
2. Did you use visual GenAI for your projects (MP), how many times have you done so?
3. What concerns might you have about including AI-generated content in your academic assignments?

Educational software questions (statements, Likert scale)

1. Overall, I am satisfied with how easy it was to use the AI system.
2. The AI system is simple to use.
3. I felt comfortable using the AI system.
4. It was easy to learn how to use the AI system.
5. I believe I became productive quickly by using the AI system.
6. The information provided with the AI system is clear.
7. It is easy to find the information I need in the AI system.
8. The information is effective in helping me complete tasks and scenarios.
9. The interface of the AI system is pleasant.
10. I like using the interface of this AI system.
11. The AI system has all the functions and capabilities I expected.
12. I believe that using the AI system helped me understand the subject better.
13. I would recommend using the AI system to fellow students.
14. Overall, I am satisfied with the AI system.

System Usability Scale Questions (statements, Likert scale)

1. I think I would like to use this AI system frequently.

2. I found the AI system unnecessarily complex.
3. I thought the AI system was easy to use.
4. I think I would need the support of an expert to use the AI system effectively.
5. I thought there was too much inconsistency in the AI system.
6. I would imagine that most people would learn to use this AI system very quickly.
7. I found the AI system cumbersome to use.
8. I felt very confident while using the AI system.
9. I needed to learn a lot of things before I could start using the AI system effectively.

Bibliography

[1] G. Eysenbach *et al.*, "The role of chatgpt, generative language models, and artificial intelligence in medical education: A conversation with chatgpt and a call for papers," *JMIR Medical Education*, vol. 9, no. 1, e46885, 2023.

[2] M. Muller, L. B. Chilton, A. Kantosalo, C. P. Martin, and G. Walsh, "Genaichi: Generative ai and hci," in *CHI conference on human factors in computing systems extended abstracts*, 2022, pp. 1–7.

[3] R. Bommasani, D. A. Hudson, E. Adeli, *et al.*, "On the opportunities and risks of foundation models," *arXiv preprint arXiv:2108.07258*, 2021.

[4] A. Voynov, K. Aberman, and D. Cohen-Or, "Sketch-guided text-to-image diffusion models," in *ACM SIGGRAPH 2023 Conference Proceedings*, 2023, pp. 1–11.

[5] I. Todorovic, "Importance of new class teaching methods in curricula development in developing countries," in *Handbook of research on enhancing innovation in higher education institutions*, IGI Global, 2020, pp. 408–427.

[6] B. Kollöffel, "Exploring the relation between visualizer–verbalizer cognitive styles and performance with visual or verbal learning material," *Computers & Education*, vol. 58, no. 2, pp. 697–706, 2012.

[7] I. Adeshola and A. P. Adepoju, "The opportunities and challenges of chatgpt in education," *Interactive Learning Environments*, pp. 1–14, 2023.

[8] Ö. Aydın and E. Karaarslan, "Openai chatgpt generated literature review: Digital twin in healthcare," *Aydın, Ö., Karaarslan, E.(2022). OpenAI ChatGPT Generated Literature Review: Digital Twin in Healthcare. In Ö. Aydın (Ed.), Emerging Computer Technologies*, vol. 2, 2022.

[9] C. K. Y. Chan and W. Hu, "Students' voices on generative ai: Perceptions, benefits, and challenges in higher education," *International Journal of Educational Technology in Higher Education*, vol. 20, no. 1, p. 43, 2023.

[10] A. Asperti, D. Evangelista, and E. Loli Piccolomini, "A survey on variational autoencoders from a green ai perspective," *SN Computer Science*, vol. 2, no. 4, p. 301, 2021.

[11] C. K. Y. Chan and K. K. Lee, "The ai generation gap: Are gen z students more interested in adopting generative ai such as chatgpt in teaching and learning than their gen x and millennial generation teachers?" *Smart Learning Environments*, vol. 10, no. 1, p. 60, 2023.

[12] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, and A. A. Bharath, "Generative adversarial networks: An overview," *IEEE signal processing magazine*, vol. 35, no. 1, pp. 53–65, 2018.

[13] P. Antoniadis and P. Antoniadis. "Latent space in deep learning." Baeldung on Computer Science. (2024), [Online]. Available: `https://www.baeldung.com/cs/dl-latent-space` (visited on 05/27/2024).

[14] R. Parihar, A. Dhiman, and T. Karmali, "Everything is there in latent space: Attribute editing and attribute style manipulation by stylegan latent space exploration," in *Proceedings of the 30th ACM International Conference on Multimedia*, 2022, pp. 1828–1836.

[15] S. Zhu, Z. Wang, Y. Zhuang, *et al.*, "Exploring the impact of chatgpt on art creation and collaboration: Benefits, challenges and ethical implications," *Telematics and Informatics Reports*, p. 100 138, 2024.

[16] J. An, W. Ding, and C. Lin, "Chatgpt," *tackle the growing carbon footprint of generative AI*, vol. 615, p. 586, 2023.

[17] C. Zhang, C. Zhang, M. Zhang, and I. S. Kweon, "Text-to-image diffusion model in generative ai: A survey," *arXiv preprint arXiv:2303.07909*, 2023.

[18] V. Kulikov, S. Yadin, M. Kleiner, and T. Michaeli, "Sinddm: A single image denoising diffusion model," in *International Conference on Machine Learning*, PMLR, 2023, pp. 17 920–17 930.

[19] H. Cao, C. Tan, Z. Gao, *et al.*, "A survey on generative diffusion models," *IEEE Transactions on Knowledge and Data Engineering*, 2024.

[20] Andrew. "Stable diffusion samplers: A comprehensive guide." (mei 2024), [Online]. Available: `https://stable-diffusion-art.com/samplers/`.

[21] A. Iskender, "Holy or unholy? interview with open ai's chatgpt," *European Journal of Tourism Research*, vol. 34, pp. 3414–3414, 2023.

[22] H. Alkaissi and S. I. McFarlane, "Artificial hallucinations in chatgpt: Implications in scientific writing," *Cureus*, vol. 15, no. 2, 2023.

[23] K. Konstantinos and E. Lambrou, "Chatgpt–another step towards the digital era or a threat to fundamental rights and freedoms?" *Pravo-teorija i praksa*, vol. 40, no. 3, pp. 1–18, 2023.

[24] A. Susarla, R. Gopal, J. B. Thatcher, and S. Sarker, "The janus effect of generative ai: Charting the path for responsible conduct of scholarly activities in information systems," *Information Systems Research*, vol. 34, no. 2, pp. 399–408, 2023.

[25] P. P. Ray, "Chatgpt: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope," *Internet of Things and Cyber-Physical Systems*, 2023.

[26] A. Abid, M. Farooqi, and J. Zou, "Persistent anti-muslim bias in large language models," in *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, 2021, pp. 298–306.

[27] R. Srinivasan and K. Uchino, "Biases in generative art: A causal look from the lens of art history," in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 2021, pp. 41–51.

[28] H. Vartiainen and M. Tedre, "Using artificial intelligence in craft education: Crafting with text-to-image generative models," *Digital Creativity*, vol. 34, no. 1, pp. 1–21, 2023.

[29] M. Hernandez-de-Menendez, C. A. Escobar Díaz, and R. Morales-Menendez, "Educational experiences with generation z," *International Journal on Interactive Design and Manufacturing (IJIDeM)*, vol. 14, no. 3, pp. 847–859, 2020.

[30] C. Seemiller and M. Grace, "Generation z: Educating and engaging the next generation of students," *About campus*, vol. 22, no. 3, pp. 21–26, 2017.

[31] M. Alier, F. García-Peñalvo, and J. D. Camba, "Generative artificial intelligence in education: From deceptive to disruptive," 2024.

[32] K. Karpouzis, D. Pantazatos, J. Taouki, and K. Meli, "Tailoring education with genai: A new horizon in lesson planning," *arXiv preprint arXiv:2403.12071*, 2024.

[33] A. Author(s), "Gazing at your gaze: Towards a reflexive creative practice through personalized ai models," *DIS '24: ACM Designing Interactive Systems Conference*, p. 18, 2024. DOI: X.X.

[34] J. Biggs and P. Moore, "Conceptions of learning and teaching," *The process of learning*, pp. 20–26, 1993.

[35] P. Kanashiro, E. S. Iizuka, C. Sousa, and S. E. F. Dias, "Sustainability in management education: A biggs' 3p model application," *International Journal of Sustainability in Higher Education*, vol. 21, no. 4, pp. 671–684, 2020.

[36] A. Kolmos, "Problem-based and project-based learning," in Jan. 2009, pp. 261–280, ISBN: 978-0-387-09828-9. DOI: 10.1007/978-0-387-09829-6_13.

[37] B. Dennison, "Transdisciplinary literacy: Seven principles that help define transdisciplinary research," *Integr. Appl. Netw*, 2017.

[38] F. Wickson, A. Carew, and A. Russell, "Transdisciplinary research: characteristics, quandaries and quality," *Futures*, vol. 38, no. 9, pp. 1046–1059, Nov. 2006. DOI: 10.1016/j.futures.2006.02.011. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S0016328706000553.

[39] R. G. Klaassen, "Interdisciplinary education: A case study," *European journal of engineering education*, vol. 43, no. 6, pp. 842–859, 2018.

[40] *Multi-, inter-, and transdisciplinarity; what is what?* [Online]. Available: https://www.uu.nl/en/education/educational-development-training/knowledge-dossiers/interdisciplinary-education-and-cel/multi-inter-and-transdisciplinarity-what-is-what.

[41] C. Lyall, A. Bruce, J. Tait, and L. Meagher, *Interdisciplinary research journeys: Practical strategies for capturing creativity*. Bloomsbury Academic, 2011.

[42] J. Balsiger, "Transdisciplinarity in the class room? simulating the co-production of sustainability knowledge," *Futures*, vol. 65, pp. 185–194, 2015.
[43] *What is civitai?* Accessed: 2024-11-04. [Online]. Available: https://civitai.com/content/guides/what-is-civitai.
[44] Civitai, *Checkpoint*, Accessed: 2024-11-06, 2024. [Online]. Available: https://wiki.civitai.com/wiki/Checkpoint.
[45] Civitai, *Everything i know about lora*, Accessed: 2024-11-06, 2024. [Online]. Available: https://civitai.com/articles/60/everything-i-know-about-lora.
[46] Civitai, *Lycoris*, Accessed: 2024-11-06, 2024. [Online]. Available: https://wiki.civitai.com/wiki/LyCORIS.
[47] Civitai Education. "Using civitai: A guide." (n.d.), [Online]. Available: https://education.civitai.com/using-civitai-a-guide/.
[48] D. Benyon, *Designing User Experience: A Guide to HCI, UX and Interaction Design*, 4th. Pearson, 2019, ISBN: 9781292155517.
[49] J. Nielsen, "Ten usability heuristics," 2005.
[50] D. R. Thomas, "A general inductive approach for qualitative data analysis," 2003.
[51] E. F. Fern, "The use of focus groups for idea generation: The effects of group size, acquaintanceship, and moderator on response quantity and quality," *Journal of marketing Research*, vol. 19, no. 1, pp. 1–13, 1982.
[52] D. L. Morgan, "Focus groups," *Annual review of sociology*, vol. 22, no. 1, pp. 129–152, 1996.