

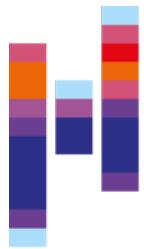
Hochschule
Flensburg
University of
Applied Sciences

Generative Artificial Intelligence

Generative Künstliche Intelligenz

M. Sc. Angewandte Informatik

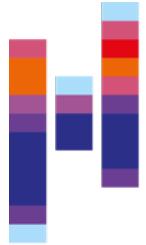




Hochschule
Flensburg
University of
Applied Sciences

But...

What is generative artificial intelligence?



A lecture at the right time

- Generative AI is one of the big trending technologies, and said to change the way we work and live.
- It is based on two key enablers:
 - The availability of large amounts of data and
 - The availability of compute.
- Both is not available for us, though ;-)



Examples of GenAI: LLMs

■ Large Language Models as Chatbots

Count the number of "r"s in the word "raspberries" for me.

Hide code ↗

Analysis ▾

Python

```
word = "raspberries"
count = word.count("r")
print(f"The number of 'r's in the word '{word}' is {count}.")
```

Code output

The number of 'r's in the word 'raspberries' is 3.

The letter "r" appears 3 times in the word "raspberries".

Like Comment Copy Share More

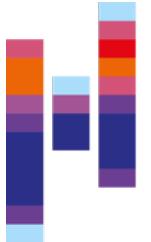


Examples of GenAI: Text-to-Video

- + A student at HS Flensburg is coming late to class. To avoid the professor being angry, he presents a magic trick.



(OpenAI sora, Aug 2025)



Hochschule
Flensburg
University of
Applied Sciences

Examples of GenAI: Text-to-Image

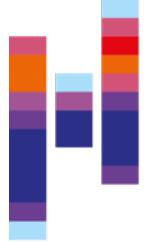
FLUX.1 Krea [dev]

FLUX.1 Krea [dev] model further tuned and customized with [Krea \[non-commercial license\]](#) [\[blog\]](#) [\[model\]](#)

A pink bunny wearing an astronaut's riding a frog on moon. In the background is Earth.

Run





Examples of GenAI: Text-to-Music

■ suno.ai Music Generation

The screenshot shows the suno.ai workspace interface. On the left, there's a sidebar with options like 'Simple' and 'Custom' modes, and a 'Write Lyrics' button. Below that are sections for 'Lyrics' and 'Styles'. The 'Lyrics' section contains a snippet of lyrics: "Splitin' up cells, oh can't you see, Mitosis keepin' us company. Win or lose, we'll stomp along. At the MyDOG challenge, sing this song! [Outro – Fiddle solo with stumps & claps] Yeehaw! Divide again, boys!". The 'Styles' section has a dropdown menu set to 'bluegrass'. Under 'Advanced Options', there are sliders for 'Weirdness' (set to 50%) and 'Style Influence' (set to 50%). A 'Create' button is at the bottom. The main area is titled 'My Workspace' and lists several generated songs:

- Do Se Bane Do (4:55) - Indian folk song (shot, tabla, harmonium, clapping circles)
- Mitoza u Scru (4:55) - Serbian turbo feld
- Mitoza u Scru (4:55) - Serbian turbo feld
- Splitin' up cells (4:55) - Serbian turbo feld
- Splitin' up cells (4:55) - Serbian turbo feld
- Well the phopase comes, chromosomes... (4:55) - bluegrass
- Code and Curiosity (Edit) (Remastered) (R... (4:55) - calm, laid back
- Code and Curiosity (Edit) (Remastered) (R... (4:55) - calm, laid back
- Code and Curiosity (Edit) (Remix) (4:55) - A calm electronic track

Each song entry includes an 'Edit' button, a 'Publish' button, and other sharing options. The bottom of the screen shows a playback bar for a song named 'Hoi-dri Mitoosaa! (Der MIDOG-Jodler)'.

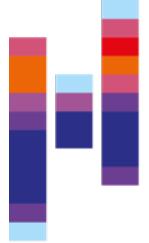
Some examples:



"Divide and Conquer"

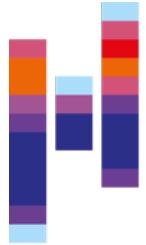


Hoi-dri Mitoosaa!
(Der MIDOG-Jodler)



Generative Artificial Intelligence: Definition

- Generative Artificial Intelligence Systems are models that can generate new and realistic data from pure entropy (randomness).
- In particular, they
 - combine randomness with patterns learned from training data, and
 - use conditioning signals (priors), such as textual descriptions, to guide the generation process.

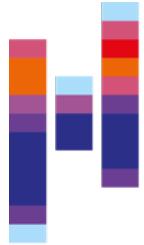


But...



Hochschule
Flensburg
University of
Applied Sciences

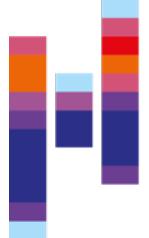
How does this all work?



Hochschule
Flensburg
University of
Applied Sciences

Plan for this semester

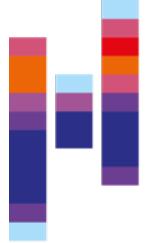
(might change a little bit, it's the first time I give this lecture)



Part 1: Introduction / Model Families

- 1. Foundations & Core Concepts
 - Generative vs. discriminative models
 - Latent variable modeling & representation learning
 - Dimensionality reduction & interpretability: PCA, t-SNE, UMAP, linear probing
 - Learning paradigms: supervised, unsupervised, self-supervised

- 2. Generative Model Families (Compact History → Modern Focus)
 - Autoregressive models (RNNs, Transformers → GPT)
 - VAEs (concepts, posterior collapse)
 - GANs (core idea, conditional GANs, StyleGAN, but kept concise)
 - Diffusion models (denoising, latent diffusion, scaling to large models)



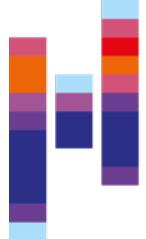
Part 2: Text and Image Generation

■ 3. Text Generation (LLMs)

- Pretraining objectives: autoregression, masked LM, causal vs. bidirectional models
- Fine-tuning strategies: supervised fine-tuning, RLHF, instruction tuning, parameter-efficient tuning (LoRA, adapters)
- Scaling laws & data efficiency
- Mixture of Experts (MoE) and sparse models
- Current trends: retrieval-augmented generation, tool use, agents

■ 4. Image Generation

- Latent Diffusion Models (Stable Diffusion, Imagen, DALL·E)
- Multi-scale & hierarchical models (e.g., Würstchen)
- Flow matching models
- Control & conditioning: text-to-image, inpainting, ControlNet
- Beyond still images: image editing, personalization, style transfer
- Limitations: compute intensity, training instability, evaluation metrics (FID, CLIP score)



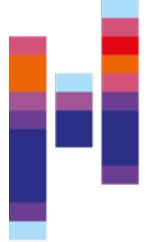
Part 3: Video and Audio Generation

■ 5. Audio Generation

- Autoregressive approaches
- Diffusion models for audio (AudioLDM, MusicLM)
- Music generation (Jukebox, MusicGen)
- Cross-modal links: text-to-speech, speech-to-speech, text-to-music

■ 6. Video Generation

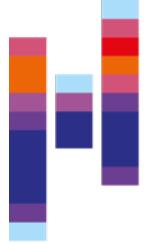
- Extension of diffusion models to temporal dimension (e.g., Imagen Video, Pika, Runway Gen-2)
- Consistency challenges (temporal coherence, motion dynamics)
- Multi-scale spatio-temporal modeling
- Emerging architectures (latent video diffusion, transformers for video)



Part 4: Multimodal Models & Discussion

- 7. Cross-Modal & Multimodal Generation
 - Joint embeddings (CLIP, ALIGN)
 - Multimodal foundation models (Flamingo, Gemini, Kosmos)
 - Agentic AI

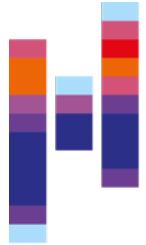
- 8. Challenges, Ethics & Future Directions
 - Training cost, scaling vs. efficiency (distillation, quantization)
 - Evaluation challenges (quality, diversity, alignment with human intent)
 - Risks: deepfakes, bias, misuse, copyright
 - Future: world models, scientific discovery, personalized generative AI



This will be fun.

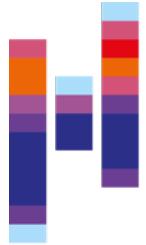
- We will also not forget to have a look at some epic outcomes of generative AI.
- And of course, these lecture slides are full of it.

- So, get strapped - this is going to be quite a ride.



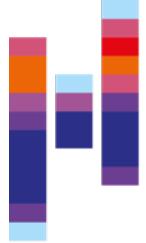
How this lecture works

- The first part of the semester is me giving talks about various topics in the field of GenAI (see slides later).
- Occasionally, I will also bring papers for you to read and present.
- Then, it's your turn:
 - Select a topic of your own choice.
 - Write your own code, do your own experiments
 - Prepare a presentation
 - Finally, submit a summary report of your findings.



Expected time investment

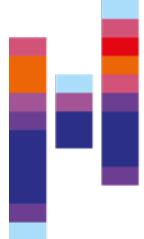
- One ECTS point accounts for 25 wall-clock hours (60 minutes) of effort.
- This lecture has 4 SWS = 6 ECTS
- You can thus expect to plan for an effort of $25 \cdot 6 = 150$ hours for this lecture.
- We have 13 on-site events (lectures), each being 180 minutes, so that's 39 hours.
- The rest of the time is:
 - Preparing and Follow-Up
 - Your own project



Selection of a topic

- I know: Selecting your own topic is hard.
- But on the other hand: You probably don't want me to select the topic for you. ;-)

- You can be totally free in the choice of topic, of course you will find inspiration in the lecture.
- Be aware: The level of complexity you choose influences the grade.
- There is a scheme for the grading (available on StudIP).
- I expect from you, until **November 1st**:
 - The topic of your talk (can be modified if it turns out to be not viable)
 - A rough time estimate



Peer review

- We also want to practice reviewing of reports / papers together.
- Reason:
A critical but well-meaning view on the works of others will make you a better writer.
- Focus:
How could this work have been improved?
What's great about this work?

