ARTIFICIAL INTELLIGENCE AND NEURAL NETWORKS

Lecture 8a – Generative AI: From Prompts to Creative Systems

Chizhi Chris ZHANG

zhangchizhi@ciomp.ac.cn
Advanced Computing and Digital Technology Research Center
University of Chinese Academy of Sciences

Spring 2026

Today's Question

What we are trying to answer

Why did generative AI suddenly become visible in classrooms, offices, design studios, and social media, and what is it really doing when it creates something?

Why this lecture matters

Earlier AI lectures often ended with prediction or classification. This lecture is about a different feeling: the machine now produces text, images, music, and drafts that people can react to directly.

What changes from AI7

AI7 focused on language systems that answer and summarize. AI8 widens the story to content generation more broadly: text, images, editing, control, and the social impact of creative tools.

From AI7 to AI8

Last time

We studied language AI as a system that turns prompts and context into answers.

Today

We step from answering into creating. The system is no longer only describing the world. It is generating new content for the user to inspect, revise, and reuse.

One sentence

AI7 was about language as interaction.
AI8 is about generation as collaboration.

From Recognition to Creation

Earlier AI was often asked to decide

- is this spam
- what object is in the image
- what number should we predict

Generative AI is often asked to make

- write a first draft
- create an illustration
- extend a melody
- turn rough ideas into concrete options

Why It Felt Like a Leap

What changed for ordinary users

AI stopped feeling like a hidden backend component and started feeling like a visible partner that can respond, revise, and offer alternatives.

Why people noticed it so quickly

The output is easy to judge with your own eyes. You do not need to trust a dashboard. You can read the paragraph, look at the poster, or compare two generated versions side by side.

That change in user experience is a big part of why generative AI spread so fast in public discussion.

Where People Meet Generative AI

Study

- explain hard concepts
- make practice questions
- summarize long readings

Work

- draft emails
- outline slides
- create mockups

Creative life

- posters
- short clips
- storyboards

What People Often Get Wrong

What people often say

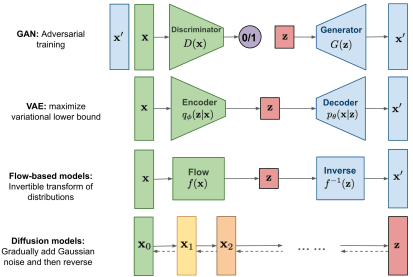
"Generative AI is just copying from the internet."

Better statement

The model learns statistical patterns from enormous training data, then composes new outputs from those learned patterns. That does not make the outputs automatically original, fair, or safe, but it is also not simple copy-paste.

This distinction matters because debates about creativity, ownership, and risk start from how we think generation works.

Generative AI Is Not One Model



Why this picture belongs early

Students often hear one large label, “generative AI”, and assume every system works the same way. In reality, there are several families with different strengths, training stories, and tradeoffs.

A Simple Pipeline

User side

prompt or condition → model → draft output

System side

representation → generation process → decoding or rendering → final result

Why this matters

The user sees one step. The system usually runs through several hidden stages before anything appears on screen.

How Text and Images Become Inputs

Text path

sentence → tokens → vectors

Image path

image → patches or latent codes → vectors

Shared idea

Different media types can still be turned into machine-friendly numerical representations. Once that happens, a model can compare, transform, and generate them.

A Prompt Is Really Task Design

Weak prompt

"Make me a picture."

Stronger prompt

"Create a clean educational illustration of coastal flooding risk for first-year students, with labels, calm colors, and a simple legend."

What improved

The second prompt gives purpose, audience, tone, and constraints. That makes the task easier for the model and makes the result easier for the user to judge.

Conditioning Means Control

Same engine, different job

The model behaves very differently when we add a prompt, a label, a sketch, an edge map, or a reference image.

What users feel

Control turns “surprise me” into “help me move in this direction”.

Why this matters

The moment a system can follow layout, pose, tone, or audience constraints, it starts to feel less like a toy and more like a draft partner.

Why Latent Space Matters

Pixel-space problem

Generating a detailed image directly at full resolution is expensive and noisy.

Latent-space strategy

Compress first, generate in a smaller latent space, then decode back to image space.

Why that was a turning point

It made generation more practical by moving the hard part of the work into a more compact representation.

A Useful Mental Picture

How many people imagine generation

Type one prompt, then a perfect image appears.

How real use usually feels

You start with a rough draft, then guide, revise, crop, regenerate, and compare. Generation is usually an editing loop, not a magic trick.

Why this picture helps

It sets the right expectation: the first output is often only the beginning of the work.

Two Big Generation Stories

Earlier story

One family asks a generator to create and another network to criticize. That was the core idea behind GANs.

Newer story

Another family starts from noise and improves it step by step. That is the intuition behind diffusion systems.

Why ordinary users should care

These two stories led to very different user experiences: one often felt striking but fragile, while the other made steady refinement and control easier.

Why Early Image Generators Felt Fragile

What often went wrong

Early tools could produce impressive images, but they also broke easily: strange artifacts, low consistency, repeated patterns, or unstable behavior from one prompt to the next.

Why this mattered outside research

People do not only want one lucky output. They want a tool they can trust across many attempts.

What the field kept chasing

Better stability, better control, and better consistency across iterations.

Diffusion Intuition

Simple story

Add noise in a controlled way, then train a model to reverse that process step by step.

Why this felt different

The model no longer has to invent the final image in one leap. It can improve a rough guess gradually.

Why diffusion became popular

It offered a path to high-quality generation with more stable training than many GAN setups.

Why Newer Image Tools Felt Better

What users experienced

Results became cleaner, more detailed, and easier to steer. More importantly, people could try several versions and still feel the system was behaving in a somewhat understandable way.

Why this happened

Step-by-step generation, stronger conditioning, and better internal representations all helped the tools feel less random.

What changed socially

Once quality and control improved together, generative AI moved from research curiosity to everyday software.

Why Speed and Cost Changed Behavior

When generation is slow and expensive

People ask for one output, wait, and hesitate to experiment.

When generation gets cheaper

People compare options, iterate more, and treat the tool as part of a workflow instead of a one-shot novelty.

Why this matters

Many big changes in AI adoption come not only from better models, but from better speed, cost, and usability.

Where Generative AI Actually Helps

Content

- first drafts
- rewrites
- image variations

Productivity

- faster ideation
- faster prototyping
- faster comparison

Education

- visual explanations
- practice materials
- alternative examples

Why Control Matters

Why users ask for more than pretty randomness

In real projects, people need pose, layout, edge structure, style, tone, and content constraints. Control turns generation from novelty into something teams can actually iterate on.

A realistic request

“Keep the same composition, leave space for the title, make the mood calmer, and give me three versions I can compare.” That is much closer to real design work than “make something cool”.

Before control

You may get individually interesting outputs, but consistency across revisions is weak.

After control

Layout and structure become part of the workflow, so revision feels more like editing and less like gambling.

A Museum Poster Story

The task

A museum team needs a poster for a student exhibition. They want three different moods, readable title placement, and a version that still looks good on mobile.

Where AI helps

It can quickly produce alternative directions, color palettes, rough compositions, and draft visuals that the team can compare.

Where people still matter

People still decide the message, choose the final style, check rights, fix errors, and ensure the design actually fits the audience.

A Better Workflow

Bad workflow

Ask for one output, trust it immediately, and publish it without checking.

Better workflow

Generate options, compare them, revise the prompt or condition, edit the result, then verify facts, rights, and fit before release.

Why this is the classroom habit to keep

Generative AI is strongest as a tool for exploration and iteration, not as a machine that removes the need for human review.

Why This Connects to Earlier Lectures

The course thread so far

We started with classification, search, language, sequence modeling, and transformers.

What AI8 adds

Generation pulls those strands together. It uses representations, sequence models, optimization, and user interaction to create outputs rather than only score inputs.

That is why this topic feels like a summary point for much of the course so far.

Plausible Is Not the Same as Correct

Reality

A generated image, caption, or paragraph can look polished and still be misleading, inaccurate, or inconsistent with the user's actual goal.

Classroom rule

The better the output looks, the more carefully you should check it when the stakes are high.

Why this is easy to forget

Humans are strongly influenced by fluency, visual polish, and confidence.

Copyright and Ownership Questions

What students should ask

- What data trained the model?
- Does the platform allow this use?
- Can this output be published commercially?
- Who is responsible if a dispute appears later?

Practical rule

Rights, attribution, and policy matter before publication, not after a complaint arrives.

Deepfakes Change Trust

The social issue

When realistic generated media becomes cheap, people can be misled not only by fake content, but also by uncertainty about whether any content is real.

Why this matters beyond engineering

This affects journalism, politics, law, education, and everyday interpersonal trust.

The real question

How do we preserve trust when synthetic media becomes easier to make than careful verification?

A Simple Safety Checklist

Before

Check the task, the audience, and the allowed use.

During

Watch for drift, bias, strange artifacts, and hidden assumptions in the prompt.

After

Verify facts, rights, privacy, and whether a human should revise the final result.

One Failure Pattern to Remember

Bad workflow

Use the first attractive result because it is fast.

Better workflow

Treat outputs as drafts, compare several versions, and keep a human decision point before anything public or important.

This single habit prevents many avoidable mistakes.

Why This Field Keeps Moving

What keeps changing

New systems appear whenever researchers find a better balance among quality, control, speed, cost, and ease of training.

Why students should expect fast change

Generative AI is not one settled technique. It is a moving stack of models, interfaces, and workflows.

The stable lesson

Even if the model family changes, the user still has to ask the same questions about fit, trust, rights, and review.

Why AI8 Leads to NN8

What AI8 focused on

Why generative AI became visible, where people use it, why prompts and control matter, and why risk grows with realism.

What NN8 will focus on

The neural mechanics behind the main model families: GANs, diffusion, conditioning, evaluation, and engineering tradeoffs.

The application story is now in place. The next step is the model story.

Summary

- Generative AI felt different because it moved AI from hidden prediction into visible collaboration.
- Prompts and conditions matter because generation becomes useful only when users can guide it.
- GANs introduced a powerful competitive training story, while diffusion later offered a more stable step-by-step generation story.
- Control and workflow matter as much as raw model power in real use.
- Rights, trust, and verification become more important as generated content becomes more realistic and easier to share.

Next Time

Where the course is going

We now move from the user-facing story of generative AI to the neural mechanisms that make the major generative model families work.

Next lecture

We will study GAN objectives, diffusion noising and denoising, conditioning channels, evaluation, and the practical tradeoffs behind modern image generation systems.

Thank You