

ACADEMIA AND THE THIRD WAVE OF ONLINE CONTENT

Campus-Wide Faculty
Event, August 2023:

“The Experience of
Faculty in the Age of
Artificial Intelligence”

RANDALL HARP



The University of Vermont

Welcome (back)!

Three functions of the University:

- **Knowledge Creation**
- **Knowledge Dissemination**
- **Knowledge Certification**


Generative AI will affect all three of these!



Three Waves of Online Content

First wave: **Access to information** is improved

Search engines like Yahoo! Search, Magellan, and Google make it possible to easily find information that has been posted online



Three Waves of Online Content

Second wave: **Sharing of information** is improved

The rise of **social media networks** like Myspace, Facebook, Twitter, and Reddit make it easier to share information (both in a one-to-one and in a one-to-many way) with other users of the internet

Three Waves of Online Content

Third wave: **Creation of information** is improved

Generative AI tools like ChatGPT, Bard, Dall-E, and Midjourney make it easier for any individual on the internet to create plausible-seeming information on any topic and in a wide range of modalities, independent of whether the information is true or corresponds to reality

-
-
- The first two waves were in response to genuine needs, but also introduced **serious problems and ethical challenges** that have yet to be resolved
 - It's less clear what need the third wave is a response to, but it's more clear what sorts of problems and ethical challenges the third wave might introduce
 - **All three core academic functions of the university will need to respond to these challenges**

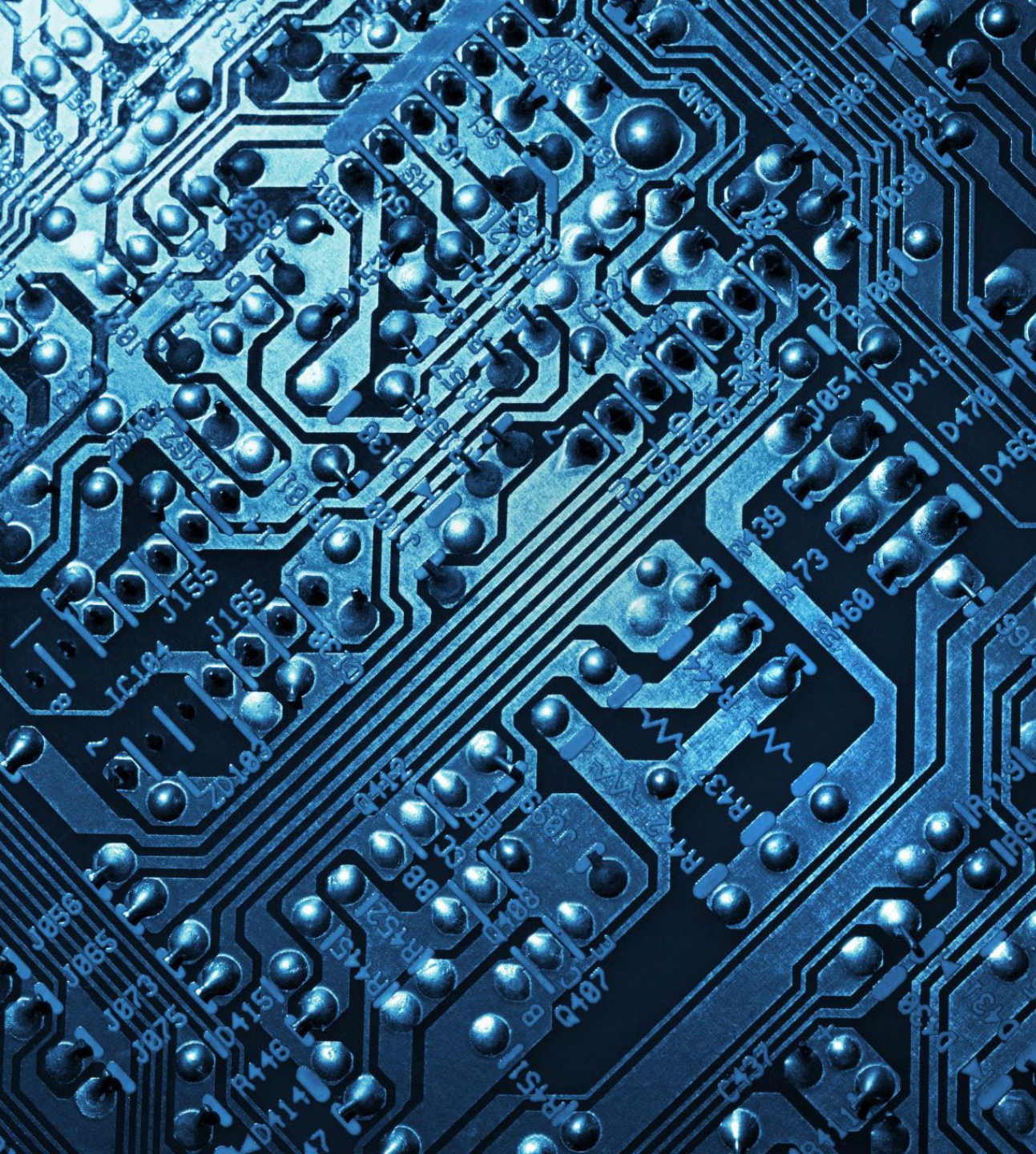
Outline of the talk

- **What is AI?**
 - What is different about generative AI (genAI)?
- **How does genAI affect knowledge transmission (learning)?**
 - How do students actually use genAI? Is there any useful way for genAI to be incorporated into teaching and learning activities in a classroom?
- **How does genAI affect knowledge certification (grading and evaluating)?**
 - Is there any way to assign work in the age of chatGPT?
- **How does genAI affect knowledge production (faculty research)?**

What is AI?



The University of Vermont



- Lots of things get called AI; the term itself is not very useful
- For our purposes, the following account of AI will suffice:

AI is an **algorithmic process**, implemented in a machine or computer, that **connects inputs with outputs**

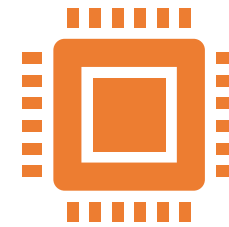




The **inputs** are anything that involve **data**



Data are just measurements or observations (objective or subjective) about the world



If something can be measured, then it can serve as the input to an AI algorithm

-
- **Outputs** can be:
 - **Classifications** (“this is a cat”; “this is a grammatical sentence in English”)
 - **Actions** or **instructions** (“Drive straight”)
 - **Judgments** (“Persons with these properties are 80% likely to recidivate within 5 years”; “these actions are likely to lead to the betterment of society”; “end users are more likely to post a response on social media to content that is inflammatory”)
 - **Etc.**
 - These outputs can take the form of text, audio, visual, actions, propositional judgments, etc.

How does AI do this?

(This is an overly simplistic story!)

Most AI involves some kind of **machine learning (ML)** in order to optimize the algorithm that goes from input to output

There are lots of different ways of doing machine learning; we can take **supervised learning** to be a representative method.

The general idea of supervised learning:

- You have a **training data set** (the inputs) which are **labeled**; the labels are human judgments of how the data should be classified, or what actions should be taken in response to the data, or what propositional judgments the data license, etc.
- An algorithm goes through the training data set and tries to label them. For every error, the algorithm goes back and tries to fix itself (this is just math—details not important); it keeps doing so until it is able to successfully label most of the training data.
- The algorithm is then introduced to a novel data set of inputs (the **validation data set**). If the algorithm produces desired outputs in response to this novel data set, the algorithm can be said to have “learned” that particular task

Standard ML cases

- Text recognition (handwriting, etc.)
- Risk prediction (person X is likely to default on a loan; person Y is likely to recidivate)
- Facial recognition
- Spam filtering (this message is spam; this message is a legitimate email)
- Recommendation algorithms (people who buy these things are also likely to buy this thing; people who watch these things are likely to watch this thing; people who favorably rate these things are likely to favorably rate this thing, etc.)

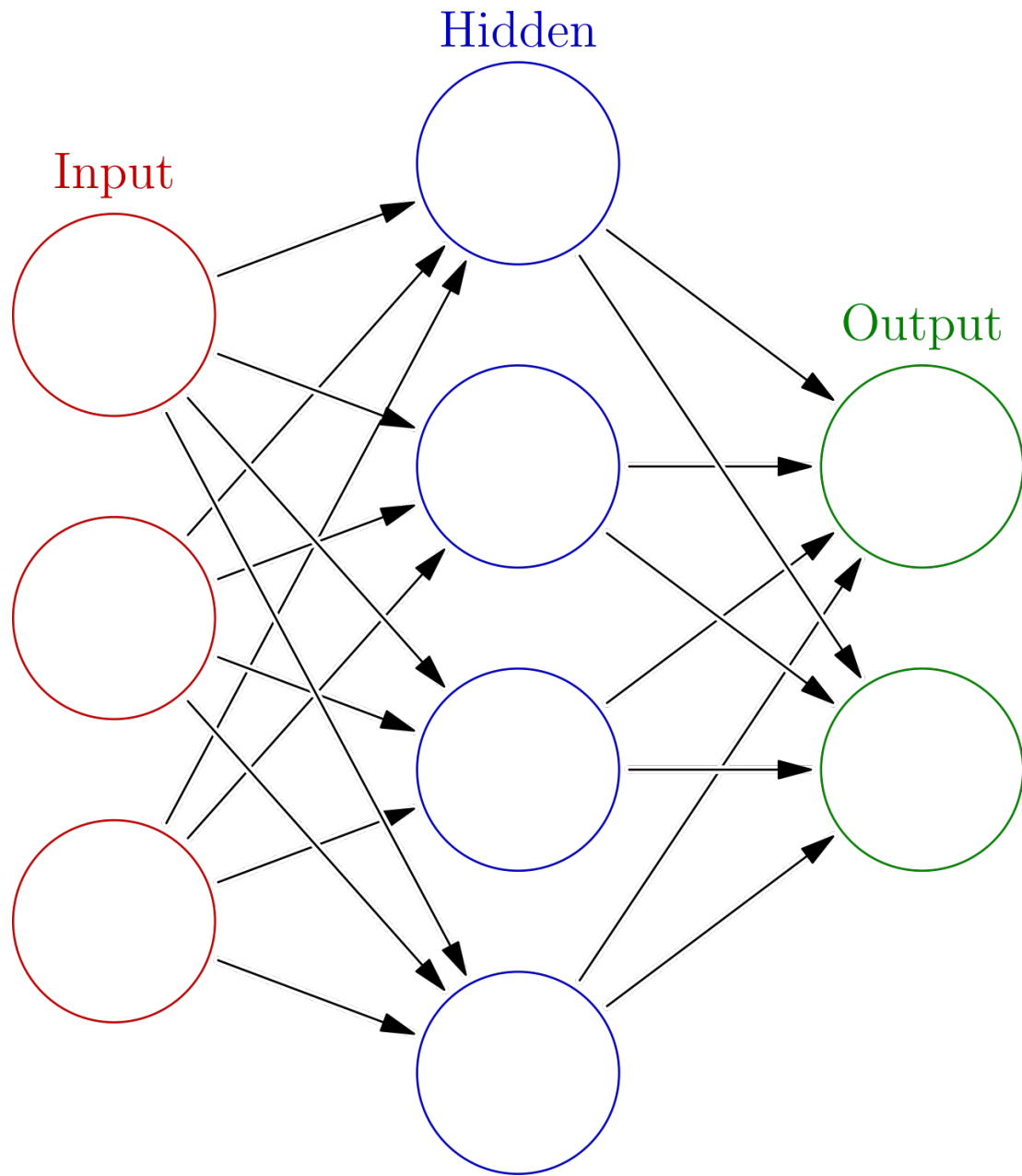
A quick pause

You might have been thinking to yourself “But wait! This sounds problematic!”

It is! (Even skipping over the math.)

A lot of work is being done by human judgments about the appropriate connections between inputs and outputs, and whether those judgments are malleable or fallible

Also, we don't know exactly why, or in virtue of what, the machine algorithm has decided to correlate inputs with outputs; the algorithms are not always **transparent**. So, even if the algorithm happens to give correct results to one validation data set, it might fail on others, or it might arrive at a “good” outcome through unjust or unfair means



By glosser.ca, Wikipedia

Deep Learning

- “**Deep learning**” is just a kind of machine learning that characteristically uses **neural networks** in order to make more optimal connections between inputs and outputs than other ML algorithms
- What is a neural network? Math. It’s just more math. (It’s not Frankensteinian!)
- In a neural network, we have **nodes** with some **strength**, and those nodes have connections (also of some strength) with other nodes; the nodes are arranged in layers, so that nodes of one layer only receive inputs from the earlier layer and only output to the later layer

In a neural net, the output is calculated (with math) based on all those weights and connections. If the output does not line up with the expected output when applied to training data, then the weights and connections are adjusted and the algorithm tries again

(This is “deep” learning because there are lots of layers; the layers run deep)

So deep learning is one way of doing ML

And ML is one way of doing AI

And AI is just a way of finding “good” ways of going from inputs to outputs

(And yes, there’s a lot buried in that word “good”)

In sum: AI is just a very fancy, very sophisticated pattern matching machine—and with deep learning techniques, and with increasingly powerful computational capacities, AI has gotten very, very good at matching patterns in general.

Generative AI

Generative AI (genAI) is a kind of AI, and so the basic idea is the same: a genAI model is rewarded when it receives an input from training data and produces an output which is the same as the “good” output

In the genAI case, though, the input is a prompt or the first part of a phrase, and the output is the response or the rest of the phrase

The genAI initially produces random responses, but after many many many (many!) iterations, the system’s weights and connections have been adjusted so that it produces plausible-seeming responses to inputs

The precise training data for genAI is unknown, but it is colloquially referred to as “the entire internet”; they are called **Large Language Models (LLMs)** in part because of the huge number of words (tokens) in their training data set

So, for any prompt, genAI attempts to produce an output which is the most similar to the expected response that would be given by “the entire internet”

(Note: the same process works for image and video and audio generation, but I’m going to skip over the details)

I like to think of genAI with the “Google Test”:

1. Take a prompt (yes you, human)
2. Type that prompt into Google
3. Read everything that is posted on the first fifteen or so links that come up in response to the prompt
4. Use your smart human brain to understand everything in those first fifteen or so pages, and **use that understanding to answer the prompt**

If you think that you can come up with a decent response to that particular prompt, then genAI probably can as well, because the information is on the internet. (Lots of stuff is on the internet!)

Note also: genAI does not produce all at once a text which is the best output for a given input; rather, genAI produces a response word by word (or, more precisely, token by token)

This is part of the reason why genAI **hallucinates** (i.e., gives an output which asserts that some event happened in reality even though the event did not happen); genAI is just choosing the statistically most likely word which comes next as an output to a given inputted prompt, one at a time—but all those individually statistically likely words can add up to an output which is unlikely in its collectivity.

In sum: genAI is a very fancy autocomplete; if you give it a prompt as input it will autocomplete that prompt in giving the most likely output, one word (token) at a time

But it is a very very very good autocomplete

And most students know this

How does genAI affect knowledge transmission?



How do (or could) students use genAI?

- Some use genAI directly to cheat—they get an output and turn it in as their own
- Some use genAI as natural language search-and-explain engines (“What wine pairs well with ramen noodles?” “Who wrote about the concept of death in neanderthal culture?”)
- Some use genAI for supplemental research and writing purposes (“How can I make this paragraph clearer?”)

I don't know how many students currently use genAI regularly or semi-regularly in some way

Some students certainly do use genAI very broadly

(Consider the aptly-titled article in Vice by Claire Woodstock, [“Students Are Using AI to Write Their Papers, Because Of Course They Are”](#))

How should students use genAI?

More precisely, how can genAI enhance the learning that should be happening in a classroom environment?

What do we want to do in a classroom?

- **Facilitate the transfer of knowledge**
and
- **improve critical thinking**

Facilitate knowledge transfer

A requirement for students **knowing** is that they can say not just **what is the case**, but (to some degree) **why it is the case**, or perhaps (to some degree) **why something that is not the case is not the case**

LLMs can help with this

Consider a suggestion by Ethan and Lilach Mollick ([“New Modes of Learning Enabled by AI Chatbots: Three Methods and Assignments”](#)):

Teach a concept. Give students clear guidance or a clear rubric about how to evaluate an explanation of that concept. Then have students ask a genAI chatbot (ChatGPT, etc.) to explain the concept in multiple ways. Have the students evaluate the chatbot responses.

(A bonus is that it helps students realize that genAI chatbots do not always give the best responses)

Improve critical thinking

Part of critical thinking involves being able to see the ways in which **conclusions** follow from **premises** via **reasons**

Current genAI is not very good at using specific bits of evidence to argue for specific conclusion, because genAI is not great about knowing specific bits of evidence. GenAI is also (sometimes, at least) not great about arguing two sides of the same position.

Consider the following prompt:

“Cite three passages in Pride and Prejudice that support the claim that Elizabeth Bennett is a shallow character who only values marriage and wealth and cite three passages that argue against the claim. Be sure to explain why the passages support or argue against the claim.”

Not a great prompt, to be clear! But consider the first paragraph of the response:

“It's important to clarify that the claim that Elizabeth Bennet is a shallow character who only values marriage and wealth is not accurate based on the character as depicted in Jane Austen's "Pride and Prejudice." Elizabeth is a complex and multidimensional character with a depth of personality and values that go beyond mere material concerns. However, I can provide you with passages that may help illustrate her values and personality more accurately.”

This might be correct! It is also obviously not what was asked for.

And consider the first bit of evidence cited in **support** of the claim that Elizabeth Bennett is shallow:

“Passage 1 (Chapter 6):

“Mr. Collins was not a sensible man, and the deficiency of nature had been but little assisted by education or society; the greatest part of his life having been spent under the guidance of an illiterate and miserly father; and though he belonged to one of the universities, he had merely kept the necessary terms, without forming at it any useful acquaintance. The subjection in which his father had brought him up, had given him originally great humility of manner; but it was now a good deal counteracted by the self-conceit of a weak head, living in retirement, and the consequential feelings of early and unexpected prosperity.”

In this passage, Elizabeth's assessment of Mr. Collins reflects her refusal to marry for the sake of financial security, despite the societal pressure to do so. She values intellectual and emotional compatibility over wealth and status.”

This passage is in Chapter 15, not Chapter 6, and it does not actually speak to the matter at hand. This is true of all of the quotes that ChatGPT cited: most actually exist (though some were completely made up), but each of the ones that were genuine quotes were attributed to incorrect chapters. Further, none of the cited passages were clearly support for or against the claim, and most were clear and obvious failures

Again, this can be a useful exercise to show students how **not** to cite evidence or engage in critical reasoning (while at the same time warning them that genAI is not great at some of these tasks).

In sum: It is sometimes said that professors don't really learn a subject until they have to teach it. Part of the reason is that when you teach it, you have to see many examples of wrong or incomplete or infelicitous ways of thinking about or explaining the subject. GenAI can give students practice seeing multiple ways of explaining or thinking about a subject, which can facilitate knowledge acquisition, and genAI can show students wrong ways to cite evidence, which can help them cite evidence better.


—

How does genAI affect knowledge certification (grading and evaluating)?

—



The University of Vermont



"Ok, but Randall: I don't feel like finding ways to build discussions of ChatGPT into my classroom pedagogy—I just don't want my students to (be able to) cheat!"

I get it.



We want students to develop critical thinking skills and use learning materials properly. We have already seen that current genAI cannot do these fluently.

Part of what we should be thinking about in our grading practice is making students show these skills more explicitly.

- **If your paper-writing and exam-writing exercises allow**, ask students to explicitly support their claims with specific quotes
- Ask students to choose specific passages and explain what is going on in that passage (All the better if the passage is not a super-famous passage that has been written about extensively on the internet—remember the Google Test)
- Ask students to **show their work!**
 - Have students give you **outlines** of their essays before they write them
 - Have students give you **1-page abstracts** of their essays
 - Have students rewrite, and **explain their rewrite decisions**

-
-
- Ask each student to spend 5 minutes in your office hours giving a “pitch” of their paper or paper argument, and then have them write up their pitch

(Note: **not all of these “show your work” exercises need to be graded!** You can register that they did them, and then refer back to them if you have later suspicions that students have made improper use of online resources)

In sum: we are not teaching strictly for an end product, we are teaching a process. Because the process for genAI is often opaque, it will always be difficult to get genAI to describe the process of something it has written. The more we evaluate our students’ process, the more we genAI-proof their work.

In summier sum: **evaluate the process!**



How does genAI affect knowledge production (faculty research)?



Remember the general principle: **don't ask genAI any question you don't know (or couldn't easily find out) the answer to!**

The general principle will already limit the role of genAI in knowledge production, at least at this point.

Also, how you use genAI will be very discipline-specific, so I don't want to speak in non-useful generalities

An attempt at a **useful** generality:

If there are tasks for which proofreading and correcting the output of genAI is faster than generating the content yourself from scratch, then you can think about using genAI

“Here's a bunch of paperwork for this grant application!”

Ok, **maybe** properly prompted genAI can help for those parts of the grant application where the difference between answering it at 100% and answering it at 60% is probably not outcome-determinative

Important caveats

Any document or text we enter in a genAI chat box becomes under the control of the genAI company

Maybe I want a genAI chatbot to give a one-paragraph summary of a paper I read and need to write a referee report for. It **might** be a useful use of genAI—except I cannot give that draft paper to a for-profit company so that they may do what they like

Likewise, I cannot upload data that contains confidential human subject research, or medical information, or trade secrets, etc.

Zoom got in trouble for initially saying they reserved the right to use all content from every zoom call as training data for LLMs

They walked back that decision, for obvious reasons, but the point remains:

Companies in this space want as much data as they can have. GenAI companies currently operate under the presumption that anything uploaded to their chatbots is their data—and so we should be very careful what we give them

That's all I've got—

Good luck out there in this brave new world!

(This video was chosen by non-generative AI as a good illustration of the text on this slide.)

(We're going to need the luck.)

