

# When Robots Learn to Write, What Happens to Learning?

Bill Hart-Davidson, Ph.D.



Scan me!

# About Me



Writing, Rhetoric, & American Cultures  
Professor Bill Hart-Davidson

Bill Hart-Davidson, Ph.D. (Purdue, 1999)

Professor in Dept. of Writing, Rhetoric & American Cultures

I study writing, broadly construed, as human activity. I tend to be interested in behavior(s) more than texts. This is because I am fascinated by writing and its centrality to our organizational lives.

If writing gets things done in the world, maybe *better* writing can make things better.

# Topics for Today

## 1. A little about Large Language Models (LLMs) and how they work...or

What the heck just happened to make GPT-3 so much better at drafting longer texts that resemble human-drafted ones? Spoiler: Transformers.

## 2. A little about writing (as humans do it and as robots do it)

Writing is intentional, goal-directed activity that sometimes results in a text. LLMs can simulate some parts of this activity better than others.

## 3. Four Proposals for Writing and Learning with AI in the Loop

How should we be thinking differently about the writing tasks we ask students to do and how we evaluate them?

# This is ChatGPT - GPT in a chatbot

- A chatbot interface connected to the GPT-3 LLM produced by OpenAI
- Trained with human feedback to be responsive to dialogic queries
- Capable of producing texts that are similar in content, tone, and formatting to those that a human would make.



Can you give me a recipe for nachos?



Sure, here's a basic recipe for nachos:

Ingredients:

- 1 bag of tortilla chips
- 1 cup of shredded cheese (cheddar or Mexican blend)
- 1 can of black or pinto beans, rinsed and drained
- 1 diced tomato
- 1 diced avocado
- 1 diced jalapeno (optional)
- 1 diced red onion (optional)
- 1 diced bell pepper (optional)
- Sour cream and salsa for serving (optional)

Instructions:

1. Preheat the oven to 375°F (190°C).
2. Spread the tortilla chips on a baking sheet.
3. Sprinkle the shredded cheese, beans, diced tomato, avocado and bell pepper over the chips.
4. Bake in the preheated oven for 5-7 minutes, or until the cheese is melted.
5. Serve with sour cream and salsa, if desired.

Enjoy!

No Meat

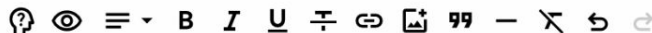


# How Does ChatGPT Know What to Write? LLM basics...

- The LLM, GPT-3, has been trained on a very large collection of texts. Hundreds of billions of words. After processing, these are called “tokens” in LLM parlance.
- The “T” in GPT stands for “transformer.” The model converts the words to tokens and the tokens to a very large graph that allows for more efficient computation than just working on a big string of words does.
- GPT-3 is known as an autoregressive model. Based on one string of tokens, it predicts what comes next using probabilities derived from its training corpus.
- GPT-3 is a “zero-shot” classifier, which means it does not need any examples of the thing it is trying to recreate. This contrasts with one-shot or few-shot models.
- ChatGPT has other layers of training too - one from feedback that evaluates responses and is used to refine future ones. This makes it a “deep learning” application.

# This is Lex - GPT3 in a Word Processor

LEX



No  
Formatting

## Nachos

*Still No Meat*

Ingredients: 2-3 cups of nacho chips - 2 cups of shredded cheese (15 ounces) of refried beans - ½ cup of salsa - ½ cup of sour cream  
Optional Toppings: jalapenos, black olives, diced tomatoes, green onions  
Instructions: 1. Preheat oven to 350 degrees F. 2. Spread out chips in a greased baking pan or dish. 3. Top chips with shredded cheese and refried beans. 4. Bake chips in preheated oven for 15 minutes. 5. Remove pan from oven and top with salsa, sour cream, and optional toppings. 6. Serve warm

Clickbaity  
Titles



1. Delicious Nacho Recipe to Satisfy Your Cravings!
2. The Perfect Plate of Nachos Every Time!
3. Unleash Your Inner Chef with this Easy Nacho Recipe
4. Nacho Mania - Ready in 15 Minutes!
5. Spicy and Cheesy - The Ultimate Nacho Experience!

# Transformer - More than Meets the Eye!

“An attention function can be described as mapping a query and a set of key-value pairs to an output, where *the query, keys, values, and output are all vectors*. The output is computed as a weighted sum of the values, where the weight assigned to each value is computed by a compatibility function of the query with the corresponding key.”

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.

## Attention Is All You Need

Ashish Vaswani\*  
Google Brain  
avaswani@google.com

Noam Shazeer\*  
Google Brain  
noam@google.com

Niki Parmar\*  
Google Research  
nikip@google.com

Jakob Uszkoreit\*  
Google Research  
usz@google.com

Llion Jones\*  
Google Research  
llion@google.com

Aidan N. Gomez\* †  
University of Toronto  
aidan@cs.toronto.edu

Łukasz Kaiser\*  
Google Brain  
lukaszkaizer@google.com

Illia Polosukhin\* ‡  
illia.polosukhin@gmail.com

### Abstract

The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train. Our model achieves 28.4 BLEU on the WMT 2014 English-to-German translation task, improving over the existing best results, including ensembles, by over 2 BLEU. On the WMT 2014 English-to-French translation task, our model establishes a new single-model state-of-the-art BLEU score of 41.8 after training for 3.5 days on eight GPUs, a small fraction of the training costs of the

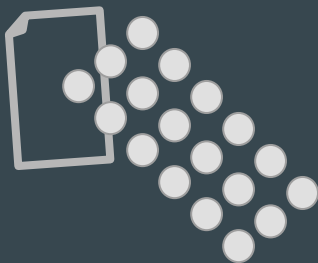
# Transformer...more than meets the eye

## Training Corpus



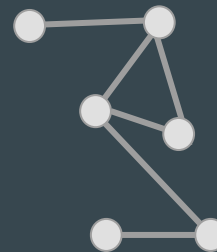
What is in the corpus and how has it been evaluated?

## “Tokens”



What steps were taken to process the words in the corpus?

## Graph



How has the model been built to create “embeddings” ?





# Hedges and the Hedge-o-Matic

## Hedge signals added

Greenhouse gas emissions from human activities **may** continue to affect Earth's climate for decades and even centuries. Humans are **likely** adding carbon dioxide to the atmosphere at a rate far greater than it is removed by natural processes, creating a long-lived reservoir of the gas in the atmosphere and oceans that is driving the climate to a warmer and warmer state.

## Actual Text

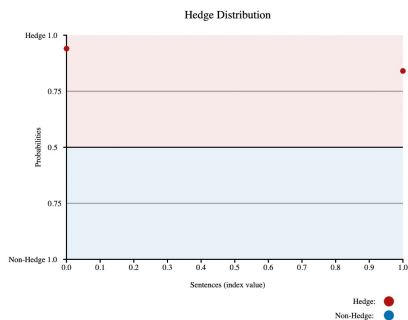
Greenhouse gas emissions from human activities continue to affect Earth's climate for decades and even centuries. Humans are adding carbon dioxide to the atmosphere at a rate far greater than it is removed by natural processes, creating a long-lived reservoir of the gas in the atmosphere and oceans that is driving the climate to a warmer and warmer state.

Source: <https://www.globalchange.gov/climate-change/whats-happening-why>

# The Hedge-o-Matic reliably identifies propositional hedges

## Hedge signals added

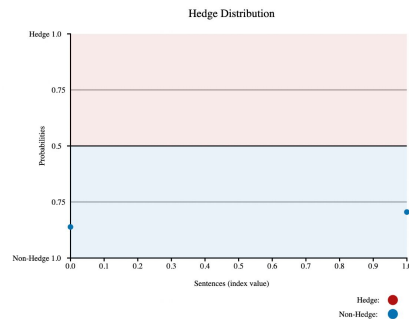
- Sentence Count: 2
- Non-hedges: 0%
- Hedges: 100%



Index	Sentence	Hedge/Non-Hedge	Probability
0	Greenhouse gas emissions from human activities may continue to affect Earths climate for decades and even centuries.	hedge	0.940283823552
1	Humans are likely adding carbon dioxide to the atmosphere at a rate far greater than it is removed by natural processes, creating a long-lived reservoir of the gas in the atmosphere and oceans that is driving the climate to a warmer and warmer state.	hedge	0.840626456741

## Actual Text

- Sentence Count: 2
- Non-hedges: 100%
- Hedges: 0%



Index	Sentence	Hedge/Non-Hedge	Probability
0	Greenhouse gas emissions from human activities will continue to affect Earths climate for decades and even centuries.	non-hedge	0.861412148583
1	Humans are adding carbon dioxide to the atmosphere at a rate far greater than it is removed by natural processes, creating a long-lived reservoir of the gas in the atmosphere and oceans that is driving the climate to a warmer and warmer state.	non-hedge	0.794917371293

Source: <https://www.globalchange.gov/climate-change/whats-happening-why>

HoM: <http://hedgeomatic.cal.msu.edu/hedgeomatic/>

# Topics for Today

1. A little about Large Language Models (LLMs) and how they work...or

What the heck just happened to make GPT-3 so much better at drafting longer texts that resemble human-drafted ones? Spoiler: Transformers.

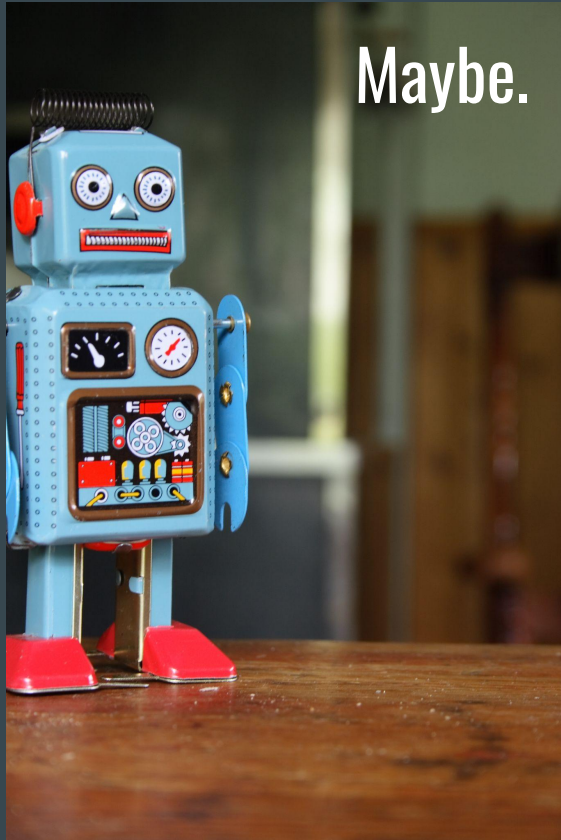
2. A little about writing (as humans do it and as robots do it)

Writing is intentional, goal-directed activity that sometimes results in a text. LLMs can simulate some parts of this activity better than others.

3. Four Proposals for Writing and Learning with AI in the Loop

How should we be thinking differently about the writing tasks we ask students to do and how we evaluate them?

# Are (writing) Robots a Threat?



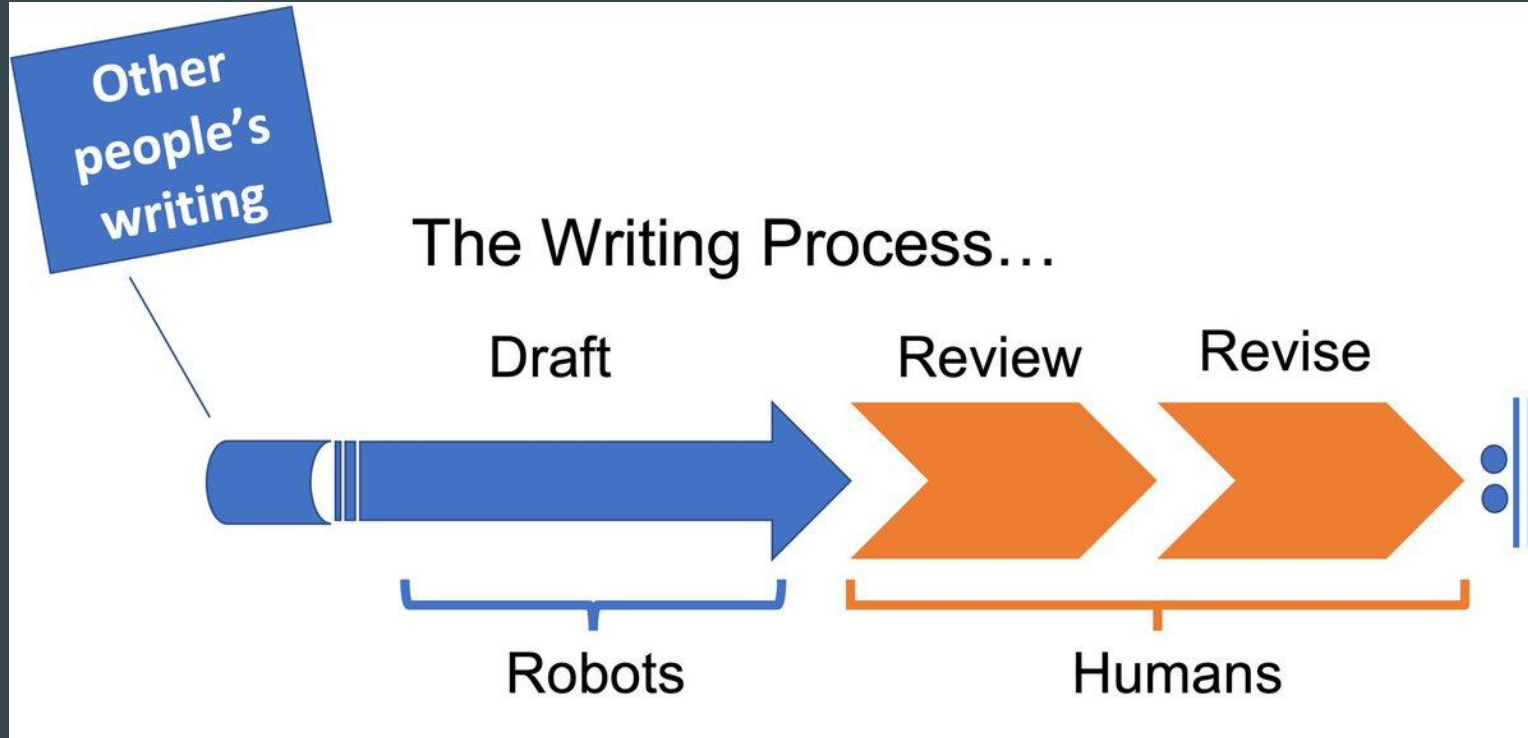
Maybe.

- Writing activity is important as **social behavior**. When robots make texts, it simulates some of this behavior. And that might cause problems...e.g. Grant proposals.

But the bigger issue for me is...

- **Writing is good practice**. Robots can do part of that practice now really easily. It may mean that humans will miss practice and miss some learning as a result.

# You are Here... or “New Writing Process Just Dropped!”



# Writing is Intentional Human Action, Not Just Text Making

- Writing is *behavior*. That behavior sometimes results in a text. Sometimes those texts are shared and sometimes they are not. But...
- People use writing to do things OTHER than make texts. Writing is usually a means, not the end.
- The situations where making a text is the GOAL are the exception, not the rule. Like poetry.
- In the words of writing studies researcher Anthony Pare - “We don’t write writing.”

# Topics for Today

1. A little about Large Language Models (LLMs) and how they work...or

What the heck just happened to make GPT-3 so much better at drafting longer texts that resemble human-drafted ones? Spoiler: Transformers.

2. A little about writing (as humans do it and as robots do it)

Writing is intentional, goal-directed activity that sometimes results in a text. LLMs can simulate some parts of this activity better than others.

3. **Four Proposals for Writing and Learning w/AI in the Loop**

How should we be thinking differently about the writing tasks we ask students to do and how we evaluate them?



# Writing and Learning - Four Proposals

1. Whenever we ask students to write, we should consider that as asking them to practice something. Writing is not a transparent window on thinking. It's a social activity that has benefits for learning. But it takes practice to realize those benefits.
2. Students need more deliberate practice in other parts of the writing process - not just drafting. Criterion-referenced review and revision are especially important.
3. Show your work. That should be our new (old?) mantra. We need to see the practice, and students need feedback on it.
4. Where LLMS, or AI are part of our workflow, we should have a **consent and disclose** approach.
  - a. Consent should precede the use of LLMS in many writing situations, including teaching and learning, as well as academic publishing
  - b. Disclosure should be(come) a regular part of our practice if we use LLMS, for example in the method sections of a research article.

# Consent and Disclosure

Where LLMs are part of our work, we should...

- 1) Make sure that we are transparent about when and how people are asked to contribute their work to systems that will reuse it for training purposes
- 2) Make sure that we understand when and how they were used to enhance the writing process and take not to be deceptive about those
- 3) Develop conventions for disclosing the use of LLMs as we do with other computer-enhanced tools, e.g. in methods sections, etc.
- 4) Develop ground rules for letting folks know what is and is not out of bounds for using LLMs in classroom (i.e. practice) situations - e.g. “I want you to do this drafting by hand...” etc.

**Thank you!**

# When Robots Learn to Write, What Happens to Learning?

...

Bill Hart-Davidson, Ph.D.



Scan me!

# References & Further Reading

Aristotle. (2022). *Art of Rhetoric*. J.H. Freese, trans. Loeb Classical Library. Harvard UP.

Erdős, P., & Rényi, A. (1961). On the strength of connectedness of a random graph. *Acta Mathematica Hungarica*, 12(1), 261-267.

Gunnarsson, B. L. (1997). The writing process from a sociolinguistic viewpoint. *Written communication*, 14(2), 139-188.

Hart-Davidson, W. (2003, October). Seeing the project: Mapping patterns of intra-team communication events. In *Proceedings of the 21st annual international conference on Documentation* (pp. 28-34).

Hart-Davidson, W., Spinuzzi, C., & Zachry, M. (2007, October). Capturing & visualizing knowledge work: Results & implications of a pilot study of proposal writing activity. In *Proceedings of the 25th annual ACM international conference on Design of communication* (pp. 113-119).

Hart-Davidson, W., & Omizo, R. (2017). Genre signals in textual topologies. In *Topologies as techniques for a post-critical rhetoric* (pp. 99-123). Palgrave Macmillan, Cham.

Kaufer, D., Geisler, C., Vlachos, P., & Ishizaki, S. (2006). Mining textual knowledge for writing education and research: The DocuScope project. *Writing and digital media*, 115-129.

Omizo, R., & Hart-Davidson, W. (2016). Hedge-O-Matic. *enculturation*, 7. Omizo, R., & Hart-Davidson, W. (2016). Finding genre signals in academic writing. *Journal of Writing Research*, 7(3), 485-509.

# References, cont.

Omizo, R. M., & Hart-Davidson, W. (2017, August). Digging text viz: an archaeological review of ACM digital library text visualizations publications (1991--2003). In *Proceedings of the 35th ACM International Conference on the Design of Communication* (pp. 1-13).

Omizo, R., Hart-Davidson, W., Nguyen, M. T., Clark, I., McDuffie, K., & Ridolfo, J. (2016). You Can Read the Comments Section Again: The Faciloscope App and Automated Rhetorical Analysis. *DH Commons Journal*.

Omizo, R. M. (2019). Participation and the Problem of Measurement. In *The rhetoric of participation: interrogating commonplaces in and beyond the classroom*. Computers and Composition Digital Press/Utah State University Press Logan, UT.

Omizo, R. M. (2020). Machining topoi: Tracking premising in online discussion forums with automated rhetorical move analysis. *Computers and Composition*, 57, 102578.

Omizo, R. (2022). Reprogramming the Faciloscope. *Reprogrammable Rhetoric: Critical Making Theories and Methods in Rhetoric and Composition*, 108.

Pare, A. Pare, A. (2009). Writing Matters: Back to the Future with Rhetoric. *Education Canada*, 49(4), 4-8.

Plato. (1952). *Plato's Phaedrus*. Cambridge :University Press.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.