NEWS

"You're gonna need a bigger boat."
JAWS BOBBY PICKETT

As class formals near, students keep eyes out for blind dates

BY OLIVIA CYRUS AND ANAYAH ACCILIEN
STAFF REPORTERS

Shogo O'Connor '29 says that during the fall of his freshman year, he had a "really big crush on this one girl."

O'Connor knew that he wanted to get to know her more, but he was repeatedly asking himself, "How?" Fortunately for him, a friend of his, who is also an acquaintance of his crush, offered to set the two of them up for the First Year Formal.

Every year, the First Year Formal spurs a wave of first years yearning to be set up by their peers on blind dates for the special night. Yet, as first years returned to campus for the spring semester, with the formal approaching in February, many described a feeling of pressure to pair up with a date. Those who had not yet been paired lamented that they had not found a date.

Similarly, next weekend's Sophomore Formal organizers are attempting to formalize matchmaking culture by using a randomized form to pair up sophomores for the occasion.

"There is nothing to lose," O'Connor said. "There's so many people on campus that it can sometimes feel hard to approach them on your own. Let's say you see someone you think is really attractive. You're not going to run to them naturally. It might not happen because there's so many people. Matchmaking culture allows you to close the gap."

First-year Cheizyn Montizor '29 said that "everyone's talking about it" and is asking their friends for "cute" peers to go formal with.

Eden Hu '29, a friend of Montizor, agreed and said that the desire to find a friend-suggested pairing for the night of Feb. 20 has even begun to show up in pleading posts on the anonymous social forum Fizz.

Both Montizor and Hu said that they are not actively looking for a one-night date but that the culture of finding a formal date breeds pressure among first years to want a date.

"Some people really think it's important for them to be like everyone else and have a date," Montizor said. "I feel like from what I've seen it's a common struggle, and I'm not sure if I've heard of any successful matchmaking."

Montizor said that for those lucky enough to make it past the initial pairing stage, the likelihood of a later long-term relationship blossoming is low due to the "superficial" nature of matchmaking.

Walker Prejean '29 said that he has low hopes that he'll be paired with a special someone come formal night.

"I'm expecting the worst," Prejean says. "My expectations are low and I'm excited to see what happens — if I even get matched at all."

Now a senior, Adaora Mobisson '26 believes the tradition of matchmaking doesn't end after formals in students' early years at Yale. Instead, it persists throughout the entirety of one's time at Yale, she said.

"A lot of people jump into those really big connections and relationships, and it's the same for your senior year too," Mobisson said. "There's the desire to get a partner before you graduate because there's this huge idea at Yale where you want a Yale husband or a Yale wife."

Emely Cardenas '29 has a boyfriend, but said that, looking from the outside-in, her friends' matchmaking processes seem like "innocent fun." She said that since the spring semester's start people have been on the prowl for a formal date as a means of reinventing themselves or trying something new.

Cardenas believes that being set up at Yale makes it easier to meet people that "fit your vibe" that people might not have otherwise known. Mobisson added that while matchmaking culture isn't exclusive to Yale, the smallness of the campus amplifies buzz around pairing up.

"In a way, having someone else do that awkward part of asking them out for you makes it easier," Cardenas said. "But there's also another pressure in that now you've been set up and there's eyes on you by people who want it to work out."

In an Instagram post published on Tuesday, the Sophomore Class Council invited the class to "Take a gamble…get randomly matched with another sophomore for SOCO formal," which is casino night themed.

Sophomore Formal attendees were asked to fill out a form by Jan. 18 with their name, gender identity and the desired identity of their match, according to Sophomore Class Council Vice President Thy Luong '28, and the latter two questions are optional. The matches are then randomly generated using an online generator or artificial intelligence. Submitters will later receive an email with a match's contact information to begin mingling before formal.

"We wanted to include a matchmaking component this year because we saw an opportunity to both promote attendance, build up hype around the formal, and encourage opportunities for connection amongst the sophomore class," Luong wrote in an email to the News.

MAIA WILSON / STAFF ILLUSTRATOR
First years are engaging in typical matchmaking ahead of First Year Formal, and the Sophomore Class Council introduced a form to formalize the matchmaking process for its class formal.

Prior to formalizing the matchmaking process, the council consulted Yale Communication and Consent Educators to create a safe and fun opportunity, according to Luong. During the formal, couples will also be able to get to know one another better with question cards, games, photo booths and film cameras.

The council has also discussed continuing the tradition in future years, similar to First Year Formal matchmaking traditions, Luong wrote.

"In line with our Casino theme, we're hoping that chance and luck might smile upon students and create space for a new connection (or the opportunity to develop an old connection!)," Luong wrote.

The First Year Formal will be held in Commons on Feb. 20.

Contact
OLIVIA CYRUS *at*
olivia.cyrus@yale.edu
and **ANAYAH ACCILIEN** *at*
anayah.accilien@yale.edu.

AI model built by Yale and Google team finds cancer treatment

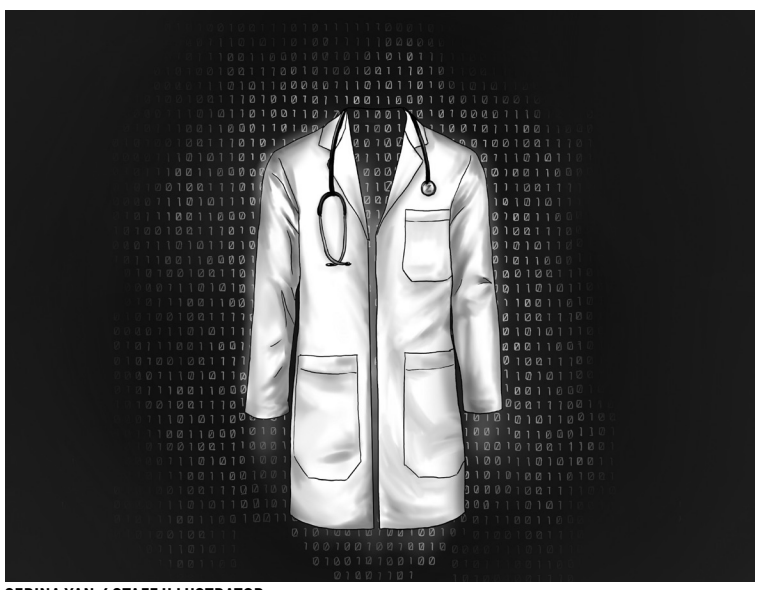
BY JONAS LOESEL
CONTRIBUTING REPORTER

When an artificial intelligence model suggested using a drug called Silmitasertib to help the human body locate cancerous tumors, Yale School of Medicine Professor David van Dijk said it was "surprising." The model wasn't relying on published literature it ingested in its training. In fact, no research had ever suggested that Silmitasertib could fight cancer in this way.

But his lab's custom AI model, called Cell2Sentence, hypothesized that the drug could increase antigen presentation — a process that helps the body's immune system identify rogue cells, which in turn helps the body fight cancer. Soon enough, his lab at the Yale School of Medicine tested the hypothesis on human skin and pulmonary cells. His lab's AI model was right.

The van Dijk lab published its findings in a preprint paper in October in collaboration with Google's top AI labs, Google DeepMind and Google Research.

"This is the first time that LLMs have been used to model gene expression, like really cellular biology, and made a prediction about the effect of a drug that was then experimentally validated," van Dijk, the lab's director, said in a telephone interview. Google's CEO, Sundar Pichai, wrote on X that the breakthrough was a "milestone" for AI in science.



SERINA YAN / STAFF ILLUSTRATOR
Collaborating with Google's top AI labs, Yale researchers developed a large-language model that identified a new treatment for cancer.

A new approach

Researchers in Yale's van Dijk lab have been working for years on a new approach to understanding the human cell. By analyzing single-cell RNA sequences, which show which genes are being expressed in each individual cell, the lab aims to understand the human body at a cellular level of detail.

"Concretely, we're trying to measure gene activity inside of cells," Syed Rizvi GRD '29, a computer science graduate student who was one of the lead authors in the October paper, said in a phone interview that month. He explained that if scientists figured out the genetic patterns within cells, they could easily identify which cells were "good or bad."

"Cells are 'thinking,' if you will, by expressing different genes," van Dijk said.

The researchers first turned to natural language processing, a branch of AI that teaches computers to understand human language, as a tool to decode the genomic activity. Single-cell RNA sequences are massive lists that count the amount of copies of each gene in a cell, but this numerical data was difficult for AI models to interpret, Rizvi said. By rendering the data as actual sentences, the lab found that the AI model quickly learned the biological patterns necessary to identify cells.

For example, a cellular sentence might look like "Gene A Gene B Gene C Gene D," continuing over the around 18,000 types of genes in the lab's datasets, Rizvi explained.

Their early models showed progress. In a paper published in October 2024, van Dijk's team announced that they were able to produce a model that could consistently identify the type of cells when given a cell sentence and even produce "biologically valid" cell sentences given a cell type input.

But their models were running on GPT-2, an older version of OpenAI's large-language models, which was originally released in 2019. The model was smaller and less complex than newer models available, including GPT-4, which was released in March of 2023.

Their model was initially limited in its ability to understand and reproduce these cell sentences. The GPT-2 model had around 774 million parameters, the adjustable weights that encode the model's knowledge and behavior. GPT-4 had 1.76 trillion parameters, making it more than 2,000 times more complex than the older model.

"The smaller models have a harder time picking up all the logic, the rules, the grammar, if you will, of this biological language," van Dijk said.

However, Rizvi said that the research computers available at Yale constrained them to using older, smaller models.

Yale operates six high performance computing clusters for its researchers. Each cluster is comprised of dozens or hundreds of individual computers, which link together to process the immense amounts of data required for large-language model training. While these clusters are much more powerful than the average laptop, they pale in comparison to the clusters operated at large tech companies, where the number of computers ranges from the hundreds of thousands to the millions.

Rizvi said that although Yale's computers were "pretty good" among academic institutions, they were unable to train the newest AI models.

Google steps in

In 2024, researchers from Google hosted an AI workshop at Yale. It was there, according to Google DeepMind scientist Shekoofeh Azizi, where staff at Google Research first met van Dijk's research team.

"The collaboration grew organically," Azizi wrote in a November email to the News.

Google Research and Google DeepMind are two of the company's heavily-funded AI research laboratories. With the Google partnership, the research team gained access to the company's massive computing resources, reportedly the largest of any company in the world — equivalent to more than a million of the type of GPUs in Yale's clusters, according to Epoch AI, an AI research institute.

"This type of research, unfortunately on the one hax=nd, you need this type of industry collaborator to be able to really scale, because that's just the thing with LLMs, they need large scale" van Dijk explained. "Fortunately, we had this possibility to work with Google."

As well as training their new model on Google's computers, the research team switched from a GPT-2 based model to Google's AI model, Gemma-2, which allowed them to expand Cell2Sentence to 27 billion parameters and feed the model significantly more data.

A key hypothesis with their Google collaboration was whether the new large-language model would perform better than the smaller one did. By scaling its AI models with Google, the team found that they indeed became more capable, as they had hoped.

"They become smarter at biology, understanding biology, generating biology, doing all sorts of different tasks," van Dijk said.

A breakthrough

The model's new scale allowed researchers to model the effects of drugs on human cells, what they called "perturbation response prediction." Armed with this ability, the researchers asked their model to identify drugs that could amplify an immune signal when a protein that signals sickness was already present.

The model identified Silmitaserib, traditionally used to inhibit cancer growth, and the researchers confirmed its hypothesis by testing the drug with human cells in a lab. The confirmation, Azizi wrote, showed that large-language models could successfully perform complex biological reasoning.

"This type of biological hypothesis is a more advanced form of biological reasoning, and shows the potential of large-scale AI and LLM-based systems to decode context-dependent effects which can accelerate drug discovery," she wrote.

Van Dijk said that the drug's developer, Senhwa Biosciences, is now studying this new feature of its product.

In the future, Azizi and van Dijk said that their work will open the door to faster, cheaper drug development across medicine. A 2024 study estimated that the mean development cost for a new drug, including the cost of failures, was around $500 million, and major pharmaceutical companies regularly spend billions in research and development each year.

"These pharma companies spend billions in doing pre-clinical research, and then it goes to trial and it fails. And then all this money is wasted," van Dijk said.

The researchers hope that their findings can help shortcut this pre-clinical phase, with AI models "guiding researchers to focus their laboratory efforts on experiments with the highest probability of success," according to Azizi.

But before that vision comes true, AI models need to keep scaling up. "We got a clear signal that scaling matters. So I want to keep on doing that," van Dijk said. Ultimately, he said, the goal is to use AI to completely model the human body.

"Now imagine you have a virtual human that can simulate everything biologically about real humans, but it doesn't complain. And it doesn't, you know, suffer. And now you can test all these drugs, many more and much faster, and have a much better filter to figure out which drugs are actually going to work in humans and are going to be safe," van Dijk said. "That is really the holy grail of drug development."

The lead authors on the October preprint were Yale's Rizvi, Yale postdoctoral associate Daniel Levine, Aakash Patel GRD '30, Shiyang Zhang GRD '31, Curtis Jamison Perry MED '18 GRD '18 and Google DeepMind's Eric Wang.

Contact
JONAS LOESEL *at*
jonas.loesel@yale.edu.