OpenAI starts offering a biology-tuned LLM

To address LLMs' tendencies toward sycophancy and overenthusiasm, OpenAI says it has tuned the model to be more skeptical, so it's more likely to tell you when something is a bad drug target. There was a lot of talk about GPT-Rosalind's "reasoning" and "expert-level" abilities. We were told that the former was defined as being able to work through complex, multi-step processes, while the latter was derived from the model's performance on a handful of benchmarks.

It's unclear whether OpenAI has tackled the hallucination issue that has plagued a variety of LLMs and can also strike when the systems are prompted to explain the steps the company took to reach its conclusions. Given past experience, it's likely we'll see a mix of glowing reports about unexpected connections the AI finds, as well as instances where it produces obviously erroneous suggestions.

For the moment, however, the company is limiting access due to concerns about the model's potential for harmful outputs if asked to do something like optimize a virus's infectivity. Only US-based entities can apply to OpenAI's trusted access deployment structure at the moment, and the company will limit who can use it. A more limited Life Sciences Research Plugin will be made generally available.

As noted above, a number of other companies have made science-focused agentic LLMs available, but those were much less focused than GPT-Rosalind, which is biology-specific. Until we start hearing reports on the effectiveness of this new model, it's difficult to evaluate whether this focus improves its utility.

https://arstechnica.com/science/2026/04/openai-starts-offering-a-biology-tuned-llm/

To teach in the time of ChatGPT is to know pain

What’s the point of building formative assessments into a course if they’re just handed off to an LLM? Suddenly, it’s a waste of time for both the student and the instructor. Small quizzes are excellent study tools to help students check their own understanding—if a student does them. Now, you can direct an “agentic” LLM browser to complete *all the quizzes in an entire course* with a single, frictionless prompt.

Should instructors preserve these sorts of assignments for students who want to benefit from them and accept the cheating, or should they eliminate the learning opportunity just to prevent cheating?

Evolution, the natural selection

Many instructors are trying to adapt to this crisis by going back to the only evaluation tools that are pretty much LLM-proof—tests like oral exams or handwritten work created under supervision in the classroom.

None of these solutions are available to instructors of asynchronous online classes. That sucks, since the availability of those classes is important. They can serve students with physical disabilities, students in rural areas

far from a campus, or students trying to obtain a degree while working full-time jobs or caring for dependents. If we have to simply give up on the idea of online classes, those are the casualties.

But even for in-person classes, adaptations to prevent LLM cheating are often concessions that reduce pedagogical quality. For example, labor-intensive oral exams didn’t become an endangered species just because of the swelling student-to-instructor ratio. Pen and paper (or keyboard and mouse) exams make it easier for each student’s experience to be the same and remove some of the potential for bias in scoring.

Writing assignments that may previously have been excellent teaching tools have obviously become the first things to end up on the chopping block. I used to have students in a natural disasters class write a plot for a big-budget Hollywood disaster movie, using both accurate and implausible physical processes. It was good practice for their writing skills; the students found it enjoyable, and it forced them to skillfully apply a lot of what they had learned.

https://arstechnica.com/science/2026/04/to-teach-in-the-time-of-chatgpt-is-to-know-pain/

Shock from Iran war has Trump’s vision for US energy dominance flailing

The closing of the Strait of Hormuz stranded tankers from Qatar and the United Arab Emirates, which together provide 20 percent of global LNG. Asia has been especially hard hit because it imports 80 percent to 90 percent of the supply from the Persian Gulf. The reopening of the strait will not restore all of the lost supply. In mid-March, Iranian missiles knocked out 17 percent of capacity at Qatar's Ras Laffan refinery, and QatarEnergy's CEO said repairs could take five years.

The United States has made an aggressive push to be a bigger part of the global LNG market, with Trump seeking to secure major purchase agreements from trade partners like Japan, the EU, and South Korea. But the eight existing US LNG export terminals are already running at full capacity. Although Trump has vowed to bring more capacity online, construction, and permitting of the complex multibillion-dollar facilities take years.

As a result, US exports of LNG, about 15 billion cubic feet of gas per day, are currently limited to only 11 percent to 13 percent of total US natural gas production. The situation leaves the United States with an abundance of its top fuel for electricity even while other countries are scrambling to stretch their supplies.

But American consumers have been coping with sharply rising electricity prices for a host of reasons unrelated to the war—mostly due to the capital build-out by utility companies, in part to accommodate the data center explosion but also to build resilience against wildfire, storms, and other climate change impacts and to replace aging infrastructure.

In their bi-monthly video series, energy analysts at the Center for Strategic and International Studies contemplated how the best example of US energy independence is almost wholly unnoticed by American consumers because of these other factors.

“So while we’re staring at the precipice of a global energy crisis, or might already be in one, the United States is going to feel that in oil markets, but we are, for the time being, by the nature of the gas system and the bountiful supply here in the United States, insulated against the gas price shocks?” asked Joseph Majkut, director of the CSIS’ Energy Security and Climate Change program.

https://arstechnica.com/science/2026/04/shock-from-iran-war-has-trumps-vision-for-us-energy-dominance-flailing/

Four astronauts are back home after a daring ride around the Moon

Slamming into the atmosphere at more than 30 times the speed of sound, NASA’s Orion spacecraft blazed a trail over the Pacific Ocean on Friday, returning home with four astronauts and safely capping humanity’s first voyage to the Moon in nearly 54 years.

Temperatures outside the capsule built up to some 5,000 degrees Fahrenheit as a sheath of plasma enveloped the Orion spacecraft, named *Integrity*, and its four long-distance travelers, temporarily blocking radio signals the Moon ship and Mission Control in Houston. Flying southwest to northeast, the spacecraft steered toward a splashdown zone southwest of San Diego, where a US Navy recovery ship held position to await the crew’s homecoming. Ground teams regained

communications with Orion commander Reid Wiseman after a six-minute blackout.

Airborne tracking planes beamed live video of Orion's descent back to Mission Control, showing the capsule jettison its parachute cover and deploy a series of chutes to stabilize its plunge toward the Pacific. Then, three larger main chutes, each with an area of 10,500 square feet, opened to slow Orion for splashdown at 8:07 pm EDT Friday (00:07 UTC Saturday).

In just 14 minutes, Orion bled off nearly 25,000 mph of velocity, subjecting the crew strapped into their seats to two brief periods of about 3.9 Gs.

The USS *John P. Murtha* amphibious transport dock ship dispatched helicopters and small boats to begin extracting Wiseman and his Artemis II crewmates: Victor Glover, Christina Koch, and Jeremy Hansen. Wiseman reported "four green crew members" inside the cockpit of the Orion spacecraft, confirming good health and high spirits after splashdown.

Koch exited the capsule first, joining Navy divers on an inflatable raft, or "front porch," assembled next to the spacecraft. Glover was next, then Hansen, a Canadian astronaut, stepped out of Orion onto the front porch. Wiseman, the captain of the ship, was last to leave his seat and join the recovery team. Two helicopters were expected to hoist the astronauts from the sea and fly them them to the *John P. Murtha*, where they were to undergo medical checks before traveling to San Diego, then back to Houston for a reunion with their families Saturday.

https://arstechnica.com/space/2026/04/four-astronauts-are-back-home-after-a-daring-ride-around-the-moon/

New paper argues history, not mantle plume, powers Yellowstone

There appear to be two separate arms originating from the same general location at the crust-mantle boundary. One branch slopes northeast to feed the Yellowstone caldera, while a second branches off toward the Snake River Plain. The branches split in a way that the volcano-free zone between the two features results.

The researchers reasoned that, whatever else was going on to provide molten material, the paths to the surface were likely to be enabled by stresses in the crust. And that was going to depend on both the existing features in the crust (obtained largely through seismic data) as well as larger-scale processes going on in the mantle underneath. So, the model included both basic geological details, known physical processes, and a bit of history in the sense of what we know about how that section of the crust came to be.

And that’s where we come back to the Farallon plate. Its remains, having been driven beneath the North American plate, are continuing to sink and move through the mantle. That, the researchers surmise, is driving a general eastward flow of material through the viscous mantle. Just east of Yellowstone, however, that flow runs into the older border of the North American plate, where the crust is thicker and denser than the portion of the continent that was put in place by the Farallon plate.

New pathways

This thick crust causes the flow of the mantle to dip downward. And that change in flow causes a series of stresses in the crust, most notably a compressive force between the older and newer sections of the North American plate, as well as a downward drag on the older section. Adding to the local stresses is the fact that all the material that erupted to form the Snake River Plain is denser than much of the surrounding rock, which generates strain on nearby rocks as it tries to sink.

https://arstechnica.com/science/2026/04/new-paper-argues-history-not-mantle-plume-powers-yellowstone/

“Oobleck” still holds some surprises

Mixing corn starch and water in appropriate amounts produces a slurry that is liquid when stirred slowly but hardens when you punch it—a substance colorfully dubbed “oobleck.” (The name derives from a 1949 Dr. Seuss children’s book, *Bartholomew and the Oobleck*.) High-speed imaging and force measurements have revealed another surprising property of oobleck drops hitting a flat surface, according to a new paper published in the journal Physical Review Letters.

As previously reported, in an ideal fluid, viscosity largely depends on temperature and pressure: Water will continue to flow regardless of other forces acting on it, such as stirring

or mixing. In a non-Newtonian fluid, the viscosity changes in response to an applied strain or shearing force, thereby straddling the boundary between liquid and solid behavior. Stirring a cup of water produces a shearing force, and the water shears to move out of the way. The viscosity remains unchanged. But for non-Newtonian fluids like oobleck, the viscosity changes when a shearing force is applied.

Ketchup, for instance, is a shear-thickening non-Newtonian fluid, which is one reason smacking the bottom of the bottle doesn’t make the ketchup come out any faster; the application of force increases the viscosity. Yogurt, gravy, mud, pudding, and thickened pie fillings are other examples. And so is oobleck.

The underlying physics principles of this simple substance are surprisingly nuanced and complex, and thus fascinating to scientists. For instance, back in 2023, molecular engineers at the University of Chicago used dense suspensions of piezoelectric nanoparticles to measure what is happening at the molecular level when oobleck transitions from liquid to solid behavior.

https://arstechnica.com/science/2026/04/oobleck-still-holds-some-surprises/

Oldest octopus fossil found to not be an octopus

Finding a radula, the chitinous, toothed, conveyor-belt-like

tongue used by mollusks to scrape food, was the definitive proof that Pohlsepia was indeed a mollusk. A more precise answer came in the form of the teeth. The synchrotron scan was sharp enough to let researchers count the tiny, individual tooth impressions across the radula's transverse rows. They observed a central rachidian tooth, pairs of lateral teeth, marginal teeth, and smooth marginal plates, adding up to at least 11 distinct elements per row. The only animals matching this exact 11-element configuration are Nautiloids, ancient, shelled relatives of modern cephalopods.

"Pohlsepia mazonensis is a nautiloid that died, decomposed, lost its protective shell and was subsequently flattened into the mud and preserved as an ambiguous stain," Clements said.

Switching titles

But while the fossil's "oldest octopus" badge is most likely gone, it managed to earn itself some new superlatives. "The first is of course the most difficult fossil to work on that I've ever worked on," Clements said. "The second would probably be the most fun fossil I've ever worked on."

The results of his study also made Pohlsepia the oldest unequivocal piece of evidence of nautiloid soft tissues ever found in the entire Paleozoic fossil record. "I'd say unequivocal because other nautiloid soft tissue fossils are a bit questionable," Clements said.

His team now plans to use the same advanced imaging on other fossils. "Mazon Creek in particular is full of very interesting and very weird fossils which have not been looked at for a long time, and you know, all of these new techniques can be used on them," Clements said.

He thinks that even for Pohlsepia itself, it's not yet the end of the story. "It's the oldest nautiloid now, so it would be interesting to study it more."

Clements’ study on the Pohlsepia mazonensis is published in Proceedings of the Royal Society B Biological Sciences: https://doi.org/10.1098/rspb.2025.2369

https://arstechnica.com/science/2026/04/oldest-octopus-fossil-found-to-not-be-an-octopus/

Orion helium leak no threat to Artemis II reentry but will require redesign

NASA’s schedule currently puts the launch of Artemis III in 2027 and Artemis IV in 2028. Kshatriya said he was confident that NASA, working with the European Space Agency and Airbus, which builds the service module, will be able to fix the valve problem in time for Artemis IV. Manufacturing of the Artemis IV service module is largely complete.

“I’m pretty sure we’re going to need to, at a minimum, tweak the design to prevent the leak rate that we have, if not fundamentally change the way the valve works,” he said.

Valves are a common bugaboo on rockets and spacecraft. Nearly every US human spaceflight program has dealt with malfunctioning or leaky valves. Boeing’s Starliner crew capsule suffered helium leaks in its propulsion system, along with other issues, during a test flight to the International Space Station in 2024. Helium valves on the Space Launch System rocket had to be replaced in the run-up to the Artemis

I and Artemis II launches. SpaceX has also scrubbed launches due to valve problems. The list goes on.

“There are a lot of options for how to take care of this problem,” Kshatriya said of the issue on the Orion spacecraft. “If anything, I’d characterize it as a production redesign risk for the Artemis IV mission, which I think we can get in front of, and which is why we put so much attention on it during this mission to make sure [we understand] what we’re seeing.”

The big lesson NASA learned on Artemis I involved the capsule’s heat shield. The ablative thermal barrier burned away unevenly as the craft reentered the atmosphere, but Orion still made it to a safe, on-target splashdown. NASA officials said they are confident the heat shield will hold up on Artemis II after adjusting the path Orion will take through the upper atmosphere. A new heat shield design will debut on Artemis III.

NASA engineers spent two years investigating the heat shield issue after Artemis I. Kshatriya does not expect the valve redesign to take as long.

“It’s not a safety of flight, safety of crew, must-work function like the heat shield investigation sent us down,” he said. “It’s going to take work to get it right, but it’s not of that magnitude.”

https://arstechnica.com/space/2026/04/nasa-homes-in-on-likely-redesign-to-fix-orion-spacecrafts-leaky-valves/

Clinical trial shows gene editing works for β-Thalassaemia, too

Almost as soon as researchers started exploring the capabilities of the CRISPR/Cas9 system, they recognized its potential use in targeted gene editing. But the intervening decades have seen slow progress as people worked to determine how to do so in a way that would be safe for use in humans. It was only a little over two years ago, decades after CRISPR's discovery, that the FDA approved the first CRISPR-based therapy, for sickle-cell anemia.

Now, following up on that success, a large Chinese collaboration has followed up with a description of an improved gene editing system that produces more focused changes and fewer mistakes. And they've used it to produce a therapy that addresses a disease that's closely related to sickle-cell anemia: β-Thalassaemia.

Gene editing and its limits

The CRISPR/Cas-9 system provides bacteria with a form of immunity. It uses specially structured RNAs (called guide RNAs) that can base-pair with a targeted sequence. The Cas-9 protein then recognizes this structure and cuts the DNA nearby. This is quite effective when the guide RNA can base-pair with a DNA virus, as the resulting cut will inactivate the virus.

There are a couple of ways to use this for DNA editing in organisms such as ourselves. Both of these take advantage of

the fact that the DNA repair systems in cells will often chew back the ends of these cuts before linking them back together again. This will frequently lead to small deletions at the site of the cut, which can be used to disable genes. The size of these deletions will vary, so you have to do some DNA sequencing to find one that disables the gene you're interested in, but doesn't do any additional damage.

Alternately, any deleted sequence can sometimes be repaired using a matching sequence, which is typically found on the other copy of the same chromosome. If the CRISPR-based cut is accompanied by lots of copies of a modified sequence, then it's possible for repair systems to insert the modifications into the genome, providing a true editing capability. But again, this process is error-prone, so people typically need to edit a bunch of cells and sequence the DNA to make sure the right changes are made.

https://arstechnica.com/science/2026/04/clinical-trial-shows-gene-editing-works-for-%ce%b2-thalassaemia-too/

Ugandan chimps split into two factions, then killed rivals

Even those factors might not have been sufficient, but the authors also identified three possible catalysts. In 2014, five adult males and one adult female died of unknown causes, although several had shown signs of illness. Those losses likely disrupted the social network by weakening social ties across clusters. There was also a new alpha male from the

western cluster the following year, coincidentally, the same time when the first sustained separation occurred. The two prior alphas had been from the central cluster, so that change in the dominance hierarchy may have exacerbated inter-group tensions.

Finally, there was a respiratory outbreak in January 2017 that killed 25 chimps, which could have sped up the final separation. “Taken together, these events suggest how networks may fracture in the face of multiple demographic and social changes,” the authors concluded.

“A hostile split among wild chimpanzees is a reminder of the danger that group divisions can present to human societies,” Brooks wrote in his perspective. “However, humans also engage with, bond, and cooperate at multiple levels across intersecting groups. The group relationships of humans are nuanced, diverse, and flexible. This flexibility enables deep cooperation but also underlies acts of violence. Humans must learn from studying the group-based behavior of other species, both in war and at peace, while remembering that their evolutionary past does not determine their future.

“If relational dynamics alone can drive polarization and lethal conflict in chimps without language, ethnicity, or ideology, then in humans, those cultural markers might be secondary to something more basic,” said Sandel. “If that’s true, then we may have the potential to reduce societal conflicts in our personal lives, and that gives me hope. As our paper concludes, it may be in the small, daily acts of reconciliation and reunion between individuals that we find opportunities for peace.”

Science, 2026. DOI: 10.1126/science.adz4944 (About DOIs).

https://arstechnica.com/science/2026/04/ugandan-chimps-split-into-two-factions-then-killed-rivals/