

# هوش مصنوعی به عنوان همکار علمی

از زیست‌شناسی تا سیاه‌چاله‌ها، ChatGPT در حال شتاب‌دهی به علم است.

ژانویه ۲۰۲۲

OpenAI

مهرشیر



## مقدمه: چرا هوش مصنوعی برای علم اهمیت دارد؟



تعداد پیام‌های ماهانه  
در حوزه علوم پیشرفته  
در سال گذشته  
**نزدیک به ۵۰ درصد**  
رشد داشته است.



**تقریباً ۱,۳ میلیون کاربر**  
**هفتگی ChatGPT**  
در سراسر جهان بر  
موضوعات علمی و ریاضیات  
پیشرفته تمرکز دارند.



**میانگین ۸,۴ میلیون**  
**پیام هفتگی**  
در موضوعات دشوار  
علوم پایه و ریاضیات

شرکت OpenAI در حال ساخت ابزارهایی است تا به پژوهشگران در تولید بینش‌ها، تسریع کشفیات علمی و اثرگذاری این بینش‌ها در دنیای واقعی کمک کند. در حال حاضر پژوهشگران، دانشجویان، اعضای هیئت علمی STEM (علوم، فناوری، مهندسی و ریاضیات) و مهندسان از ChatGPT برای مطالعه و ترکیب منابع علمی و فنی، کدنویسی و دیباگ کردن، تحلیل داده و برنامه‌ریزی آزمایش‌ها استفاده می‌کنند. هر هفته، ChatGPT نزدیک به ۸.۴ میلیون پیام در مورد موضوعات پیشرفته در علوم و ریاضیات دریافت می‌کند. این پیام‌ها در حال حاضر از سوی حدود ۱.۳ میلیون کاربر هفتگی در سراسر جهان ارسال می‌شوند.

بر اساس گزارش یونسکو، تنها حدود ۱۰ درصد از جمعیت جهان خود را به عنوان دانشمند معرفی می‌کنند، با این حال آن‌ها تأثیری فراتر از این نسبت دارند. پژوهش‌های علمی موتور محرک پیشرفت به سوی آینده‌ای سالم‌تر، مرفه‌تر و تاب‌آورتر هستند. داروهای جدید، فناوری‌های جدید و صنایع جدید از دانش نوینی نشئت می‌گیرند که به کار گرفته شده است. گروه کوچکی از فیزیک‌دانان در اوایل قرن بیستم، پایه‌های مکانیک کوانتوم را از طریق پژوهش‌های انتزاعی بنا نهادند که دهه‌ها بعد، زیربنای بخش اعظم اقتصاد دیجیتال مدرن را تشکیل داد؛ اقتصادی که اکنون ده‌ها تریلیون دلار ارزش‌گذاری می‌شود. پژوهش‌های بنیادی؛ یعنی کارهایی که پیش از مشخص شدن بازدهی‌شان انجام می‌شوند، سرچشمه این دانش بودند. در سال ۱۹۴۷، دانشمندان آزمایشگاه‌های Bell بر اساس بینش‌های حاصل از فیزیک کوانتومی، اولین ترانزیستور را ساختند. این ترانزیستور به یکی از اجزای سازنده کامپیوترها، تلفن‌ها و فناوری دیجیتال امروزی تبدیل شد. همچنین GPS و همگام‌سازی ساعت‌های اتمی، به بینش‌های اینشتین در مورد نسبیت متکی است.

با این حال، در بسیاری از حوزه‌ها، تداوم پیشرفت دشوارتر است. اقتصاددانان و تحلیل‌گران پژوهشی به کاهش «بهره‌وری پژوهش» اشاره می‌کنند، به این معنا که برای تولید همان تعداد بینش به افراد، زمان و هزینه بیشتری نیاز است. نیمه‌هادی‌ها مثال بارزی در این زمینه هستند. حفظ پایداری «قانون مور» (Moore's Law) برای دو برابر کردن تعداد ترانزیستورها در یک مدار مجتمع هر دو سال یک‌بار، نیازمند افزایش چشمگیری در پژوهش‌های علمی بوده است، به‌طوری که برآورد می‌شود تعداد پژوهشگران مورد نیاز در حال حاضر بیش از ۱۸ برابر تعداد مورد نیاز در اوایل دهه ۱۹۷۰ باشد. با پیچیده‌تر شدن دانش، هر نسل جدید از پژوهشگران تنها برای رسیدن به مرزهای دانش با بار کار سنگین‌تری روبرو می‌شود که این امر زمان آموزش آن‌ها را طولانی‌تر و تخصص‌هایشان را محدودتر می‌کند. از نظر نهادی، ساختار پژوهش به سمت تیم‌های بزرگ‌تر تغییر جهت داده است که با افزایش هزینه‌های سربار برای پروپوزال‌های دریافت کمک‌هزینه، تطابق با مقررات، گزارش‌دهی و هزینه‌های هماهنگی همراه است.

در حوزه پزشکی، پیشرفت‌های علمی جان افراد بی‌شماری را نجات داده است. به‌عنوان مثال امید به زندگی در سراسر جهان از حدود ۳۲ سال در ۱۹۰۰ به حدود ۷۳ سال در ۲۰۲۳ (و به بیش از ۷۸ سال در ایالات متحده) افزایش یافته است. اما بار باقی‌مانده سایر بیماری‌ها همچنان سنگین است. سازمان بهداشت جهانی گزارش می‌دهد که بیماری‌های غیرواگیر مانند سکت، بیماری‌های قلبی، سرطان و دیابت همچنان عامل حدود ۷۴ درصد مرگ‌ومیرها در سطح جهانی هستند. حتی زمانی که سرعت پیشرفت بالاست، تبدیل ایده‌های جدید به درمان‌های در دسترس زمان می‌برد. در ایالات متحده، به‌طور میانگین بین ۱۰ تا ۱۵ سال از کشف هدف درمانی تا تأیید مقرراتی یک داروی جدید طول می‌کشد؛ تأخیری که بر بیمارانی که به درمان‌های جدید و بهتر نیاز دارند، تحمیل می‌شود.

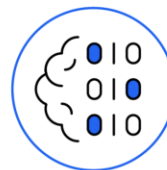
تسریع روند پیشرفت، باعث نجات جان انسان‌ها و بهبود کیفیت زندگی آن‌ها خواهد شد. هوش مصنوعی در حال حاضر به رفع تنگنایی که باعث کندی علم می‌شوند کمک می‌کند. پژوهش‌های مدرن در میان رشته‌های مختلف پراکنده شده و توسط محدودیت‌های شناختی و لجستیکی مقید شده‌اند؛ از جمله: خواندن و هضم حجم عظیمی از منابع پژوهشی برای تعیین آنچه مشخص شده است، ترجمه ایده‌ها به ریاضیات و کد، تنظیم تحلیل‌ها و شبیه‌سازی‌ها، بررسی محاسبات، جست‌وجو در فضاها، طراحی عظیم و تصمیم‌گیری در مورد اینکه کدام آزمایش‌های آینده امیدوارکننده‌تر هستند. در صورت استفاده صحیح، هوش مصنوعی می‌تواند به‌عنوان یک همکار با توان عملیاتی بالا برای تفکر، محاسبات و استدلال ساختاریافته عمل کند، چرخه از فرضیه تا آزمایش را کوتاه کرده و ظرفیت پژوهشگرانی که به‌تنهایی یا در تیم‌ها کار می‌کنند را حتی در فراسوی مرزهای بین‌رشته‌ای، افزایش دهد.

«کوین ویل» (Kevin Weil)، معاون بخش علوم در OpenAI این فرصت را این‌گونه توصیف می‌کند: «هوش مصنوعی روزبه‌روز بیش از گذشته به‌عنوان یک همکار علمی مورد استفاده قرار می‌گیرد و ما شاهد رشد تأثیر آن در محیط‌های پژوهشی واقعی هستیم. پژوهشگران بیشتری برای پیشرفت در مسائل حل‌نشده، تفسیر داده‌های پیچیده و تکرار سریع‌تر در کارهای آزمایشی از سیستم‌های استدلال پیشرفته استفاده می‌کنند. میزان این استفاده در طول سال گذشته به‌سرعت رشد کرده است و نتایج آن رفته‌رفته در حوزه‌های مختلف نمایان می‌شوند. ما هنوز در مراحل اولیه هستیم، اما سرعت پذیرش و کیفیت کار نشان می‌دهد که علم در حال ورود به یک فاز شتاب‌دهی جدید است.»

از همین رو OpenAI با شرکای پژوهشی در آژانس‌های دولتی، آزمایشگاه‌های ملی، دانشگاه‌ها و بخش پزشکی همکاری می‌کند. برخی از این شرکای پژوهشی شامل: وزارت انرژی ایالات متحده، آزمایشگاه ملی Lawrence Livermore، مراکز کنترل و پیشگیری از بیماری‌های ایالات متحده، دانشگاه‌های هاروارد، MIT، آکسفورد، A&M و بیمارستان کودکان بوستون هستند.

گزارش حاضر به شرح موارد زیر می‌پردازد:

- چگونه ابزارهای هوش مصنوعی در جریان‌های کاری روزمره پژوهش از جمله ترکیب منابع پژوهشی، کدنویسی و دیباگ‌کردن، تحلیل داده، پشتیبانی از شبیه‌سازی و برنامه‌ریزی آزمایش‌ها مورد استفاده قرار می‌گیرند.
- نتایج اولیه در مورد ظرفیت هوش مصنوعی برای پشتیبانی از دستاوردهای مهم جدید چه چیزی را نشان می‌دهند.
- دانشمندان رشته‌های مختلف چگونه از ChatGPT برای پیشرفت در حوزه کاری خود استفاده کرده‌اند.
- پیشنهاد‌های سیاستی برای حمایت از پیشرفت مستمر هوش مصنوعی در علم و ریاضیات.



## مقیاس کار علمی و ریاضی در ChatGPT

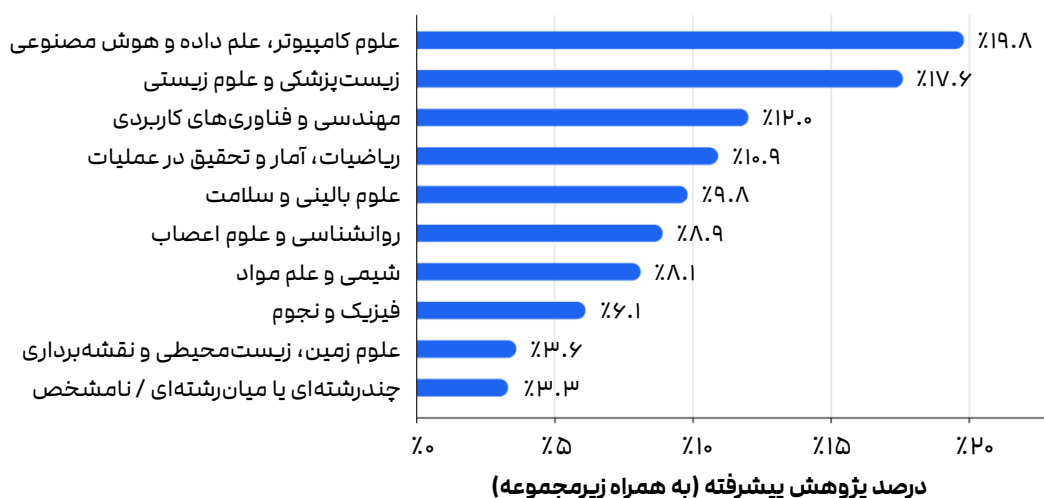
گروهی کوچک اما تأثیرگذار از پژوهشگران از مدل‌های هوش مصنوعی OpenAI برای وظایف پیچیده‌ای اعم از استخراج‌های فنی، ریاضیات پیشرفته، شبیه‌سازی و مدل‌سازی مهندسی و سایر حل‌مسئله‌های پیشرفته استفاده می‌کنند. این گروه شامل دانشمندان و ریاضی‌دانان، از کاندیداهای دکترا و محققان پسادکترا گرفته تا پژوهشگران شاغل و اعضای هیئت علمی STEM می‌شود.

بر اساس یک تحلیل داخلی از یک نمونه تصادفی از مکالمات ناشناس ChatGPT از ژانویه تا دسامبر ۲۰۲۵، میانگین تعداد پیام‌های هفتگی در موضوعات پیشرفته علوم و ریاضیات با رشد حدوداً ۴۷ درصدی، از ۵.۷ میلیون پیام به نزدیک به ۸.۴ میلیون پیام در طول سال ۲۰۲۵ رسید. تا ژانویه ۲۰۲۶، نزدیک به ۱.۳ میلیون کاربر هفتگی در مورد موضوعات پیشرفته در علم و ریاضی با ChatGPT گفت‌وگو می‌کنند.

مجموع این نشانه‌ها از جمله ده‌ها میلیون پرامپت ماهانه در زمینه علوم پایه و ریاضیات پیشرفته که توسط گروهی بزرگ و در حال رشد که از این سیستم برای کارهای جدی علمی و مهندسی به‌منظور منفعت جامعه و حمایت از رشد اقتصادی استفاده می‌کنند؛ نشان می‌دهد که ChatGPT چگونه در حال شتاب‌دهی به پژوهش‌های پیشرفته است.

در ادامه نحوه تفکیک استفاده از ChatGPT در رشته‌های مختلف در میان گروه کاربران پژوهش‌محور آورده شده است:

## علوم پیشرفته: سهم پژوهش پیشرفته به تفکیک رشته (تقسیم بندی علوم فیزیکی، ۲۰۲۵)



## دانشمندان و ریاضی‌دانان در واقعیت چگونه از ChatGPT استفاده می‌کنند؟

دانشمندان، ریاضی‌دانان و مهندسان از ChatGPT به‌عنوان یک همکار فنی که همیشه در دسترس است استفاده می‌کنند؛ ابزاری که با آن می‌توانند محاسبات را تکرار کنند، ایده‌ها را به کد تبدیل کنند، مفروضات را ارزیابی کنند و مطالب پیچیده را در قالب مدل‌های ذهنی کارآمد فشرده کنند. در تحلیل OpenAI، پرامپت‌های علوم پایه «پیشرفته» به‌عنوان مواردی تعریف می‌شوند که پژوهش‌محور بوده و برای پاسخگویی کارآمد، نیازمند تخصص در سطح تحصیلات تکمیلی یا سطح پژوهشی هستند. رفتار کاربری این گروه نسبت به شیوه رفتار کاربران عادی تفاوت‌های بسیاری دارد که مستقیماً با جریان‌های کاری پژوهش‌های مدرن مطابقت دارد.

وظایف پژوهشی در حوزه‌هایی زیر خوشه‌بندی می‌شوند: کدنویسی (پیش‌نویس، بازنویسی و دیباگ)، تحلیل داده (پاک‌سازی و ادغام مجموعه‌داده‌ها، اجرای آمار، تفسیر نتایج)، استدلال ریاضی (استخراج، استراتژی‌های اثبات، بررسی جبری، محاسبات طولانی، ترجمه فرمول‌بندی‌ها)، و مرور و ترکیب منابع پژوهشی (یافتن منابع، مرور منابع جدید).

کاربران پژوهشگر علوم و ریاضیات پیشرفته در مقایسه با کاربران معمولی ChatGPT:

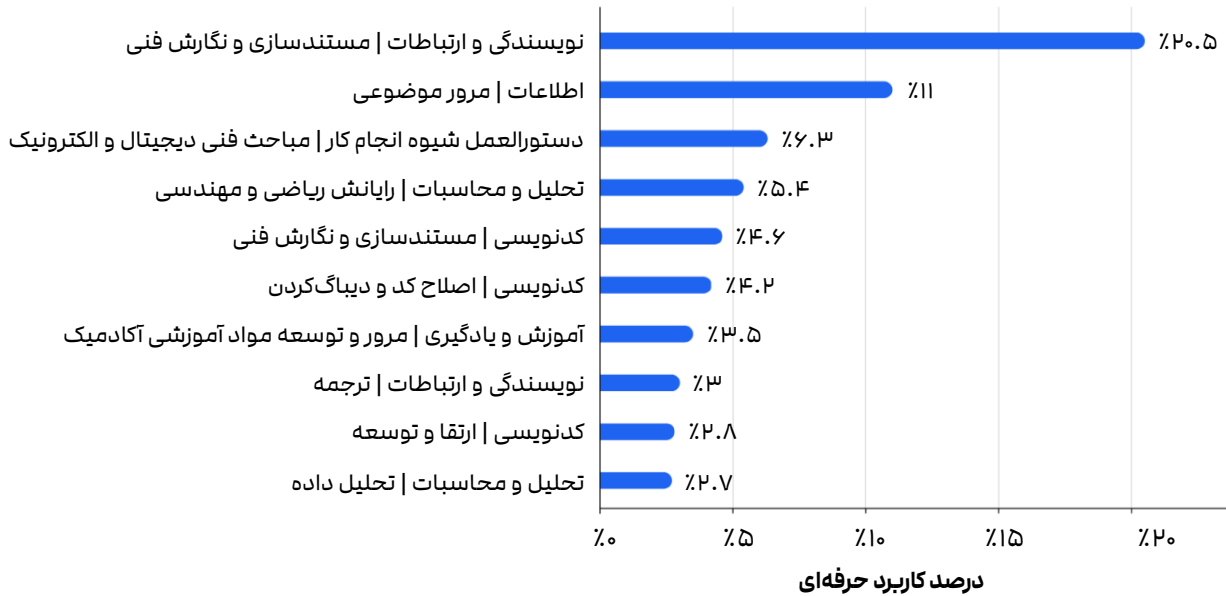
تقریباً ۳.۵ برابر بیشتر از متوسط پایه پیام ارسال می‌کنند.

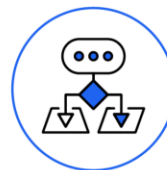
تقریباً ۱۲ برابر بیشتر، پیام‌های مرتبط با کدنویسی ارسال می‌کنند.

به طور میانگین در هفته ۹ پرامپت برای مرور کلی در مقایسه با ۱.۵ پرامپت ارسال می‌کنند.

نحوه استفاده کاربران حرفه‌ای از ChatGPT به تفکیک پرکاربردترین وظایف:

## ۱۰ کاربرد حرفه‌ای و پیشرفته به تفکیک قابلیت





## هوش مصنوعی در مرزهای ریاضیات و علم

### قابلیت‌های هوش مصنوعی پیشگام در ریاضیات

#### پیشرفت‌های اخیر

طی دو سال گذشته، مدل‌های زبانی بزرگ از عملکرد اولیه و نابرابر در محاسبات پایه، به مدیریت استدلال ریاضی چندمرحله‌ای که می‌تواند در پژوهش‌های علوم ریاضی مفید واقع شود، ارتقا یافته‌اند. بخش عمده‌ای از این بهبود، نتیجه روش‌های تقویت استدلال گام‌به‌گام و همچنین یکپارچگی بهتر با ابزارهایی مانند ماشین حساب و اجرای کد برای محاسبات دقیق بود. هم‌زمان با بهبود مدل‌ها، روند توسعه بنچمارک‌ها نیز به سمت آزمون‌های سخت‌تری تغییر جهت داد که برای سنجش استدلال عمیق‌تر و درعین حال کاهش موفقیت‌های صرفاً ناشی از «تطبیق الگو» (Pattern Matching) طراحی شده‌اند.

در ۲۰۲۵ و اوایل ۲۰۲۶، بیشترین تأثیرگذاری ناشی از «مقیاس‌پذیری رایانش در زمان آزمایش» یا «تفکر آهسته» بوده است. به جای انتخاب سریع به یک مسیر، مدل محاسبات بیشتری را صرف بررسی مسیرهای جایگزین و خودبررسی (Self-Checking) می‌کند. هم‌زمان، رویکردهای آموزشی که به نتایج قابل‌تأیید، مانند تولید یک پاسخ نهایی صحیح یا کد قابل اجرا پاداش می‌دهند، ریاضیات و کدنویسی را به سمت قابل‌اعتمادتر شدن سوق داده‌اند و به اندازه‌ای درست عمل می‌کنند که با هدایت انسان مفید باشند.

یکی از نشانه‌های این تغییر در المپیاد بین‌المللی ریاضی در جولای ۲۰۲۵ نمایان شد، زمانی که یک مدل OpenAI در کنار DeepMind عملکردی در سطح مدال طلا را در مجموعه مسائل ۲۰۲۵ به دست آورد.

#### قابلیت‌های فعلی

پیشرفت پیوسته GPT-5.2 در قابلیت‌های ریاضی نتیجه زنجیره استدلال طولانی‌تر و قوی‌تر، روندهای تأیید سیستماتیک و استفاده بهتر از ابزارهای قابل‌بررسی است. GPT-5.2 Thinking توانست در آزمون ریاضی AIME 2025 که برای سنجش حل مسئله چندمرحله‌ای طراحی شده است، بدون استفاده از ابزارهای خارجی نمره کامل را کسب کند. خانواده GPT-5.2 شامل مدل‌های Thinking و Pro فراتر از عملکرد در سطح مسابقات، به سمت اکتشافات ریاضی از جمله کار بر روی مسائل لاینحل قدیمی پیش رفته است.

در ارزیابی‌های مبتنی بر سبک پژوهشی، مجموعه مسائل FrontierMath که به Google-Proof به معنی مقاوم در برابر جست‌وجوی گوگل مشهور هستند، به‌گونه‌ای طراحی شده است که تنها برای متخصصان واقعی قابل‌دسترسی و حتی قابل‌حل باشد. در این بنچمارک، GPT-5.2 Thinking توانسته ۴۰.۳ درصد از مسائل سطوح ۱ تا ۳ را حل کند. بااین‌حال، عملکرد در سخت‌ترین سطح همچنان کاهش یافت و GPT-5.2 Pro در سطح ۴ آزمون FrontierMath امتیازی معادل ۳۱ درصد را در مجموعه‌ای از مسائل که می‌توان آن‌ها را به عنوان «پروژه‌های پژوهشی کوچک» توصیف کرد، کسب کرد.

دومین جهش بزرگ در قابلیت‌ها زمانی رخ داد که GPT-5.2 با جریان‌های کاری فرمولاسیون ادغام شد. در یک اقدام برجسته، GPT-5.2 اثبات‌هایی را به زبان طبیعی تولید می‌کند و از یک مدل زبانی بزرگ شخص ثالث به نام Aristotle، برای فرموله کردن آن اثبات‌ها در Lean استفاده کرد. Lean یک دستیار اثبات است که در آن اثبات‌ها به شکلی نوشته می‌شوند که کامپیوتر آن‌ها را گام‌به‌گام بررسی می‌کند و سیستم شکاف‌ها را در طول فرموله‌سازی تشخیص داده و اصلاح می‌کند.

اهمیت این ادغام‌ها از آن جهت است که یکی از موقعیت‌هایی که برای مدل‌های زبانی همیشه با شکست روبه‌رو بوده زمانی است که یک راه‌حل ریاضیاتی «درست به نظر می‌رسد»، اما در واقع این‌طور نیست؛ یعنی استدلال‌هایی که معقول به نظر می‌رسند، اما نقاط ضعف ظریفی دارند. اثبات‌های بررسی‌شده با Lean استاندارد اطمینان به یک اثبات را با اجبار به انجام مراحل صریح و بررسی‌شده به‌صورت مکانیکی تحت یک چارچوب رسمی مشخص، به طور قابل‌توجهی ارتقا می‌دهند.

### مسائل اردوش و راه‌حل‌های هوش مصنوعی

«پل اردوش» (Paul Erdős) ریاضی‌دان مجارستانی (۱۹۱۳-۱۹۹۶) و جهانگردی بود که زندگی‌اش را با یک چمدان می‌گذراند و در جست‌وجوی همکاران پژوهشگر ریاضی خود از دانشگاهی به دانشگاه دیگر می‌رفت. «مسائل اردوش» مجموعه عظیمی از پرسش‌ها و حدسیاتی هستند که او مطرح کرد؛ از معماهای به‌ظاهر ساده تا مسائلی که هنوز در برابر بهترین تکنیک‌های ریاضی مقاومت می‌کنند. این مسائل مانند نشانگرهای مسیر برای ریاضیات مدرن عمل کرده‌اند؛ یعنی روشن کردن آنچه هنوز درک نکرده‌ایم، پایه‌ریزی کلی برنامه‌های پژوهشی و کشاندن نسل‌هایی از ریاضی‌دانان به‌سوی لبه‌های دانش.

در اوایل سال ۲۰۲۶، مدل GPT-5.2 با کمک ابزارهایی مانند Aristotle و Lean راه‌حل‌هایی برای چندین مسئله حل‌نشده اردوش ارائه کرد که توسط «ترنس تائو» (Terence Tao) نیز تأیید شدند. مسائل شماره ۲۸۱، ۷۲۸ و ۷۲۹ اکنون به‌عنوان اثبات‌شده و شماره ۳۹۷ به‌عنوان رد شده فهرست شده‌اند. درحالی‌که ریاضی‌دانان هشدار می‌دهند که درجه سختی مسائل اردوش به‌شدت متفاوت است، این راه‌حل‌ها نشان‌دهنده توانایی روبه‌رشد مدل‌های OpenAI برای انجام پژوهش واقعی ریاضی و شیوه جدیدی از تعاملات با حداقل راهنمایی است.

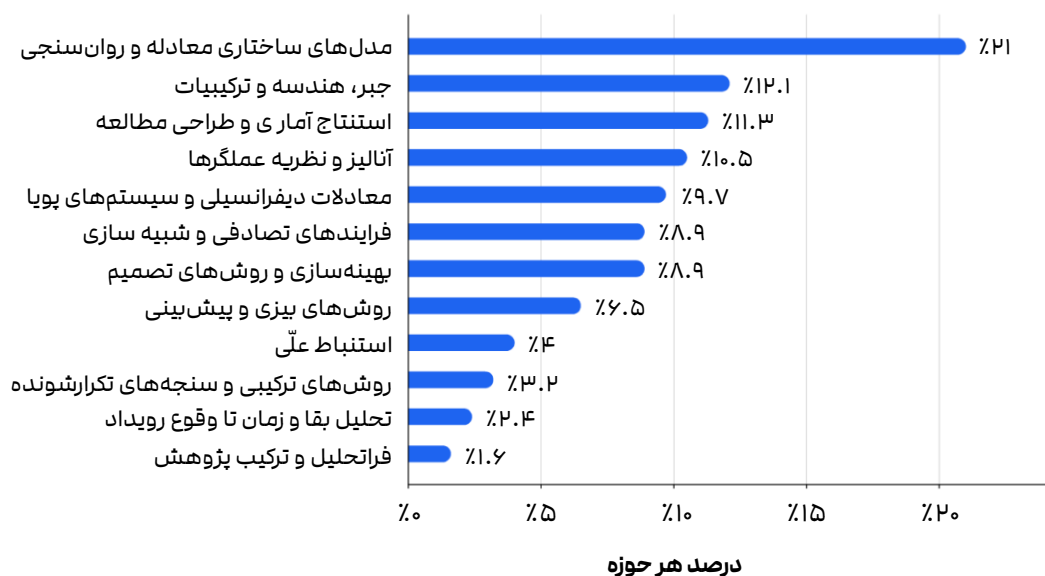
### چشم‌انداز ظرفیت‌های کوتاه‌مدت

برخی از اکتشافات مهم ریاضی می‌توانند به شکل پیوند دادن روش‌های شناخته‌شده به یکدیگر برای یافتن استدلال صحیح باشند. در حال حاضر GPT-5.2 می‌تواند این کار را در بسیاری از موارد انجام دهد. سایر اکتشافات شامل اختراع انواع کاملاً جدیدی از ریاضیات هستند؛ درست همان‌طور که نیوتن حساب دیفرانسیل و انتگرال را برای درک دینامیک (نیروها، جرم و انرژی که تغییرات در حرکت را توضیح می‌دهند) ابداع کرد. این موضوع فراتر از توانایی مدل‌های فعلی هوش مصنوعی است. اما نوع سوم از اکتشافات مهم شامل ایجاد ارتباط بین دو حوزه و انتقال سازوکارها، نتایج و ابزارهای شناخته‌شده یک حوزه به حوزه دیگر است (مانند هندسه جبری که از جبر مجرد و هندسه کلاسیک پدید آمد). نمونه‌های متوسطی از این مورد تاکنون با هوش مصنوعی انجام شده و OpenAI بر این باور است که اهمیت این ارتباطات و زیرشاخه‌های ایجاد شده از آن‌ها در آینده نزدیک افزایش خواهد یافت.

بخش عمده‌ای از ارزش‌آفرینی کوتاه‌مدت هوش مصنوعی در تغییر و تحول جریان‌های کاری خواهد بود. GPT-5.2 می‌تواند مسیرهای راه‌حل را پیشنهاد داده و تکرار کند، درحالی‌که ابزارهای خارجی صحت آن‌ها را از طریق محاسبات دقیق یا فرموله کردن تضمین می‌کنند. این امر با روندهای گسترده‌تر رویکردهای ترکیبی و فرمولاسیون خودکار همسو است، جایی که ریاضیات غیررسمی به زبان‌های رسمی مانند Lean ترجمه می‌شود تا صحت آن‌ها به صورت مکانیکی قابل تأیید باشد. GPT-5.2 در حال حاضر برای بررسی پیشنهاد (بررسی آنچه شناخته شده) به منظور یافتن لبه‌های دانش و آشکارساختن منابع غیرمنتظره یا مبهم مفید است. این ابزار در ارائه اثبات‌ها، نقد و ساده‌سازی آن‌ها و پیشنهاد استراتژی‌های اثبات کارایی دارد.

اگر این مسیر ادامه یابد، تأثیر کوتاه‌مدت GPT-5.2 احتمالاً به صورت یک ارتقای گسترده در بهره‌وری برای پژوهشگران ریاضی و همچنین تیم‌های علمی و مهندسی (از آنجایی‌که ریاضیات جزء اصلی بسیاری از کارهای علمی و مهندسی است) ظاهر خواهد شد. این امر به شکل ترجمه سریع‌تر از یک شرح مسئله درهم‌ریخته به یک بیان ریاضی تمیز، کاهش محدودیت‌های از قلم افتاده در استخراج‌های چندمرحله‌ای، اشکال‌زدایی مطمئن‌تر محاسبات و اثبات‌ها و رشد شمار نتایجی که به درستی فرموله می‌شوند، تجلی خواهد یافت.

### علوم پیشرفته؛ ریاضیات، آمار و تحقیق و در عملیات: درصد استفاده در هر رشته (تقسیم بندی علوم فیزیکی، ۲۰۲۵)



## قابلیت‌های هوش مصنوعی پیشگام در علم

در رشته‌هایی مانند فیزیک، شیمی و زیست‌شناسی، مدل‌های زبانی بزرگ در کلاس ChatGPT از استدلال فنی و جریان‌های کاری پژوهشی مبتنی بر ابزار و همچنین نگارش علمی پشتیبانی می‌کنند. بنچمارک‌هایی مانند GPQA که مجموعه‌ای از پرسش‌های سطح تحصیلات تکمیلی و Google-proof تألیف‌شده توسط متخصصان هستند، در ابتدا برتری قابل‌توجه انسان را نشان می‌دادند؛ به طوری که متخصصان به دقت ۶۵ درصدی دست یافتند و خط پایه مدل GPT-4 به ۳۹ درصد رسید. OpenAI اکنون دقت ۹۳.۲ درصدی را در بنچمارک GPQA Diamond برای مدل GPT-5.2 Pro و ۹۲.۴ درصدی را برای مدل GPT-5.2 Thinking (بدون فعال‌سازی هیچ ابزاری و با حداکثر تلاش استدلالی) گزارش می‌دهد که نشان‌دهنده خط پایه بالاتری برای پاسخگویی به پرسش‌های علمی در سطح تحصیلات تکمیلی در بسیاری از رشته‌های علمی است. به موازات آن، خودکارسازی کارهای پرزحمت و ناگفته پژوهش مانند جستجوی منابع، گردآوری کتاب‌شناسی و گزارش‌دهی اداری معمول وقت محدود و ارزشمند دانشمندان را برای کارهای ارزشمندتر آزاد می‌کند. این مدل‌ها در ترکیب با بازیابی اطلاعات و ابزارهای قابل‌اجرایی که محاسبات قابل‌بررسی و اعتبارسنجی گام‌به‌گام را امکان‌پذیر می‌سازند، در حال تبدیل شدن به هماهنگ‌کننده‌های (Orchestrators) قابل‌اعتمادی جریان کاری برای برنامه‌ریزی، تحلیل و مستندسازی علمی هستند و زمینه را برای شتاب‌دهی در بسیاری از حوزه‌ها فراهم می‌کنند.

### فیزیک

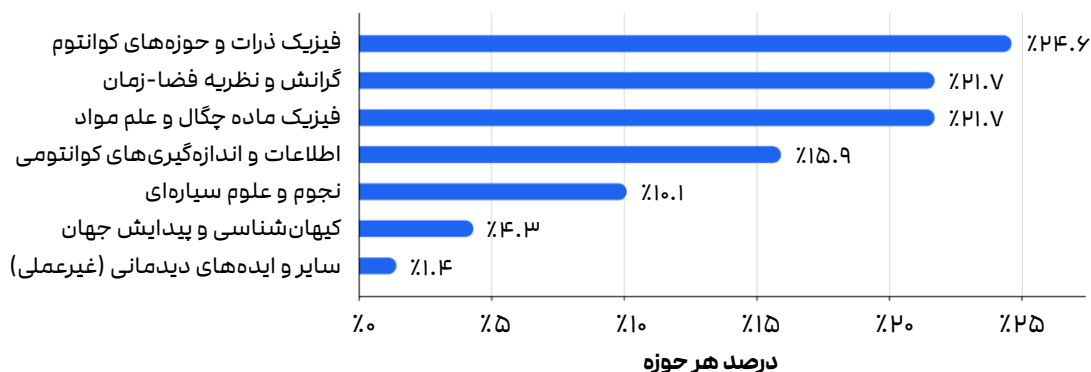
OpenAI اخیراً تفاهم‌نامه‌ای با وزارت انرژی ایالات متحده (DOE) برای حمایت از همکاری در زمینه هوش مصنوعی و محاسبات پیشرفته به‌منظور پیشبرد طرح‌های این وزارتخانه از جمله مأموریت Genesis با کاربردهایی در حوزه انرژی مانند پژوهش‌های همجوشی هسته‌ای منعقد کرده است.

در فیزیک، LLMها در نهادهای علمی بزرگ و متعددی از جمله بسیاری از آزمایشگاه‌های ملی ایالات متحده، به‌عنوان یک لایه یکپارچه‌ساز بر روی ساختار و پشته‌های عملیاتی (Operations Stacks) پیچیده و پایگاه‌های دانش داخلی مورد‌استفاده قرار می‌گیرند و در کنار ابزارهای موجود یادگیری ماشین برای شبیه‌سازی، کاهش بلادرنگ داده‌ها و کنترل آزمایشی به تحلیل و تصمیم‌گیری تحت محدودیت‌های سخت‌گیرانه کمک می‌کنند. LLMها می‌توانند گزارش‌های شیف‌ت و هشدارها را درک کنند، به پرسش‌هایی که از مستندات داخلی سرچشمه می‌گیرند پاسخ دهند و به هدایت کار به سمت ابزار تحلیل، شبیه‌سازی یا کنترل مناسب کمک کنند که همه این‌ها تحت محدودیت‌های سخت‌گیرانه ایمنی، زمان‌بندی و منابع انجام می‌شود. این امر در همان راستای تاریخچه طولانی استفاده از یادگیری ماشین تخصصی در فیزیک است. نمونه‌ای از این کاربردهای پیشین، «مدل‌های عصبی جانشین» (Neural Surrogate Models) است که شبیه‌سازی‌های مبتنی بر معادله را زمانی که محاسبات کامل بسیار کند است تقریب می‌زنند، فیلترسازی و بازسازی بلادرنگ در آشکارسازهای ذرات که ده‌ها میلیون برخورد را در ثانیه مشاهده می‌کنند. همچنین کنترل‌کننده‌های مبتنی بر یادگیری ماشین که آهنرباهای الکتریکی متعدد را در آزمایش‌های همجوشی Tokamak هماهنگ کرده و درعین‌حال در محدوده محدودیت‌های سخت‌افزاری و ایمنی باقی می‌مانند.

دستاوردهای کوتاه‌مدت در فیزیک احتمالاً در محیط‌های با توان عملیاتی بالا و تصمیم‌محور متمرکز خواهند شد؛ جایی که توجه متخصصان و زمان به‌عنوان گلوگاه عمل می‌کنند. دستیارهای هوش مصنوعی که می‌توانند به مستندات داخلی یک آزمایشگاه ارجاع دهند و بررسی‌های خودکار را اجرا کنند، قادرند هشدارهای زنده آزمایش، یادداشت‌ها و فایل‌های لاگ را به برنامه‌های پژوهشی اولویت‌بندی‌شده برای گام بعدی و خروجی‌های تحلیلی تکرارپذیر مانند نوت‌بوک‌ها، اسکریپت‌ها و گزارش‌ها تبدیل کنند.

در فیزیک نظری، LLMها همچنان به ارزش آفرینی به عنوان شرکای فکری ادامه خواهند داد و ضمن کاهش بار شناختی پژوهشگران، فضای اکتشاف را گسترش می‌دهند. هنگامی که یک محاسبه پیچیده به بن‌بست می‌رسد، LLMها می‌توانند راه‌هایی را برای قاب‌بندی مجدد مسئله ارائه دهند، مراحل میانی را پیشنهاد کنند و بررسی‌های سازگاری سریعی را انجام دهند که زمان درگیرشدن فیزیک‌دانان را کاهش می‌دهد. گاه این امر به بینش‌هایی مانند یک شرط جاافتاده، یک فرمول‌بندی تمیزتر یا یک رابطه مفید بین عبارت‌ها منجر می‌شود که می‌تواند به عنصر معناداری از یک مقاله تبدیل شود. ضریب اثرگذاری بزرگ‌تر ممکن است از ترکیب پژوهش‌ها در مقیاس وسیع ناشی شود؛ جایی که مدل‌ها مقالات و زیرشاخه‌ها را در ادبیات بررسی می‌کنند تا ارتباطات جدید را به صورت خودکار آشکار سازند.

### علوم پیشرفته؛ فیزیک و نجوم: درصد استفاده در هر رشته (تقسیم بندی علوم فیزیکی، ۲۰۲۵)

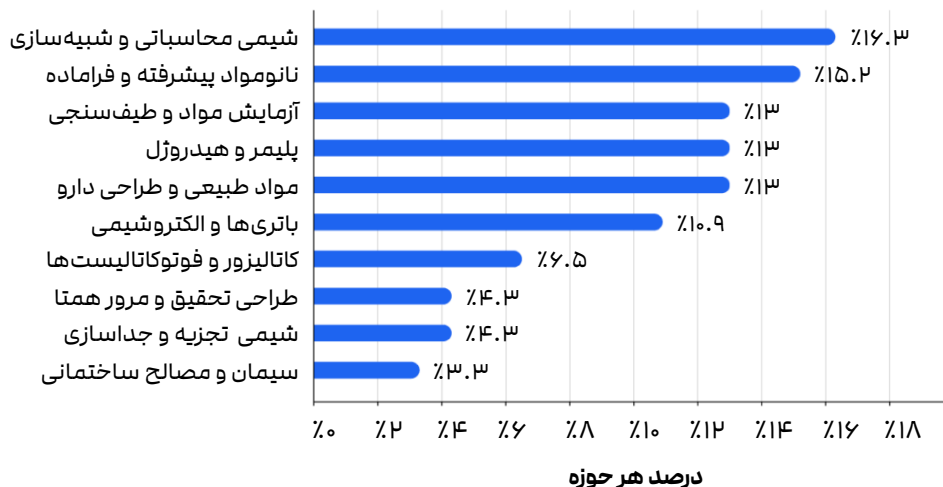


### شیمی

کاربردهای ChatGPT در شیمی از پاسخ‌گویی تک‌مرحله‌ای به سمت جریان‌های کاری چندمرحله‌ای که بین زبان طبیعی و بازنمایی‌های شیمیایی ترجمه می‌کنند و برای اعتبارسنجی و بازیابی به ابزارهای خارجی متکی هستند، پیش رفته است. بنچمارک ChemBench که در سال ۲۰۲۵ در نشریه Nature Chemistry منتشر شد، بیش از ۲۷۰۰ پرسش نوشته‌شده توسط متخصصان را گردآوری کرد و نشان داد که مدل‌های پیشرو به طور میانگین عملکرد بهتری نسبت به شیمی‌دانان انسانی داشتند؛ درحالی‌که هنوز در برخی وظایف اساسی با مشکل مواجه بوده و بعضاً با اعتماد به نفس کاذب پاسخ‌های اشتباه تولید می‌کردند.

سیستم‌های پیشرو هوش مصنوعی در شیمی از یک جریان کاری ترکیبی استفاده می‌کنند. یک LLM همه‌منظوره به برنامه‌ریزی کار چندمرحله‌ای و هماهنگی ابزارها کمک می‌کند، درحالی‌که مدل‌های تخصصی که ساختار مولکولی را درک می‌کنند، وظیفه پیش‌بینی و شبیه‌سازی را بر عهده دارند. یک مثال کلیدی، «شبکه‌های عصبی گرافی» (GNN) پیشرفته هستند. این مدل‌ها با یک مولکول مانند یک شبکه، با اتم‌ها به عنوان گره‌ها و با پیوندها به عنوان اتصالات برخورد می‌کنند؛ بنابراین سیستم می‌تواند بیاموزد که چگونه تغییرات داخلی بر کل ساختار تأثیر می‌گذارد. GNNهای جدیدتر به گونه‌ای طراحی شده‌اند که پیش‌بینی‌های آن‌ها در هنگام چرخش یا جابه‌جایی یک مولکول در فضای سه‌بعدی ثابت باقی می‌ماند که این ویژگی آن‌ها را برای یادگیری قوانین انرژی مورد نیاز برای اجرای شبیه‌سازی‌های مولکولی سریع و دقیق بسیار مناسب می‌سازد. درحالی‌که مدل‌های گراف مولکولی با داده‌ها و پیش‌آموزش بیشتر مقیاس می‌یابند، نتایج همچنان بهبود پیدا می‌کنند؛ اما بنچمارک‌های سخت‌تر برای استدلال شیمیایی از جمله سازوکارهای واکنش‌های آلی مانند oMeBench، بر نیاز به نظارت انسانی تأکید دارند.

### علوم پیشرفته؛ شیمی و علم مواد: درصد استفاده در هر رشته (تقسیم بندی علوم فیزیکی، ۲۰۲۵)

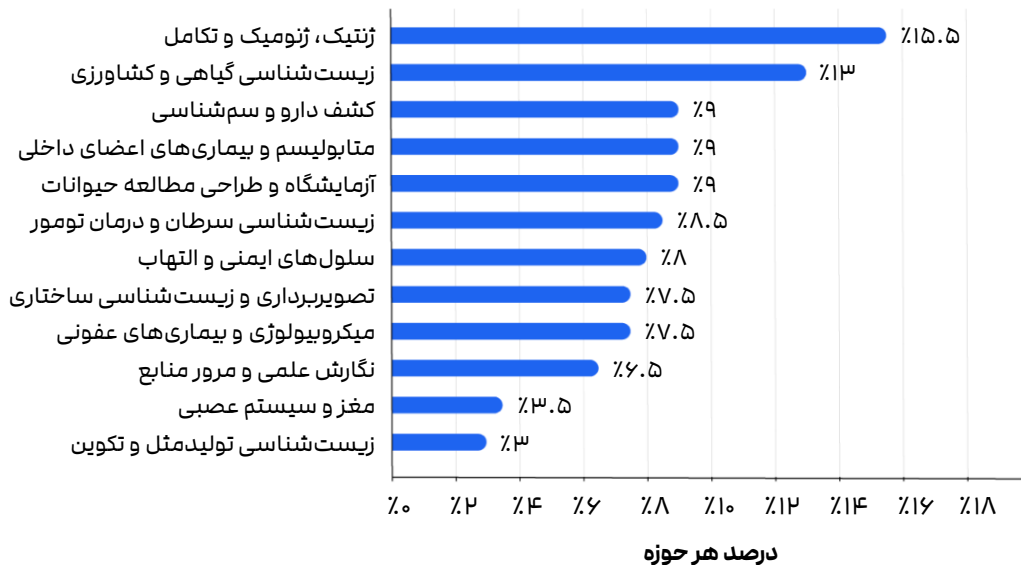


### زیست‌شناسی

عمده کاربردهای ChatGPT در زیست‌شناسی مربوط به جریان‌های کاری چندمرحله‌ای است که پرسش‌های مبتنی بر زبان طبیعی را با منابع علمی ساختاریافته مانند پایگاه‌های داده ژنومیک، مخازن پروتئین و منابع زیست‌پزشکی ترکیب می‌کنند که اغلب همراه با کد و ابزارهای بازایی برای ردیابی و اعتبارسنجی مورد استفاده قرار می‌گیرند. بنچمارک GeneTuring که یک ارزیابی ژنومیک است که در سال ۲۰۲۵ در نشریه Briefings in Bioinformatics توسعه یافته است، ۱۶۰۰ پرسش را در ۱۶ نوع وظیفه گردآوری کرد و ۴۸۰۰۰ پاسخ را از پیکربندی‌های مختلف مدل‌ها به صورت دستی ارزیابی کرد. بهترین نتایج از یک چیدمان تقویت‌شده با ابزار (Tool-augmented setup) به دست آمد که یک مدل همه‌منظوره را با دسترسی مستقیم به رابط‌های برنامه‌نویسی کاربردی (API) مرکز ملی اطلاعات بیوتکنولوژی (NCBI) جفت می‌کرد. این امر نشان می‌دهد که زمانی که مدل‌های زبانی به داده‌های مرجع معتبر متصل شوند و بتوانند روند کار خود را نشان دهند، قابلیت اطمینان بهبود می‌یابد.

همانند شیمی، پژوهش‌های پیشرفته مبتنی بر هوش مصنوعی در زیست‌شناسی نیز بر ساختار و پشته‌های ترکیبی متکی هستند. مدل‌های زبانی همه‌منظوره به برنامه‌ریزی و هماهنگی تحلیل کمک می‌کنند، درحالی‌که مدل‌های پایه تخصصی که بر روی توالی‌ها و ساختارهای بیولوژیکی آموزش دیده‌اند، موتور محرک پیش‌بینی و طراحی هستند. در علم پروتئین، AlphaFold 3 با پیش‌بینی ساختار سه‌بعدی مشترک کمپلکس‌هایی که می‌توانند شامل پروتئین‌ها، DNA، RNA و مولکول‌های کوچک در یک معماری مبتنی بر انتشار (Diffusion-based Architecture) باشند، گامی روبه‌جلو به سوی مدل‌سازی بیومولکولی یکپارچه محسوب می‌شود.

**علوم پیشرفته؛ زیست پزشکی و علوم زیستی:**  
**درصد استفاده در هر رشته (تقسیم بندی علوم فیزیکی، ۲۰۲۵)**





## هوش مصنوعی در مرزهای ریاضیات و علم

### امست ریو (Emest Ryu) - ریاضی‌دان



ارنست ریو در سال ۲۰۲۳ از روی کنجکاوی به سراغ ChatGPT رفت و شاهد پیشرفت آن بود تا جایی که سال ۲۰۲۵ توانست نتیجه‌ای قابل انتشار تولید کند. کار آکادمیک ریو بر روی بهینه‌سازی متمرکز است؛ یعنی ریاضیات پشت الگوریتم‌های کارآمد و قابل اعتمادی که در اقتصادهای مدرن از برنامه‌ریزی لجستیک گرفته تا حفظ پایداری بال‌های هواپیما کاربرد دارد.

زمانی که مدل‌های زبانی بزرگ در سال ۲۰۲۳ به سرعت به محبوبیت دست یافتند، ریو اولین آزمایش‌های خود را آغاز کرد؛ آیا یک مدل می‌تواند «مسائل کلامی» دنیای واقعی را به مدل‌های بهینه‌سازی دقیق، از جمله تمام محدودیت‌های پنهان ترجمه کند و سپس آن‌ها را به یک حل‌کننده (Solver) تحویل دهد؟ به عنوان مثال، زمان‌بندی یک فصل بیسبال نیازمند محدودیت‌های دشوار (مثلاً هیچ تیمی دو بازی را هم‌زمان انجام ندهد) و محدودیت‌های آسان‌تر (مثلاً روزهای استراحت برای سفر که در صورت لزوم می‌توانند نقض شوند) است. آن مدل اولیه با مدیریت دقیق محدودیت‌هایی که این کار می‌طلبید درگیر بود و گاهی اوقات محدودیت‌ها را حذف می‌کرد و در زمان‌بندی‌های بزرگ‌تر و واقع‌بینانه شکست می‌خورد.

نقطه عطف در سال ۲۰۲۵ و پس از ورود مدل‌های استدلالی و کسب مدال طلا توسط OpenAI در المپیاد بین‌المللی ریاضی اتفاق افتاد و همان دسته از مسائل زمان‌بندی که ریو قبلاً آزمایش کرده بود، حل شدند. این موفقیت ریو را بر آن داشت تا LLM‌ها را در کارهای روزمره ریاضی به کار گیرد. ریو در حین نوشتن درس‌نامه‌ها، از ChatGPT می‌خواست تا نتایجی را که می‌دانست درست هستند؛ اما آن‌ها را در ذهن نداشت، اثبات کند.

در نهایت، او ChatGPT را در پژوهش امتحان کرد. ریو مسئله‌ای مرتبط با شتاب‌دهی نستروف (Nesterov Acceleration)؛ تکنیکی شناخته‌شده برای سرعت بخشیدن به بهینه‌سازی را انتخاب کرد و نسخه‌ای از مسئله را برگزید که به اندازه کافی حل نشده باقی‌مانده بود که دیگران آن را امتحان کرده بودند، اما به اندازه کافی ساده بود که ممکن بود اثبات کوتاهی برای آن وجود داشته باشد. برای سه شب متوالی، او از ساعت ۸ شب تا نیمه‌شب کار کرد و در شب سوم هوش مصنوعی را به نقطه‌ای رساند که توانست مسئله را حل کند.

همکاری آن‌ها شبیه یک پژوهش واقعی بود. مدل اثبات اولیه‌ای با یک اشتباه محاسباتی تولید کرد؛ بنابراین ریو شروع به تکرار و اصلاح کرد. او خطا را اصلاح کرد، مراحل میانی صحیح را در یک پرامپت حفظ کرد، مسیرهای بن‌بست را رها کرد و مدل را به سمت رویکردهای دیگر سوق داد. ریو این روند را به عنوان دویدن در یک هزارتو توصیف می‌کند؛ جایی که گاهی اوقات فقط برای اینکه آن‌ها را خالی بیابید به سمت راهروها می‌پیچید و درها را باز می‌کنید، درحالی‌که نقشه ذهنی از آنچه شکست می‌خورد و آنچه امیدوارکننده به نظر می‌رسد را حفظ می‌کنید. ChatGPT به ریو کمک کرد تا سرعت دویدن خود در این هزارتو را ۳ تا ۱۰ برابر افزایش دهد.

در شب سوم، مدل جهش کوچک اما معناداری را انجام داد که به اندازه کافی «متفاوت به نظر می‌رسید» تا قفل اثبات را باز کند. ریو گفت که او استدلال را پیش از اینکه آن را به صورت عمومی با جامعه بهینه‌سازی به اشتراک بگذارد، «پیش از سه بار بررسی کرد» و سپس از یک دانشجو خواست آن را تأیید کند. جامعه علمی نیز با شگفتی و هیجان نسبت به آن واکنش نشان دادند. از آن نقطه، نتیجه زمان‌پیوسته با یک پرامپت واحد به بیان الگوریتم زمان‌گسسته ترجمه شد و یک خلاصه جدید و کوتاه یک‌صفحه‌ای باقی ماند که استاندارد یک دستاورد قابل انتشار را برآورده می‌کرد.

از آن زمان، ریو به تیم داده‌های مصنوعی OpenAI پیوسته است، جایی که تمرکز اصلی او بهبود قابلیت‌های ریاضی مدل است.

### الکس لوپساسکا (Alex Lupsasca) - فیزیک‌دان



الکس لوپساسکا همان‌گونه به سراغ هوش مصنوعی رفت که بسیاری از فیزیک‌دانان با ادعاهای بزرگ روبرو می‌شوند؛ شک‌گرایی مؤدبانه و مجموعه‌ای از آزمایش‌ها. اوایل سال ۲۰۲۵، او ChatGPT را امتحان کرد و آن را برای کارهای اداری معمول که مرتباً در دانشگاه پیش می‌آیند، مفید یافت. اما آن را به عنوان ابزاری برای بخش سخت کار؛ یعنی تبدیل قوانین فیزیک به پیش‌بینی‌های ملموس و قابل اعتبارسنجی نمی‌دید.

لوپساسکا می‌گوید واقعیت رایج در نشر آکادمیک این است که یک پروژه پژوهشی اغلب ممکن است به پیش‌نویسی از معادلات و بافت‌های ارتباطی خام منجر شود که کار یک استاد دانشگاه را به کار یک ویراستار تبدیل می‌کند، به خصوص زمانی که به زبانی به‌غیر از زبان مادری خود می‌نویسند. ChatGPT می‌تواند یک پیش‌نویس خام را بگیرد و آن را به نثری علمی و تمیزتر تبدیل کند که باعث صرفه‌جویی قابل‌توجهی در زمان می‌شود. لوپساسکا به‌رغم تجربه‌اش می‌گوید ابزارهای هوش مصنوعی تقریباً همیشه در نوشتن از او بهتر هستند.

با این حال، با بهبود مدل‌ها او همچنان به آزمایش قابلیت‌های بیشتر ادامه داد. آنچه نظر او را تغییر داد، تماشای سیستمی بود که سرانجام مجموعه مسائل در سطح تحصیلات تکمیلی فیزیک را با سرعت بالا حل می‌کرد. لوپساسکا به یک سؤال با بیان ساده از نسبت عام اشاره کرد: با شروع از یک مدل کتاب درسی شناخته‌شده برای میدان‌های الکترومغناطیسی در نزدیکی یک سیاه‌چاله (به اصطلاح راه حل Wald)، قدرت میدان مغناطیسی درست در لبه سیاه‌چاله، یعنی افق رویداد چقدر است؟ یک دانشجوی مبتدی تحصیلات تکمیلی ممکن است ساعت‌ها برای حل آن زمان نیاز داشته باشد، اما مدل کل راه حل را در عرض چند ثانیه تولید کرد.

سپس او متوجه چیز عجیب‌تری شد. مدل اغلب به پاسخ‌های صحیحی می‌رسید، اما آن‌ها را به زبان ریاضی ناآشنا بیان می‌کرد، گویی گویش‌های متعددی از فیزیک و ریاضی را آموخته بود و می‌توانست هر کدام را که نتیجه را بهتر فشرده می‌کرد، انتخاب کند.

آزمایشش بزرگ‌تر از تحقیقات خود لوپساسکا حاصل شد. او اخیراً «تقارن‌های پنهان» جدیدی از معادله‌ای که بر پاسخ جزرومدی یک سیاه‌چاله حاکم است استخراج کرده بود که تقریباً مشابه جزرومد اقیانوس‌ها توسط ماه است. این تقارن‌ها توضیح می‌دهند که چرا یک اثر جزرومدی معروف از بین می‌رود. وقتی او این معادله را با حداقل راهنمایی به GPT-5 Pro تحویل داد، حدود ۱۸ دقیقه فکر کرد و همان تولیدکننده‌های تقارنی را بازگرداند که او سال‌ها برای دستیابی به مهارت‌های یافتن آن‌ها و ماه‌ها کار مستقیم روی آن‌ها وقت صرف کرده بود. او گفت: «باورکردنی نیست و به وضوح قرار است همه کارهایی را که ما انجام می‌دهیم تغییر دهد.»

او از آن زمان این الگو را با همکاران شکاک خود تکرار کرده است: در «سرن» (CERN)، یک فیزیکدان مسئله‌ای در سبک تکالیف ارائه داد که معمولاً به دانشجویان دکترا یک هفته فرصت برای حل آن می‌داد و مدل در چند دقیقه راه‌حل مفصلی تولید کرد. «الیوت کوآتارت» (Elliot Quataert)، اخترفیزیکدان در «اسپن» (Aspen) سعی کرد آن را با یک سیگنال‌گذاری معمایی فریب دهد و مدل به‌درستی یک مگنتار (ستاره نوترونی با میدان مغناطیسی شدید) را شناسایی کرد، سپس مشاهدات بعدی و پیش‌نویس یک چکیده را ارائه داد.

این سیر تکاملی، از شک‌گرایی تا اشتیاق مسری، باعث شد لوپ‌سازکا به OpenAI بپیوندد. او اکنون فراتر از موفقیت‌های موردی، به سمت شتاب‌دهی علمی تکرارپذیر و توسعه ابزارهای بهتر برای خواندن و توضیح مقالات، جریان‌های کاری قوی‌تر از یک پنجره چت واحد و تنظیمات آموزشی که مرزهای فیزیک جدید را در قابلیت‌های مدل جاسازی می‌کند، پیش می‌رود. هدف او این است که پیامدهای بینش‌های فیزیکی را سریع‌تر آشکار سازد تا پژوهشگران زمان کمتری را در جبر گیر کنند و زمان بیشتری را صرف شناسایی سؤال بعدی که ارزش پرسیدن دارد کنند تا در نهایت بزرگ‌ترین رازهایی که رشته او را تسخیر کرده‌اند را کشف کنند.

## RetroBioSciences - زیست‌شناسی

شرکت RetroBioSciences در حال پیگیری یک چالش کلیدی در زمینه طول عمر است: عملی و مقیاس‌پذیر کردن جوان‌سازی سلولی. یکی از اهرم‌های کلیدی، برنامه‌ریزی مجدد سلولی با استفاده از چهار فاکتور OSKM است که می‌تواند جنبه‌هایی از سن سلولی را بازنشانی کند؛ اما این فرایند در عمل و به‌خصوص در سلول‌های پیرتر، کند و ناکارآمد است. مهندسی پروتئین سنتی می‌تواند این عوامل را بهبود بخشد؛ اما به گواه تاریخ، پیشرفت در این زمینه سال‌ها آزمون و خطا را طلبیده است که نمونه بارز آن «ابرسویه» (Super Variants) قبلی هستند که کشف آن‌ها نیازمند کارهای آزمایشگاهی طولانی و پرهزینه بود.

OpenAI با RetroBio همکاری کرد تا با ساخت یک مدل پایه‌ای متمرکز بر پروتئین به نام GPT-4B Micro، این جدول زمانی را فشرده کند. GPT-4B Micro برای انجام از وظایف گسترده مهندسی پروتئین از جمله ویرایش هدفمند پروتئین از طریق پر کردن (In-filling) و تولید ساختار سه‌بعدی و تاخوردگی پروتئین طراحی شده است. این مدل به‌جای آموزش صرف بر روی توالی‌ها، بر روی چندین مدل‌الیه بیولوژیکی از جمله توالی‌های پروتئینی، ساختارهای سه‌بعدی توکنیزه‌شده، اطلاعات تکاملی از ترازسازی توالی‌های چندگانه، داده‌های تعامل پروتئین-پروتئین و متون علمی مرتبط و به‌صورت نیمه‌آموزش‌دیده (Mid-Trained) آموزش دیده است. هدف از این ترکیب، ایجاد یک مدل واحد بود که بتواند در میان جریان‌های کاری رایج طراحی پروتئین انعطاف‌پذیر باشد. کارهایی مانند تولید توالی‌ها، انجام ویرایش‌های پر کردن میانی برای حفظ مناطق بحرانی درحالی‌که مناطق دیگر مجدداً طراحی می‌شوند و تولید توکن‌های ساختاری که با ترکیبات سه‌بعدی محتمل مطابقت دارند.

در همکاری با RetroBio، این مدل در یک جریان کاری «آزمایشگاه در حلقه» (Lab-in-the-Loop) برای مهندسی فاکتورهای OSKM بهبودیافته استفاده شد. یک استراتژی پرامپت‌نویسی مبتنی بر سری‌های تکاملی باعث شد سیستم قابل‌هدایت باشد. سیستم الگوهای را برای همه پروتئین‌های مرتبط یاد گرفت و توانست توالی‌های جدیدی را مشروط به مثال‌ها پیشنهاد دهد. در عمل، مدل هزاران توالی بالقوه تولید کرد که برای حفظ دامنه‌های ضروری و حفظ تنوع فیلتر شدند و سپس به چند صد سویه تقلیل یافتند که برای یک آزمایشگاه معمولی قابل‌آزمایش بود. RetroBio درنهایت DNA این سویه‌ها را سنتز کرد، آن‌ها را از طریق ساختارهای لنتی‌ویروسی (Lentiviral) به فیبروبلاست‌های انسانی انتقال داد، سلول‌ها را برای حدود ۱۴-۱۰ روز مجدداً برنامه‌ریزی کرد و از نشانگرهای سطح سلول به علاوه توالی‌یابی استفاده کرد تا رتبه‌بندی کند که کدام سویه‌ها در جمعیت‌هایی که با موفقیت مجدداً برنامه‌ریزی شده‌اند، غنی‌تر شده‌اند.

مستندات اولیه، تغییرات مورفولوژی تسریع‌شده و نشانگرهای چندتوانی (Pluripotency-Marker) بهبودیافته با نرخ موفقیت بالایی را نشان دادند. برای برخی از اهداف، حدود نیمی از سویه‌های تولیدشده توسط مدل از پروتئین‌های نوع مهاجم بهتر عمل کردند، حتی زمانی که توالی‌ها به طور چشمگیری با همتایان طبیعی خود متفاوت بودند.

آزمایش‌های بعدی بر روی بهترین سویه‌های منفرد، عملکردی قابل‌مقایسه با و در برخی موارد فراتر از بهترین فاکتورهای مهندسی‌شده قبلی در کلاس خود را با کلونی‌های شبه سلول بنیادی بزرگ‌تر و متعددتر نشان دادند. RetroBio همچنین مراحل اعتبارسنجی پایین‌دستی مرتبط با ترجمه را گزارش کرد، از جمله بررسی‌هایی مبنی بر اینکه سلول‌های بازبرنامه‌ریزی‌شده ظرفیت تمایز موردانتظار را حفظ کرده و هیچ ناهنجاری کروموزومی آشکاری در غربالگری‌های ایمنی اولیه را نشان ندادند. سنجش‌های جداگانه نشان دادند که برخی سویه‌ها تاب‌آوری را در برابر آسیب‌های القاشده به DNA بهبود بخشیدند؛ شاخصی که با مکانیسم‌های مرتبط با جوان‌سازی پیوند دارد، حتی علی‌رغم اینکه این ویژگی به‌صراحت در طول طراحی غربالگری نشده بود.

در مجموع، این همکاری الگوی ملموسی را برای زیست‌شناسی مبتنی بر هوش مصنوعی به تصویر می‌کشد. مدل‌های OpenAI به RetroBio کمک می‌کنند تا فضاهای وسیع طراحی پروتئین را بهتر و مؤثرتر جست‌وجو کند، نرخ موفقیت آزمایشگاهی را افزایش دهد و چرخه‌های تکرار را از چند سال به چند ماه کوتاه کند؛ امری که بیانگر نوعی از شتاب‌دهی علمی در دنیای واقعی است که نشان می‌دهد هوش مصنوعی چگونه به‌عنوان یک همکار در جریان‌های کاری پژوهشی عمل می‌کند.



## سیاست‌گذاری: شتاب‌دهی به کار علمی مبتنی بر هوش مصنوعی

OpenAI قصد دارد تمرکز خود را بر تقویت نوآوری علمی آمریکا با استفاده از هوش مصنوعی حفظ کند تا اطمینان حاصل کند که سال ۲۰۲۶، سال هوش مصنوعی و علم خواهد بود. OpenAI چهار رکن شامل مهارت‌آموزی هوش مصنوعی، دسترسی به داده‌ها، زیرساخت‌های مدرن هوش مصنوعی و اعطای دسترسی مقیاس‌پذیر به سیستم‌های پیشرو هوش مصنوعی به جامعه علمی را برای مدیریت پایدار توصیه می‌کند. OpenAI این چهار رکن را در طرح پیشنهادی اخیر خود به دفتر سیاست‌گذاری علم و فناوری کاخ سفید (OSTP) با جزئیات بیشتری شرح داده است.

**گسترش مهارت‌آموزی هوش مصنوعی** برای آماده‌سازی نیروی کار و دانشمندان آینده ایالات متحده. راه‌اندازی یک برنامه ملی نیروی کار هوش مصنوعی که از برنامه‌های درسی هوش مصنوعی K-12 پشتیبانی می‌کند، برنامه‌های مدرک و گواهینامه هوش مصنوعی را در کالج‌های محلی و دانشگاه‌ها گسترش می‌دهد و آموزش‌های کوتاه‌مدت برای نیروی انسانی در اواسط دوره شغلی ایجاد می‌کند. کنگره و سازمان‌ها می‌توانند تخصیص بودجه‌هایی را برای توسعه برنامه‌های درسی و آموزش مربیان در زمینه هوش مصنوعی و علم داده برای مدارس در هر ایالت مجاز کنند، در کنار تشکیل یک یگان ویژه مهارت‌های هوش مصنوعی یا صندوق اعطای کمک‌هزینه که کارگاه‌ها و آموزش‌های رایگان هوش مصنوعی را از طریق کتابخانه‌ها، مراکز کاریابی و مدارس به جوامع ارائه می‌دهد.

**در دسترس قرار دادن داده‌ها و گسترش مشارکت‌های پژوهشی** باز برای تسریع کشفیات. تحقیقات علمی با کمک هوش مصنوعی می‌تواند دهه‌ها اکتشاف را در چند سال فشرده کند، به‌خصوص زمانی که با دسترسی اکتشافی گسترده به ابزارهای جدید هوش مصنوعی و داده‌های جدید همراه شود. OpenAI از تأسیس منبع ملی پژوهش هوش مصنوعی (National AI Research Resource) به‌عنوان یک پلتفرم مشترک که دسترسی پژوهشگران دانشگاهی و غیرانتفاعی را به محاسبات در مقیاس بزرگ و مجموعه‌داده‌های با کیفیت بالا فراهم می‌کند، حمایت می‌کند و خواستار تصویب و اجرای سریع آن با بودجه و حکمرانی قدرتمند است. آژانس‌ها باید مجموعه‌داده‌های ارزشمند را شناسایی کرده و آن‌ها را در قالبی قابل‌خواندن توسط ماشین برای تحقیق و توسعه (R&D) هوش مصنوعی در دسترس قرار دهند؛ از جمله ایجاد یک کاتالوگ متمرکز داده‌های آموزش هوش مصنوعی، درحالی‌که حریم خصوصی و امنیت با حذف هویت برای داده‌های حساس و با پیش‌فرض دسترسی باز برای کاربردهای پژوهشی معتبر حفظ می‌شود.

مدرن‌سازی زیرساخت‌های هوش مصنوعی به‌عنوان یک دارایی استراتژیک ملی از جمله انرژی، محاسبات و تراشه‌ها. محاسبات ابزار حیاتی برای اکتشافات مبتنی بر هوش مصنوعی است و اطمینان از انجام محاسبات انبوه در خاک ایالات متحده به زیرساخت‌های مدرن و اضافات کافی به شبکه برق ایالات متحده بستگی دارد تا پیشرفت علمی آمریکا با کمبود انرژی محدود نشود. سیاست‌گذاری فدرال باید هاب‌های زیرساخت هوش مصنوعی را با استفاده از اختیارات قانون تراشه‌ها و علم (CHIPS and Science Act) تأسیس کند، مناطق نوآوری هوش مصنوعی را با پشتیبانی اولویت‌دار از دیتاسنترهای با ظرفیت بالا و سرمایه‌گذاری‌های مرتبط با انرژی تعیین کند و آن‌ها را با صدور مجوزهای تسهیل‌شده همراه سازد. مشارکت‌های دولتی-خصوصی می‌توانند به تأمین ظرفیت در سیستم‌های پیشرفته هوش مصنوعی برای پژوهشگرانی که از بودجه فدرال استفاده می‌کنند، در کنار اقدام برای تقویت زنجیره تأمین نیمه‌هادی‌های موردنیاز برای مکان‌یابی، ساخت و تولید تراشه‌های هوش مصنوعی در ایالات متحده، کمک کنند

ارائه دسترسی گسترده و مقیاس‌پذیر به سیستم‌های پیشگام هوش مصنوعی و تقویت اکوسیستم نوآوری. دولت فدرال باید یک برنامه تخصیص دسترسی ملی به هوش مصنوعی پیشگام (National Frontier AI Access Allocation) ایجاد کند که به پژوهشگران در سراسر دانشگاه‌ها، آزمایشگاه‌های ملی خود دولت و مؤسسات غیرانتفاعی، دسترسی به سیستم‌های هوش مصنوعی پیشرفته را در مقیاسی کافی برای آزمایش‌های مستمر، توسعه روش‌ها و اعتبارسنجی اعطا کند؛ دسترسی باید دارای پایگاه گسترده، دستیابی آسان و به‌گونه‌ای طراحی شود که از اکتشافات باز در کنار پروژه‌های پژوهشی تعریف‌شده حمایت کند و استفاده از هوش مصنوعی را به‌عنوان یک منبع پژوهشی ملی مشترک، مشابه زمان استفاده از تلسکوپ یا ساعات استفاده از ابرکامپیوترها در نظر بگیرد. برای بهبود کاربرد و کارایی این فناوری در اقتصاد، سیاست فدرال باید برنامه‌های پژوهش نوآوری کسب‌وکارهای کوچک (SBIR) را برای پوشش پذیرش هوش مصنوعی گسترش دهد و انکوباتورهای نوآوری منطقه‌ای ایجاد کند که متخصصان فناوری هوش مصنوعی را با نیازهای صنایع محلی پیوند دهند.