

[블로그](#) >

2023년: AI와 컴퓨팅이 획기적인 발전을 이루는 해

2023년 12월 22일 금요일

게시자: Google DeepMind 및 Google Research 수석 과학자 Jeff Dean, Google DeepMind CEO Demis Hassabis, Google Research, Technology & Society SVP James Manyika

올해는 인공지능(AI) 연구와 실제 적용 분야에서 놀라운 발전을 이룬 한 해였습니다.

진행 중인 연구가 AI를 더욱 발전시키면서 올해 1월에 발표된 "우리가 AI에 집중하는 이유(및 목적)"라는 제목으로 발표된 우리의 [관점을 되돌아보며 다음과 같이 언급했습니다](#).

우리는 유용하고 유익한 응용 프로그램을 개발 및 출시하고, 인간의 가치에 기반을 둔 윤리적 원칙을 적용하고, 연구, 경험, 사용자 및 더 넓은 커뮤니티를 통해 학습하면서 접근 방식을 발전시키는 데 있어 표준을 선도하고 설정하기 위해 최선을 다하고 있습니다.

또한 우리는 AI를 올바르게 구현하는 것(혁신하고 사람들과 사회에 광범위하게 접근 가능한 혜택을 제공하는 동시에 위험을 완화하는 것과 관련)이 우리와 연구원, 개발자, 사용자(개인, 기업 및 다른 조직), 정부, 규제 기관 및 시민.

우리는 대담하고 책임감 있게 개발하고 제공하는 데 중점을 두고 있는 AI 기반 혁신이 유용하고 매력적이며 모든 곳의 사람들의 삶을 지원하고 개선할 수 있는 잠재력을 가지고 있다고 확신합니다. 이것이 바로 우리를 강요하는 것입니다.

이번 연말 리뷰 게시물에서는 2023년 내내 이러한 문단을 안전하게 실행하기 위한 Google Research와 Google DeepMind의 노력 중 일부를 살펴보겠습니다.

제품 및 기술의 발전

올해는 생성적 AI가 몇 년 전만 해도 상상할 수 없을 정도의 창의성과 속도로 이미지, 음악, 스토리, 상상할 수 있는 모든 것에 대한 매력적인 대화를 만들어 세계의 관심을 사로잡은 해였습니다.

지난 2월, 우리는 창의적인 아이디어를 탐색하고 간단하게 설명하는 데 사용할 수 있는 도구인 [Bard를 처음 출시했습니다](#). 텍스트를 생성하고, 언어를 번역하고, 다양한 종류의 창의적인 콘텐츠를 작성하는 등의 작업을 수행할 수 있습니다.

바다의 거대한 거리는 우리 의식(의식)의 거리를 **1000배** 늘려줍니다.



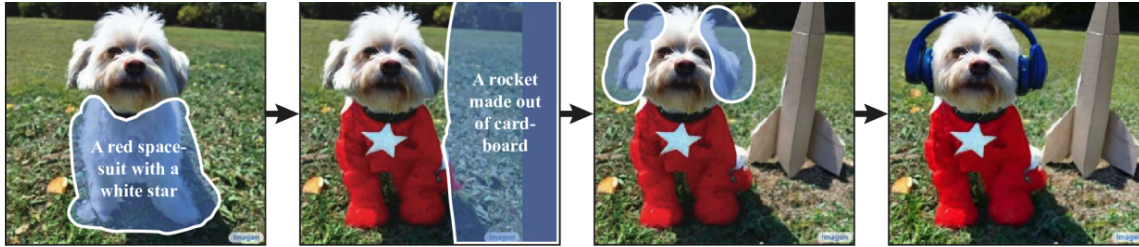
PaLM 2

PaLM 2를 다양한 목적에 맞게 미세 조정하고 지침을 조정함으로써 PaLM 2를 다음과 같은 수많은 Google 제품 및 기능에 통합할 수 있었습니다.

- 다국어 기능을 지원하는 Bard 업데이트입니다. Bard는 처음 출시된 이후 현재 40개 이상의 언어와 230개 이상의 국가 및 지역 에서 사용할 수 있으며 확장 프로그램을 통해 Bard는 Gmail, Google 지도, YouTube 등 매일 사용하는 Google 도구에서 관련 정보를 찾고 표시할 수 있습니다.
- SGE(검색 생성 경험)는 LLM을 사용하여 정보를 구성하는 방법과 사람들이 정보를 탐색하도록 돕는 방법을 재구성하여 핵심 검색 제품에 대한 보다 유동적인 대화형 상호 작용 모델을 만듭니다. 이 작업은 주로 정보 검색에 초점을 맞춘 검색 엔진 경험을 훨씬 더 많은 것(검색, 종합, 창의적인 생성 및 이전 검색의 지속 가능)으로 확장하는 동시에 사용자와 그들이 찾는 웹 콘텐츠 간의 연결 지점 역할을 계속합니다.
- AudioLM 및 MuLAN을 기반으로 하는 텍스트-음악 모델인 MusicLM 은 텍스트, 흥얼거림, 이미지 또는 비디오와 노래에 대한 음악 반주로 음악을 만들 수 있습니다.
- 사용자가 Google Workspace 및 Google Cloud를 사용할 때 도움을 제공하는 AI 기반 공동작업자인 Duet AI. 예를 들어 Google Workspace의 Duet AI는 사용자가 작성, 이미지 생성, 스프레드시트 분석, 이메일 및 채팅 메시지 초안 작성 및 요약, 회의 요약을 돕습니다. Google Cloud의 Duet AI는 사용자가 애플리케이션을 코딩, 배포, 확장, 모니터링하는 것은 물론 사이버 보안 위협을 식별하고 해결을 가속화하는 데 도움이 됩니다.
- 그리고 다른 많은 발전 .

작년 6월에는 텍스트-이미지 생성 모델 **Imagen**을 출시한 후 **Imagen Editor**를 출시했습니다. Imagen Editor는 영역 마스크와 자연어 프롬프트를 사용하여 생성 이미지를 대화형으로 편집하여 모델 출력을 훨씬 더 정확하게 제어할 수 있는 기능을 제공합니다.





올해 말에 우리는 좋은 조명, 프레임, 노출 및 선명도와 같은 품질에 대한 인간 선호도를 기반으로 전문적인 이미지 미학 모델을 통해 출력을 개선한 Imagen 2를 출시했습니다.

지난 10월, 우리는 사람들이 말하기 연습을 하고 언어 능력을 향상시키는 데 도움이 되는 기능을 출시했습니다. 이 기능을 가능하게 한 핵심 기술은 Google 번역 팀과 협력하여 개발한 Deep Aligner라는 새로운 딥 러닝 모델이었습니다. 이 단일 새로운 모델은 테스트된 모든 언어 쌍에서 정렬 품질을 획기적으로 향상시켜 HMM(Hidden Markov 모델) 기반 정렬 접근 방식에 비해 평균 정렬 오류율을 25%에서 5%로 줄였습니다.

In November, in partnership with [YouTube](#), we announced [Lyria](#), our most advanced AI music generation model to date. We released two experiments designed to open a new playground for creativity, DreamTrack and music AI tools, in concert with [YouTube's Principles for partnering with the music industry on AI technology](#).

Then in December, we launched [Gemini](#), our most capable and general AI model. Gemini was built to be multimodal from the ground up across text, audio, image and videos. Our initial family of Gemini models comes in three different sizes, Nano, Pro, and Ultra. Nano models are our smallest and most efficient models for powering on-device experiences in products like Pixel. The Pro model is highly-capable and best for scaling across a wide range of tasks. The Ultra model is our largest and most capable model for highly complex tasks.





In a [technical report](#) about [Gemini models](#), we showed that Gemini Ultra's performance exceeds current state-of-the-art results on 30 of the 32 widely-used academic benchmarks used in LLM research and development. With a score of 90.04%, Gemini Ultra was the first model to outperform human experts on [MMLU](#), and achieved a state-of-the-art score of 59.4% on the new [MMMU](#) benchmark.

Building on [AlphaCode](#), the first AI system to perform at the level of the median competitor in competitive programming, we [introduced AlphaCode 2](#) powered by a specialized version of Gemini. When evaluated on the same platform as the original AlphaCode, we found that AlphaCode 2 solved 1.7x more problems, and performed better than 85% of competition participants

At the same time, [Bard got its biggest upgrade](#) with its use of the Gemini Pro model, making it far more capable at things like understanding, summarizing, reasoning, coding, and planning. In six out of eight benchmarks, Gemini Pro outperformed GPT-3.5, including in MMLU, one of the key standards for measuring large AI models, and [GSM8K](#), which measures grade school math reasoning. Gemini Ultra will come to Bard early next year through Bard Advanced, a new cutting-edge AI experience.

Gemini Pro is also available on [Vertex AI](#), Google Cloud's end-to-end AI platform that empowers developers to build applications that can process information across text, code, images, and video. [Gemini Pro was also made available in AI Studio](#) in December.

To best illustrate some of Gemini's capabilities, we produced a [series of short videos](#) with explanations of how Gemini could:

- [Unlock insights in scientific literature](#)

- [Explain reasoning in math and physics](#)
- [Reason about user intent to generate bespoke experiences](#)

ML/AI Research

In addition to our advances in products and technologies, we've also made a number of important advancements in the broader fields of machine learning and AI research.

At the heart of the most advanced ML models is the Transformer model architecture, [developed by Google researchers in 2017](#). Originally developed for language, it has proven useful in domains as varied as [computer vision](#), [audio](#), [genomics](#), [protein folding](#), and more. This year, our work on [scaling vision transformers](#) demonstrated state-of-the-art results across a wide variety of vision tasks, and has also been useful in building [more capable robots](#).

Expanding the versatility of models requires the ability to perform higher-level and multi-step reasoning. This year, we approached this target following several research tracks. For example, [algorithmic prompting](#) is a new method that teaches language models reasoning by demonstrating a sequence of algorithmic steps, which the model can then apply in new contexts. This approach improves accuracy on one middle-school mathematics benchmark from 25.9% to 61.1%.



By providing algorithmic prompts, we can teach a model the rules of arithmetic via in-context learning.

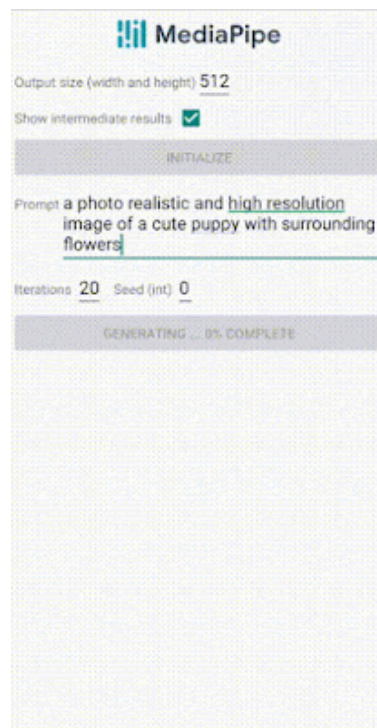
In the domain of visual question answering, in a collaboration with UC Berkeley researchers, we showed how we could [better answer complex visual questions](#) ("Is the carriage to the right of the horse?") by combining a visual model with a language model trained to answer visual questions by synthesizing a program to perform multi-step reasoning.

We are now using a [general model that understands many aspects of the software development life cycle](#) to automatically generate code review comments, respond to code review comments, make performance-improving suggestions for pieces of code (by

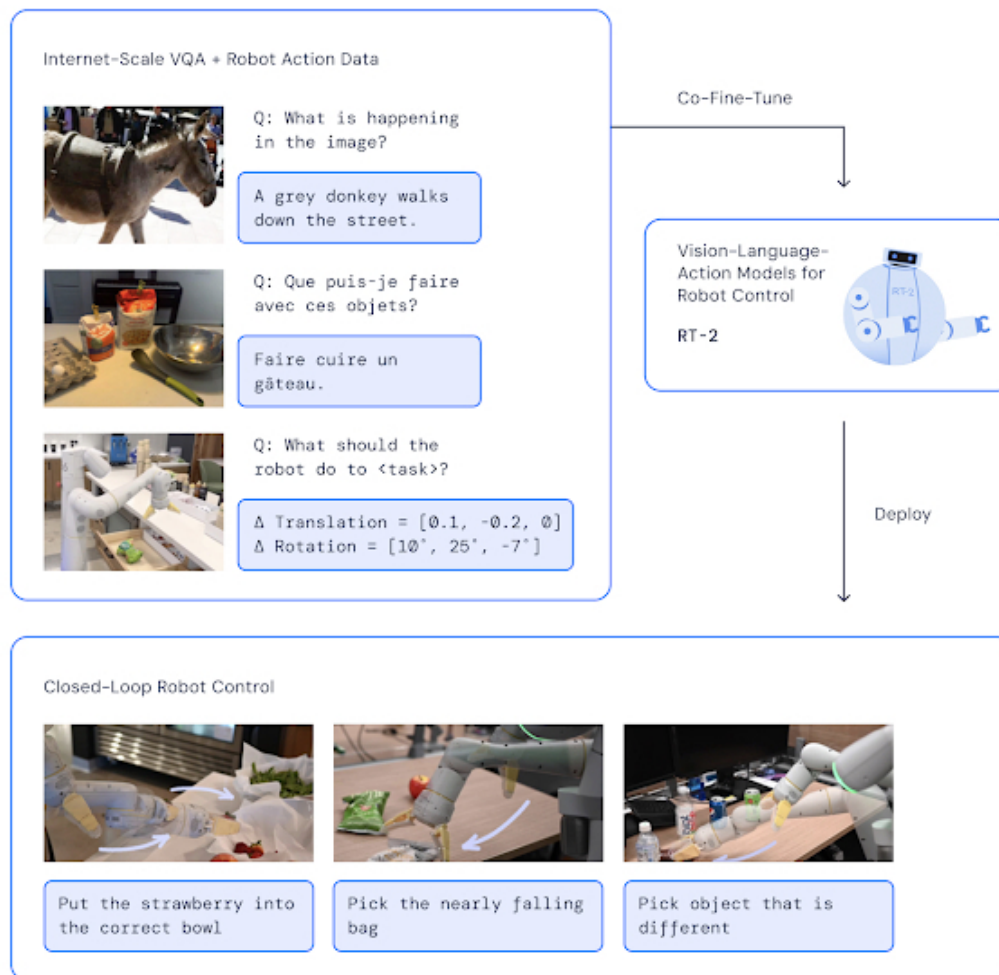
In a multi-year research collaboration with the Google Maps team, we were able to scale inverse reinforcement learning and apply it to the [world-scale problem of improving route suggestions](#) for over 1 billion users. Our work culminated in a 16–24% relative improvement in global route match rate, helping to ensure that routes are better aligned with user preferences.

We also continue to work on techniques to improve the inference performance of machine learning models. In work on [computationally-friendly approaches to pruning connections in neural networks](#), we were able to devise an approximation algorithm to the computationally intractable best-subset selection problem that is able to prune 70% of the edges from an image classification model and still retain almost all of the accuracy of the original.

In work on [accelerating on-device diffusion models](#), we were also able to apply a variety of optimizations to attention mechanisms, convolutional kernels, and fusion of operations to make it practical to run high quality image generation models on-device; for example, enabling “a photorealistic and high-resolution image of a cute puppy with surrounding flowers” to be generated in just 12 seconds on a smartphone.



models into **PaLM-E**, an embodied multi-modal model for robotics, and **Robot Transformer 2** (RT-2), a novel vision-language-action (VLA) model that **learns** from both web and robotics data, and translates this knowledge into generalized instructions for robotic control.



RT-2 architecture and training: We co-fine-tune a pre-trained vision-language model on robotics and web data. The resulting model takes in robot camera images and directly predicts actions for a robot to perform.

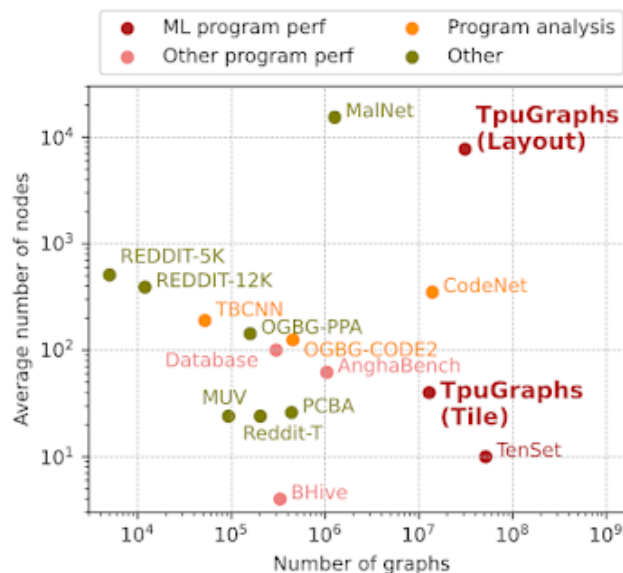
Furthermore, we showed how **language can also be used to control the gait of quadrupedal robots** and explored the **use of language to help formulate more explicit reward functions** to bridge the gap between human language and robotic actions. Then, in **Barkour** we benchmarked the agility limits of quadrupedal robots.

Algorithms & optimization

optimization, and privacy.

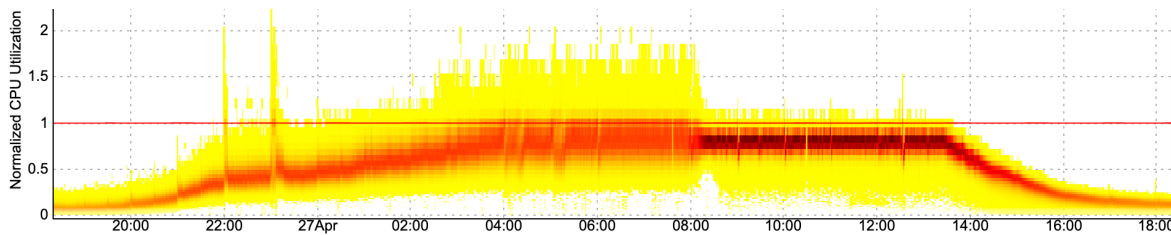
We introduced [AlphaDev](#), an AI system that uses reinforcement learning to discover enhanced computer science algorithms. AlphaDev uncovered a faster algorithm for sorting, a method for ordering data, which led to improvements in the LLVM libc++ sorting library that were up to 70% faster for shorter sequences and about 1.7% faster for sequences exceeding 250,000 elements.

We developed a novel model to [predict the properties of large graphs](#), enabling estimation of performance for large programs. We released a new dataset, [TPUGraphs](#), to accelerate [open research in this area](#), and showed how we can use [modern ML to improve ML efficiency](#).



The TPUGraphs dataset has 44 million graphs for ML program optimization.

We developed a new [load balancing](#) algorithm for distributing queries to a server, called [Prequal](#), which minimizes a combination of requests-in-flight and estimates the latency. Deployments across several systems have saved CPU, latency, and RAM significantly. We also designed a new [analysis framework](#) for the classical caching problem with capacity reservations.



Heatmaps of normalized CPU usage transitioning to [Prequal](#) at 08:00.

We improved state-of-the-art in clustering and [graph algorithms](#) by developing new techniques for [computing minimum-cut](#), [approximating correlation clustering](#), and [massively parallel graph clustering](#). Additionally, we introduced [TeraHAC](#), a novel hierarchical clustering algorithm for trillion-edge graphs, designed a [text clustering algorithm](#) for better scalability while maintaining quality, and designed the most efficient [algorithm for approximating the Chamfer Distance](#), the standard similarity function for multi-embedding models, offering >50× speedups over highly-optimized exact algorithms and scaling to billions of points.

We continued optimizing Google's large embedding models (LEMs), which power many of our core products and recommender systems. Some new techniques include [Unified Embedding](#) for battle-tested feature representations in web-scale ML systems and [Sequential Attention](#), which uses attention mechanisms to discover high-quality sparse model architectures during training.

Beyond auto-bidding systems, we also studied auction design in other complex settings, such as [buy-many mechanisms](#), [auctions for heterogeneous bidders](#), [contract designs](#), and innovated [robust online bidding algorithms](#). Motivated by the application of generative AI in collaborative creation (e.g., joint ad for advertisers), we proposed a [novel token auction model](#) where LLMs bid for influence in the collaborative AI creation. Finally, we show how to [mitigate personalization effects in experimental design](#), which, for example, may cause recommendations to drift over time.

The Chrome Privacy Sandbox, a multi-year collaboration between Google Research and Chrome, has publicly launched several APIs, including for [Protected Audience](#), [Topics](#), and [Attribution Reporting](#). This is a major step in protecting user privacy while supporting the open and free web ecosystem. These efforts have been facilitated by fundamental research on [re-identification risk](#), [private streaming computation](#), [optimization](#) of privacy caps and budgets, [hierarchical aggregation](#), and training models with [label privacy](#).

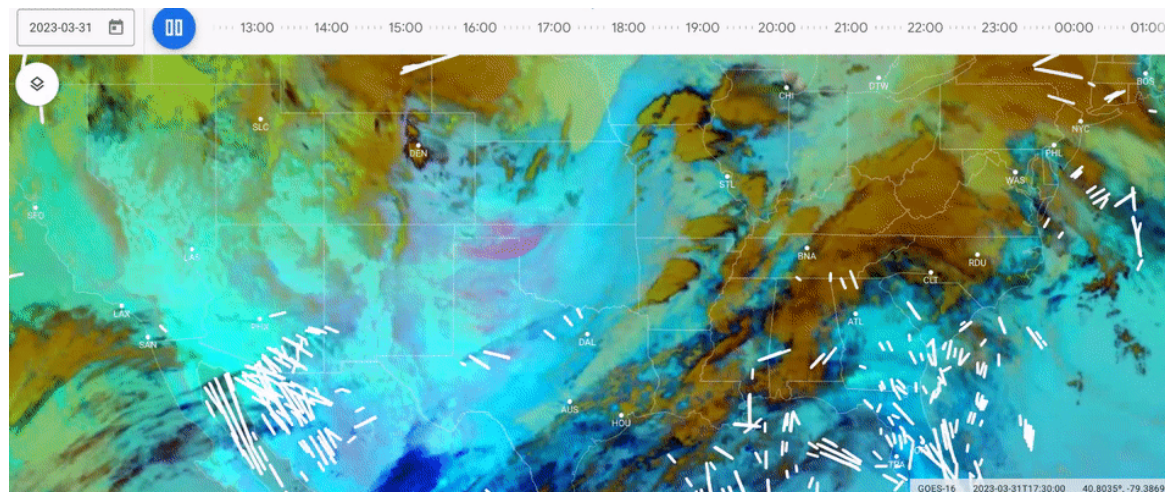
Science and society

and lead to major advances in diverse areas including bioengineering, [materials science](#), [weather prediction](#), [climate forecasting](#), [neuroscience](#), [genetic medicine](#), and [healthcare](#).

Sustainability and climate change

In [Project Green Light](#), we partnered with 13 cities around the world to help improve traffic flow at intersections and reduce stop-and-go emissions. Early numbers from these partnerships indicate a potential for up to 30% reduction in stops and up to 10% reduction in emissions.

In our [contrails work](#), we analyzed large-scale weather data, historical satellite images, and past flights. We [trained an AI model](#) to predict where contrails form and reroute airplanes accordingly. In partnership with American Airlines and Breakthrough Energy, we used this system to demonstrate contrail reduction by 54%.



Contrails detected over the United States using AI and GOES-16 satellite imagery.

We are also developing novel technology-driven approaches to [help communities with the effects of climate change](#). For example, we have [expanded our flood forecasting coverage to 80 countries](#), which directly impacts more than 460 million people. We have initiated a [number of research efforts](#) to help mitigate the increasing danger of wildfires, including [real-time tracking of wildfire boundaries](#) using satellite imagery, and work that [improves emergency evacuation plans](#) for communities at risk to rapidly-spreading wildfires. Our [partnership](#) with American Forests puts data from our [Tree Canopy](#) project to work in their [Tree Equity Score](#) platform, helping communities identify and address unequal access to trees.

outperform traditional numerical weather simulations up to twenty-four hours. In the area of medium-term, global weather forecasting, our work on [GraphCast](#) showed significantly better prediction accuracy for up to 10 days compared to [HRES](#), the most accurate operational deterministic forecast, produced by the [European Centre for Medium-Range Weather Forecasts](#) (ECMWF). In collaboration with ECMWF, we released [WeatherBench-2](#), a benchmark for evaluating the accuracy of weather forecasts in a common framework.

A selection of GraphCast's predictions rolling across 10 days



A selection of GraphCast's predictions rolling across 10 days showing specific humidity at 700 hectopascals (about 3 km above surface), surface temperature, and surface wind speed.



Health and the life sciences

The potential of AI to dramatically improve processes in healthcare is significant. Our initial [Med-PaLM](#) model was the first model capable of achieving a passing score on the U.S. medical licensing exam. Our more recent [Med-PaLM 2 model](#) improved by a further 19%, achieving an expert-level accuracy of 86.5%. These [Med-PaLM models](#) are language-based, enable clinicians to ask questions and have a dialogue about complex medical conditions, and are [available](#) to healthcare organizations as part of [MedLM](#) through Google Cloud.

In the same way our general language models are evolving to handle multiple modalities, we have recently shown research on a [multimodal version of Med-PaLM](#) capable of interpreting medical images, textual data, and other modalities, [describing a path](#) for how we can realize the exciting potential of AI models to help advance real-world clinical care.



Med-PaLM

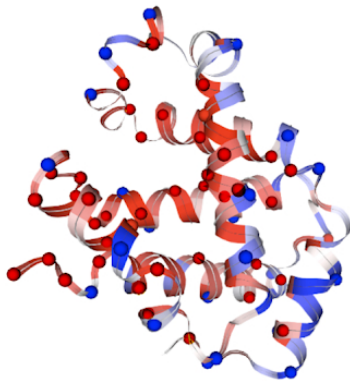
Med-PaLM M is a large multimodal generative model that flexibly encodes and interprets biomedical data including clinical language, imaging, and genomics with the same model weights.



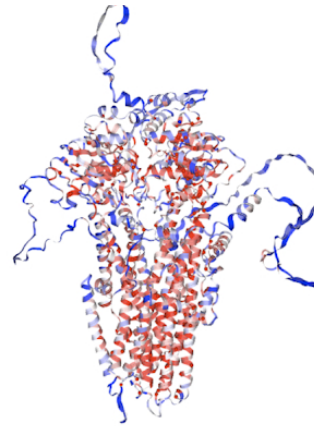
We have also been working on [how best to harness AI models in clinical workflows](#). We have shown that [coupling deep learning with interpretability methods](#) can yield new insights for clinicians. We have also shown that self-supervised learning, with careful consideration of privacy, safety, fairness and ethics, [can reduce the amount of de-identified data needed](#) to train clinically relevant medical imaging models by 3×–100×, reducing the barriers to adoption of models in real clinical settings. We also released an [open source mobile data collection platform](#) for people with chronic disease to provide tools to the community to build their own studies.

AI systems can also discover completely new signals and biomarkers in existing forms of medical data. In work on [novel biomarkers discovered in retinal images](#), we demonstrated that a number of systemic biomarkers spanning several organ systems (e.g., kidney, blood, liver) can be predicted from external eye photos. In other work, we showed that combining [retinal images and genomic information](#) helps identify some underlying factors of aging.

In the genomics space, we worked with 119 scientists across 60 institutions to create a [new map of the human genome](#), or pangenome. This more equitable pangenome better represents the genomic diversity of global populations. Building on our ground-breaking [AlphaFold](#) work, our work on [AlphaMissense](#) this year provides a catalog of predictions for 89% of all 71 million possible [missense variants](#) as either likely pathogenic or likely benign.



Hemoglobin subunit beta (HBB)



Cystic fibrosis transmembrane conductance regulator (CFTR)

*Examples of AlphaMissense predictions overlaid on AlphaFold predicted structures (red – predicted as pathogenic; blue – predicted as benign; grey – uncertain). Red dots represent known pathogenic missense variants, blue dots represent known benign variants. **Left:** HBB protein. Variants in this protein can cause sickle cell anaemia. **Right:** CFTR protein. Variants in this protein can cause cystic fibrosis.*

We also shared [an update](#) on progress towards the next generation of AlphaFold. Our latest model can now generate predictions for nearly all molecules in the [Protein Data Bank](#) (PDB), frequently reaching atomic accuracy. This unlocks new understanding and significantly improves accuracy in multiple key biomolecule classes, including ligands (small molecules), proteins, nucleic acids (DNA and RNA), and those containing post-translational modifications (PTMs).

On the neuroscience front, we [announced a new collaboration](#) with Harvard, Princeton, the NIH, and others to map an entire mouse brain at synaptic resolution, beginning with a first phase that will focus on the [hippocampal formation](#) – the area of the brain responsible for memory formation, spatial navigation, and other important functions.

Quantum computing

Quantum computers have the potential to solve big, real-world problems across science and industry. But to realize that potential, they must be significantly larger than they are today, and they must reliably perform tasks that cannot be performed on classical computers.

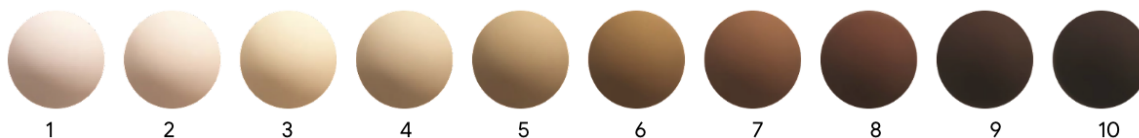
This year, we took an important step towards the development of a large-scale, useful quantum computer. Our breakthrough is the first demonstration of [quantum error correction](#), showing that it's possible to reduce errors while also increasing the number of qubits. To enable real-world applications, these qubit building blocks must perform more reliably, lowering the error rate from ~ 1 in 10^3 typically seen today, to ~ 1 in 10^8 .

Design for Responsibility

Generative AI is having a transformative impact in a wide range of fields including healthcare, education, security, energy, transportation, manufacturing, and entertainment. Given these advances, the importance of designing technologies consistent with our [AI Principles](#) remains a top priority. We also recently published case studies of [emerging practices in society-centered AI](#). And in our annual [AI Principles Progress Update](#), we offer details on how our Responsible AI research is integrated into products and risk management processes.

Proactive design for Responsible AI begins with identifying and documenting potential harms. For example, we recently [introduced](#) a [three-layered](#) context-based framework for comprehensively evaluating the social and ethical risks of AI systems. During model design, harms can be mitigated with the use of [responsible datasets](#).

We are [partnering with Howard University](#) to build high quality African-American English (AAE) datasets to improve our products and make them work well for more people. Our research on [globally inclusive cultural representation](#) and our publication of the [Monk Skin Tone scale](#) furthers our commitments to equitable representation of all people. The insights we gain and techniques we develop not only help us improve our own models, they also power [large-scale studies of representation in popular media](#) to inform and inspire more inclusive content creation around the world.



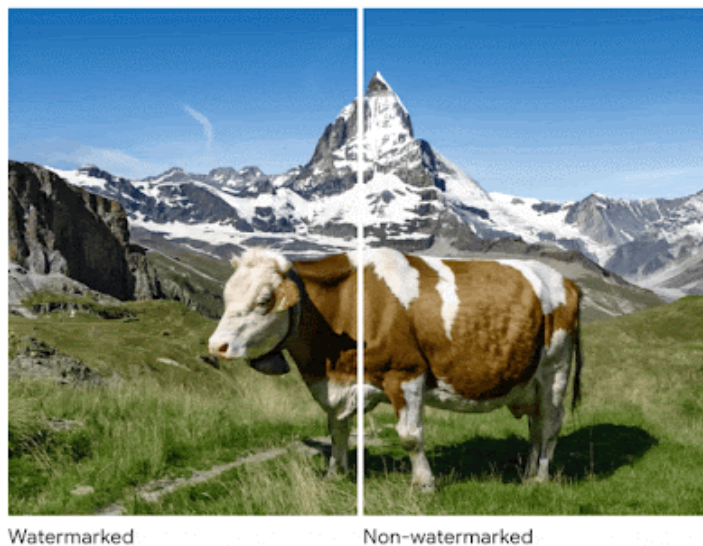
Monk Skin Tone (MST) Scale. See more at skintone.google.

With advances in generative image models, [fair and inclusive representation of people](#) remains a top priority. In the development pipeline, we are working to [amplify underrepresented voices and to better integrate social context knowledge](#). We proactively address potential harms and bias using [classifiers and filters](#), [careful dataset analysis](#), and in-model mitigations such as fine-tuning, [reasoning](#), [few-shot prompting](#), [data augmentation](#) and [controlled decoding](#), and our research showed that generative AI enables [higher quality safety classifiers](#) to be developed with far less data. We also released [a powerful way to better tune models with less data](#) giving developers more control of responsibility challenges in generative AI.

[Image classifiers](#), we found that we can identify mislabeled training examples. This makes it possible to reduce the noise in training data, leading to significant improvements in model accuracy.

We initiated several efforts to improve safety and transparency about online content. For example, we introduced [SynthID](#), a tool for watermarking and identifying AI-generated images. SynthID is imperceptible to the human eye, doesn't compromise image quality, and allows the watermark to remain detectable, even after modifications like adding filters, changing colors, and saving with various lossy compression schemes.

We also launched [About This Image](#) to help people assess the credibility of images, showing information like an image's history, how it's used on other pages, and available metadata about an image. And we [explored safety methods](#) that have been developed in other fields, learning from established situations where there is low-risk tolerance.



SynthID generates an imperceptible digital watermark for AI-generated images.

Privacy remains an essential aspect of our commitment to Responsible AI. We continued improving our state-of-the-art privacy preserving learning algorithm [DP-FTRL](#), developed the DP-Alternating Minimization algorithm ([DP-AM](#)) to enable personalized recommendations with rigorous privacy protection, and defined a new [general paradigm](#) to reduce the privacy costs for many aggregation and learning tasks. We also proposed a scheme for [auditing differentially private machine learning systems](#).

are [useful for a range of downstream tasks](#). We [proposed](#) a new algorithm for DL training of large embedding models that provides efficient training on TPUs without compromising accuracy.

We also teamed up with a broad group of academic and industrial researchers to organize the [first Machine Unlearning Challenge](#) to address the scenario in which training images are forgotten to protect the privacy or rights of individuals. We shared a mechanism for [extractable memorization](#), and [participatory systems](#) that give users more control over their sensitive data.

We continued to expand the world's largest corpus of atypical speech recordings to >1M utterances in [Project Euphonia](#), which enabled us to train a [Universal Speech Model](#) to [better recognize atypical speech by 37%](#) on real-world benchmarks.

We also built an [audiobook recommendation system](#) for students with reading disabilities such as dyslexia.

Adversarial testing

Our work in adversarial testing [engaged community voices](#) from historically marginalized communities. We partnered with groups such as the [Equitable AI Research Round Table](#) (EARR) to ensure we represent the diverse communities who use our models and [engage with external users](#) to identify potential harms in generative model outputs.

We [established a dedicated Google AI Red Team](#) focused on testing AI models and products for security, privacy, and abuse risks. We showed that attacks such as “[poisoning](#)” or [adversarial examples](#) can be applied to production models and surface additional risks such as memorization in both [image](#) and [text generative models](#). We also demonstrated that defending against such attacks can be challenging, as merely applying defenses can cause other [security and privacy leakages](#). We also introduced model evaluation for [extreme risks](#), such as offensive cyber capabilities or strong manipulation skills.

Democratizing AI through tools and education

As we advance the state-of-the-art in ML and AI, we also want to ensure people can understand and apply AI to specific problems. We released [MakerSuite](#) (now [Google AI Studio](#)), a web-based tool that enables AI developers to quickly iterate and build lightweight AI-powered apps. To help AI engineers better understand and debug AI, we released [LIT 1.0](#), a state-of-the-art, open-source debugger for machine learning models.



[code assistance](#) to all users at no cost — making Colab an even more helpful and integrated experience in data and ML workflows.



One of the most used features is “Explain error” — whenever the user encounters an execution error in Colab, the code assistance model provides an explanation along with a potential fix.

To ensure AI produces accurate knowledge when put to use, we also recently introduced [FunSearch](#), a new approach that generates verifiably true knowledge in mathematical sciences using evolutionary methods and large language models.

For AI engineers and product designers, we’re updating the [People + AI Guidebook](#) with generative AI best practices, and we continue to design [AI Explorables](#), which includes [how and why models sometimes make incorrect predictions confidently](#).

Community engagement

We continue to advance the fields of AI and computer science by publishing much of our work and participating in and organizing conferences. We have published more than 500 papers so far this year, and have strong presences at conferences like ICML (see the [Google Research](#) and [Google DeepMind](#) posts), ICLR ([Google Research](#), [Google DeepMind](#)), NeurIPS ([Google Research](#), [Google DeepMind](#)), ICCV, CVPR, ACL, CHI, and [Interspeech](#). We are also working to support researchers around the world, participating in events like the [Deep Learning Indaba](#), [Khipu](#), supporting [PhD Fellowships in Latin America](#), and more. We also worked with partners from 33 academic labs to pool data from 22

Google has spearheaded an industry-wide effort to develop [AI safety benchmarks](#) under the [MLCommons](#) standards organization with participation from several major players in the generative AI space including OpenAI, Anthropic, Microsoft, Meta, Hugging Face, and more. Along with others in the industry we also [co-founded](#) the [Frontier Model Forum](#) (FMF), which is focused on ensuring safe and responsible development of frontier AI models. With our FMF partners and other philanthropic organizations, we launched a \$10 million [AI Safety Fund](#) to advance research into the ongoing development of the tools for society to effectively test and evaluate the most capable AI models.

In close partnership with [Google.org](#), we [worked with the United Nations](#) to build the [UN Data Commons for the Sustainable Development Goals](#), a tool that tracks metrics across the 17 [Sustainable Development Goals](#), and [supported projects](#) from NGOs, academic institutions, and social enterprises on [using AI to accelerate progress on the SDGs](#).

The items highlighted in this post are a small fraction of the research work we have done throughout the last year. Find out more at the [Google Research](#) and [Google DeepMind](#) blogs, and our [list of publications](#).



Future vision

As multimodal models become even more capable, they will empower people to make incredible progress in areas from science to education to entirely new areas of knowledge.

발전은 빠른 속도로 계속되고, 해가 갈수록 우리의 제품과 연구가 발전함에 따라 사람들은 AI에 대한 더욱 흥미롭고 창의적인 용도를 발견하게 될 것입니다.

[우리가 AI에 중점을 두는 이유\(및 목표\)](#)에서 설명한 대로 우리가 시작한 올해 검토를 마무리합니다.

대담하고 책임감 있게 추구한다면 AI가 모든 곳의 사람들의 삶을 변화시키는 기반 기술이 될 수 있다고 믿습니다. 이것이 바로 우리를 흥분시키는 것입니다!

이번 연도 리뷰는 [Google Research 블로그](#)와 [Google DeepMind 블로그](#)에 교차 게시됩니다.



이전 게시물

2023년 12월 19일

VideoPoet: 제로샷 비디오 생성을 위한 대규모 언어 모델



2023년 12월 19일

시뮬레이션을 통해 이벤트 후 교통 흐름 경로를 밝힙니다.



2023년 12월 15일

기계 학습을 위한 기계 학습의 발전



Google

은둔

자귀

구글 소개

구글 제품