≡ **Google**   The Keyword

100 things we announced at I/O 2024

May 15, 2024 · 11 min read

Share

Phew — it's been a busy couple of days.

Molly McHugh-Johnson
Contributor, The Keyword

📄✨ **Read AI-generated summary** ⌄

Listen to article 19 minutes

A lot happened at I/O 2024! Whether you were most into the latest Gemini app updates, felt especially excited about what's coming for developers or can't wait to try the latest generative AI tools, there was something for just about everyone. Don't believe us? Below, we rounded up 100 things we announced over the last two days.

AI moments and model momentum

1. We introduced Gemini 1.5 Flash: a lighter-weight model that's designed to be fast and efficient to serve at scale. 1.5 Flash is the fastest Gemini model served in the API.

2. We've significantly improved 1.5 Pro, our best model for general performance across a wide range of tasks.

3. Both 1.5 Pro and 1.5 Flash are available in public preview with a 1 million token context window on Google AI Studio and Vertex AI.

4. 1.5 Pro is also available with a 2 million token context window to developers via waitlist in Google AI Studio and Vertex AI.

Context lengths of leading foundation models compared with Gemini 1.5's 2 million token capability.

5. We shared Project Astra: our vision for the future of AI assistants.

6. We announced Trillium, the sixth generation of our custom AI accelerator, the Tensor Processing Unit (TPU). It is the most performant TPU to date.

7. Compared to TPU v5e, Trillium TPUs achieve a 4.7x increase in peak compute performance per chip.

8. They're also our most sustainable generation: Trillium TPUs are over 67% more energy-efficient compared to TPU v5e.

9. And we demoed an early prototype of Audio Overviews for NotebookLM, which uses a collection of uploaded materials to create a verbal discussion personalized for the user.

10. We announced that Grounding with Google Search — a tool that connects the Gemini model with world knowledge, a wide possible range of topics or up-to-date information on the internet — is now generally available on Vertex AI.

11. We added audio understanding in the Gemini API and AI Studio, so Gemini 1.5 Pro can now reason across image and audio for videos uploaded in AI Studio.

12. Starting with Pixel, applications using Gemini Nano with Multimodality will be able to understand the world the way people do — not just through text input but also through sight, sound and spoken language.

Generative media models and Labs experiments

13. We announced Imagen 3, our highest-quality image generation model yet.

14. Imagen 3 understands natural language and intent behind your prompts and incorporates small details from longer prompts. This helps it generate an incredible level of detail, producing photorealistic, lifelike images with far fewer distracting visual artifacts than our prior models.

15. Imagen 3 is also our best model yet for rendering text — a challenge for image generation models.

16. We rolled out Imagen 3 to Trusted Testers in ImageFX and you can sign up to join the waitlist.

17. Imagen 3 will also be coming to Vertex AI this summer.

18. Then we announced Veo, our most capable video generation model yet. It generates high-quality 1080p resolution videos that can go beyond a minute, in a wide range of cinematic and visual styles.

19. We'll also bring some of Veo's capabilities to YouTube Shorts and other products in the future.

20. We showed off what Veo can help artists do by collaborating with filmmakers — including Donald Glover, who experimented with Veo for a film project.

21. We highlighted Music AI Sandbox, a suite of music AI tools that allow people to create new instrumental sections from scratch, transfer styles between trackers and much more. You can find some brand new songs from these collaborations — including one from Wyclef Jean and another from Marc Rebillet — on YouTube now.

22. And be sure to check out *Infinite Wonderland*, an experience where artists and Google creatives experimented together to fine-tune an AI model to endlessly reimagine the visual world of the novel "Alice's Adventures in Wonderland." Readers of *Infinite Wonderland* can generate seemingly infinite images for each one of the 1,200 sentences in the book based on each artist's respective style.

23. We announced VideoFX, our newest experimental tool that uses Google DeepMind's generative video model, Veo, and lets you turn an idea into a video clip.

24. It also comes with a Storyboard mode that lets you iterate scene by scene and add music to your final video.

25. We added in more editorial controls to ImageFX — a top feature request from the community — so you can add, remove or change elements by simply brushing over your image.

26. ImageFX will also use Imagen 3 to unlock more photorealism with richer details and fewer visual artifacts and more accurate text rendering.
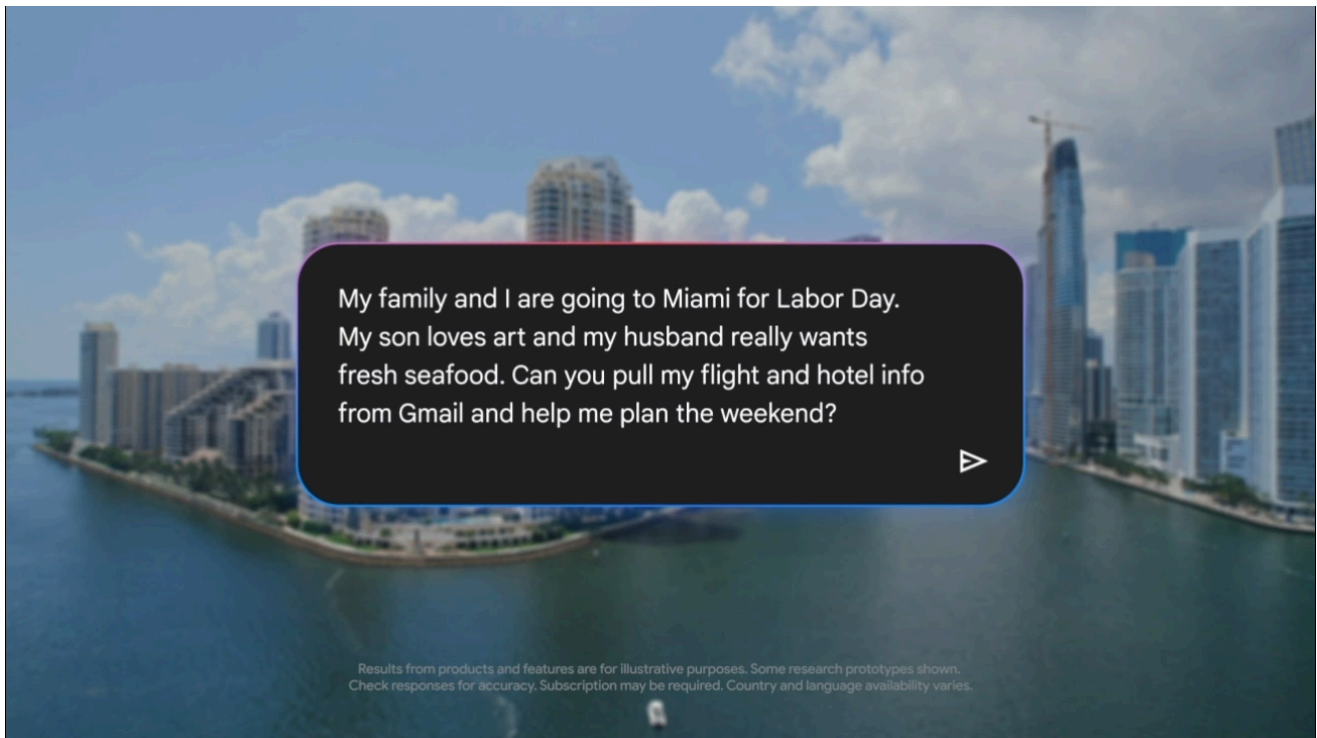
27. MusicFX has a new feature called "DJ Mode" that helps you mix beats by combining genres and instruments, using the power of generative AI to bring music stories to life.

28. As of this week, ImageFX and MusicFX are now available in over 100 countries through Labs.

New ways to get more done with the Gemini app

29. We're bringing Gemini 1.5 Pro, our cutting edge model, to Gemini Advanced subscribers — which means Gemini Advanced now has a 1 million token context window and can do things like make sense of 1,500-page PDFs.

30. This *also* means Gemini Advanced now has the largest context window of any commercially available chatbot in the world.

31. We added the ability to upload files via Google Drive or directly from your device right into Gemini Advanced.

32. Soon, Gemini Advanced will help you analyze your data to quickly uncover insights and build charts from uploaded data files like spreadsheets.

33. Great news for travelers: Gemini Advanced has a new planning feature that goes beyond a list of suggested activities and will actually create a custom itinerary just for you.

34. Then there's Gemini Live for Gemini Advanced subscribers, a new, mobile-first conversational experience that uses state-of-the-art speech technology to help you have more natural, intuitive spoken conversations with Gemini.

35. Gemini Live lets you choose from 10 natural-sounding voices it can respond to you with; plus, you can speak at your own pace or interrupt mid-response with clarifying questions.

36. Gemini in Google Messages now lets you chat with Gemini in the same app where you message your friends.

37. Gemini Advanced subscribers will soon be able to create Gems, customized versions of Gemini designed for whatever you dream up. Simply describe what you want your Gem to do and how you want it to respond and Gemini will take those instructions and create a Gem for your specific needs.

38. And look out for more Google tools being connected to Gemini, including Google Calendar, Tasks, Keep and Clock.

Updates that make Search do the work for you

39. We're using a new Gemini model customized for Google Search to bring together Gemini's advanced capabilities — including multi-step reasoning, planning and multimodality — with our best-in-class Search systems.

40. AI Overviews in Search are rolling out to everyone in the U.S. beginning this week with more countries coming soon.

41. And multi-step reasoning capabilities are coming soon to AI Overviews in Search Labs for English queries in the U.S. So rather than breaking your question into multiple searches, you can ask complex

questions like "find the best yoga or pilates studios in Boston and show details on their intro offers and walking time from Beacon Hill."

42. Soon, you'll be able to adjust your AI Overview with options to simplify the language or break it down in more detail, when you're new to a topic or trying to get to the heart of a subject.

43. Search is also getting new planning capabilities. For example, meal and trip planning with customization will launch later this year in Search Labs, followed soon by more categories like parties and fitness.

44. Thanks to advancements in video understanding, you now have the ability to ask questions with a video. Search can take a complex visual question and figure it out for you, then explain next steps and offer resources with an AI Overview.

45. And soon, generative AI in Search will also create an AI-organized results page when you're searching for fresh ideas. These AI-organized search result pages will be available when you're searching for categories like dining, recipes, movies, music, books, hotels, shopping and more.

Help from Gemini models in Workspace and Photos

46. Gemini 1.5 Pro is now available in the side panel in Gmail, Docs, Drive, Slides and Sheets via Workspace Labs — and it's rolling out to our Gemini for Workspace customers and Google One AI Premium subscribers next month.
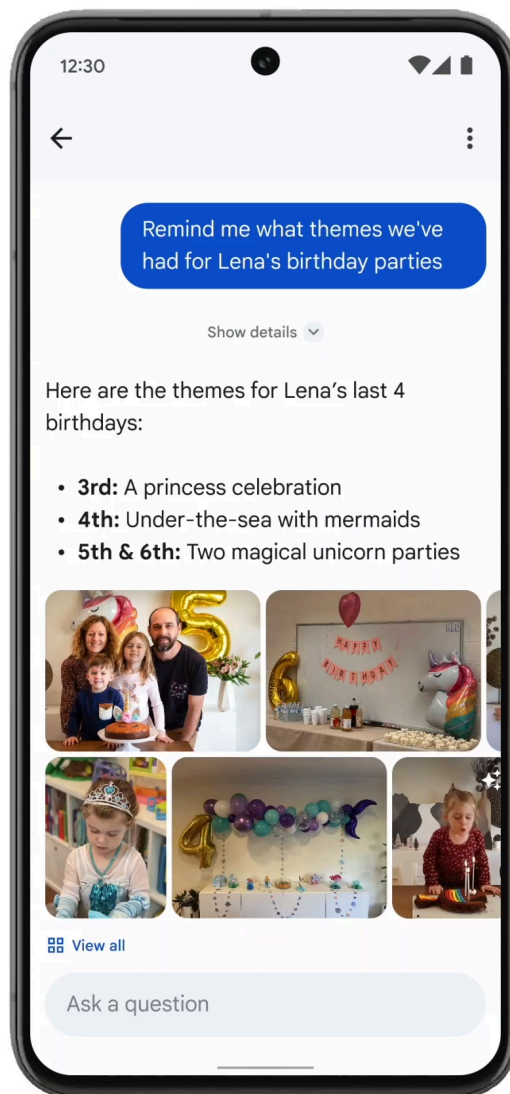
47. You'll be able to use Gmail's side panel to summarize emails to get the most important details and action items.

48. In addition to summaries, Gmail's mobile app will soon use Gemini for two other new features: Contextual Smart Reply and Gmail Q&A.

49. In the coming weeks, Help me write in Gmail and Docs will support Spanish and Portuguese.

50. Later this year in Labs, you can even ask Gemini to automatically organize email attachments in Drive, generate a sheet with the data and then analyze it with Data Q&A.

51. A new experimental feature in Google Photos called Ask Photos makes it even easier to look for specific memories or recall information included in your gallery. The feature uses Gemini models, and it's rolling out over the coming months.

UI subject to change. Sequences simulated. Check responses for accuracy.

52. You can also use Ask Photos to create a highlight gallery from a recent trip, and it will even write personalized captions for you to share on social media.

Android advancements

53. Starting with Pixel later this year, Gemini Nano — Android's built-in, on-device foundation model — will have multimodal capabilities. Beyond just processing text input, your Pixel phone will also be able to understand more information in context like sights, sounds and spoken language.

54. Talkback, an accessibility feature for Android devices that helps blind and low-vision people use touch and spoken feedback to better interact with their devices, is being improved thanks to Gemini Nano with Multimodality.

55. A new, opt-in scam protection feature that will use Gemini Nano's on-device AI to help detect scam phone calls in a privacy preserving way. Look out for more details later this year.

56. We announced that Circle to Search is currently available on more than 100 million Android devices, and we're on track to double that by the end of the year.
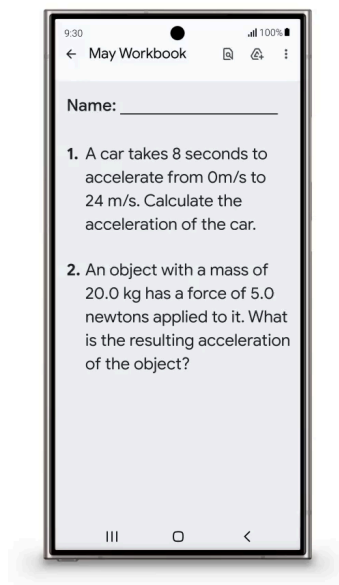
57. Soon, you'll be able to use Gemini on Android to create and drag and drop generated images into Gmail, Google Messages and more, or ask about the YouTube video you're viewing.

58. If you have Gemini Advanced, you'll also have the option to "Ask this PDF" to get an answer quickly without having to scroll through multiple pages.

59. Students can now use Circle to Search for homework help directly from select Android phones and tablets. This feature is powered by LearnLM — our new family of models based on Gemini, fine-tuned for learning.

60. Later this year, Circle to Search will be able to solve even more complex problems involving symbolic formulas, diagrams, graphs and more.

61. Oh, and we introduced the second beta of Android 15.

62. Theft Detection Lock uses powerful Google AI to sense if your device has been snatched and quickly lock down your information on your phone.

63. Private space is coming to Android 15, which lets you choose apps to keep secure inside a separate space that requires an extra layer of authentication to open.

64. And if a separate lock screen isn't enough for your private spaces, you can hide the existence of it altogether.

65. Later this year, Google Play Protect will use on-device AI to help spot apps that attempt to hide their actions to engage in fraud or phishing.

66. We're bringing an updated messaging experience to Japan with RCS in Google Messages.

67. Soon in the U.S., you'll be able to create a digital version of passes that just contain text. Simply take a photo of a pass (like an insurance card or event ticket) and easily add it to your Google Wallet for quick access.

68. We showed off how augmented reality content will be available directly in Google Maps, laying the foundation for an extended reality (XR) platform we're building in collaboration with Samsung and Qualcomm for the Android ecosystem.

69. You can now catch up on episodes of your favorite shows on Max and Peacock or start a game of Angry Birds on select cars with Google built-in.

70. We are also bringing Google Cast to cars with Android Automotive OS, starting with Rivian in the coming months, so you can easily cast video content from your phone to the car.

71. Later this year, battery life optimizations are coming to watches with Wear OS 5. For example, running an outdoor marathon will consume up to 20% less power when compared to watches with Wear OS 4.

72. Wear OS 5 will also give fitness apps the option to support more data types like ground contact time, stride length and vertical oscillation.

73. It's now easier to pick what to watch on Google TV and other Android TV OS devices with personalized AI-generated descriptions, thanks to our Gemini model.

74. These AI-generated descriptions will also fill in missing or untranslated descriptions for movies and shows.

75. Here's a fun stat: Since launch, people have made over 1 billion Fast Pair connections.

76. Later this month, you'll be able to use Fast Pair to connect and find items like your keys, wallet or luggage in the Find My Device app with Bluetooth tracker tags from Chipolo and PebbleBee (with more partners to come).

Developments for developers

77. You can join the Gemini API Developer Competition and be a part of discovering the most helpful and groundbreaking AI apps. The prize: an electrically retrofitted custom 1981 DeLorean.

78. We introduced PaliGemma, our first vision-language open model optimized for visual Q&A and image captioning.

79. We previewed the next version of Gemma, Gemma 2. It's built on a whole new architecture and will include a larger 27B parameter instance which outperforms models twice its size and runs on a single TPU host.

80. Gemini models are now available to help developers be more productive in Android Studio, IDX, Firebase, Colab, VSCode, Cloud and Intellj.

81. Gemini 1.5 Pro is coming to Android Studio later this year. Equipped with a large context window, this model leads to higher-quality responses and unlocks use cases like multimodal input.

82. Google AI Studio is now available in more than 200 countries including the U.K. and E.U.

83. Parallel function calling and video frame extraction are now supported by the Gemini API.

84. And with the new context caching feature in the Gemini API, coming next month, you'll be able to streamline workflows for large prompts by caching frequently used context files at lower costs.

85. Android now provides first-class support for Kotlin multiplatform to help developers share their apps' business logic across platforms.

86. Resizable Emulator, Compose UI check Mode and Android Device Streaming powered by Firebase are new products that can all help developers build for all form factors.

87. Starting with Chrome 126, Gemini Nano will be built into the Chrome Desktop client.

88. View Transitions API for multi-page apps, a much-requested feature, is now available so developers can easily build smooth, fluid app-like navigation regardless of site architecture.

89. Project IDX, our new integrated developer experience for full-stack, multiplatform apps, is now open for everyone to try.

90. Firebase released Firebase Genkit in beta, which will make it even easier for developers to build generative AI experiences into their apps.

91. Firebase also released Firebase Data Connect, a new way for developers to use SQL with Firebase (via Google Cloud SQL). This will not only bring SQL workflows to Firebase, but also reduce the amount of app code developers need to write.

92. We took developers under the hood in a deep-dive conversation about the technology and research powering our AI with James Manyika, Jeff Dean and Koray Kavukcuoglu.
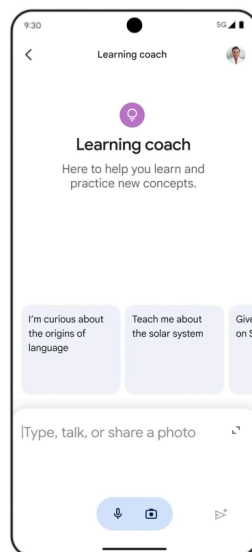
Responsible AI progress

93. We're enhancing red teaming — a proven practice where we proactively test our own systems for weakness and try to break them — through a new technique we're calling "AI-Assisted Red Teaming."

94. We're also expanding SynthID to two new modalities: text and video.

95. SynthID text watermarking will also be open-sourced in the coming months through our updated Responsible Generative AI toolkit.

96. We announced LearnLM, a new family of models based on Gemini and fine-tuned for learning. LearnLM is already powering a range of features across our products, including Gemini, Search, YouTube and Google Classroom.

97. We'll be partnering with experts from institutions like Columbia Teachers College, Arizona State University, NYU Tisch and Khan Academy to refine and expand LearnLM beyond our products.

98. And we also worked with MIT RAISE to develop an online course that equips educators to effectively use generative AI in the classroom.

99. We've built a new experimental tool called Illuminate to make knowledge more accessible and digestible.

100. Illuminate can generate a conversation consisting of two AI-generated voices, providing an overview of the key insights from research papers. You can sign up to try it today at labs.google.

POSTED IN:

AI   Developers   Gemini App   Gemini Features   Gemini Models

99. We've built a new experimental tool called Illuminate to make knowledge more accessible and digestible.

100. Illuminate can generate a conversation consisting of two AI-generated voices, providing an overview of the key insights from research papers. You can sign up to try it today at labs.google.