GOOGL Earnings Call – FY2024 Q4

Date: February 4, 2025

Operator

Welcome, everyone. Thank you for standing by for the Alphabet Fourth Quarter and Fiscal Year 2024 Earnings Conference Call. At this time, all participants are in a listen-only mode. After the speaker presentation, there will be a question-and-answer session. [Operator Instructions] I would now like to hand the conference over to your speaker today, Jim Friedland, Senior Director of Investor Relations. Please go ahead.

Jim Friedland

Thank you. Good afternoon, everyone, and welcome to Alphabet's fourth quarter 2024 earnings conference call. With us today are Sundar Pichai; Philipp Schindler; and Anat Ashkenazi. Now I'll quickly cover the Safe Harbor. Some of the statements that we make today regarding our business, operations and financial performance may be considered forward-looking. Such statements are based on current expectations and assumptions that are subject to a number of risks and uncertainties. Actual results could differ materially. Please refer to our Forms 10-K and 10-Q, including the risk factors. We undertake no obligation to update any forward-looking statement. During this call, we will present both GAAP and non-GAAP financial measures. A reconciliation of non-GAAP to GAAP measures is included in today's earnings press release, which is distributed and available to the public through our Investor Relations website located at abc.xyz/investor. Our comments will be on year-over-year comparisons unless we state otherwise. And now I'll turn the call over to Sundar.

Sundar Pichai

Thanks, Jim, and hello, everyone. We delivered another strong quarter in Q4, driven by our leadership in AI and our unique full stack. We're making dramatic progress across compute, model capabilities, and in driving efficiencies. We're rapidly shipping product improvements, and seeing terrific momentum with consumer and developer usage. And we're pushing the next frontiers, from AI agents, reasoning and deep research, to state-of-the-art video, quantum computing and more. The company is in a great rhythm and cadence, building, testing, and launching products faster than ever before. This is translating into product usage, revenue growth, and results. In Search, AI overviews are now available in more than 100 countries. They continue to drive higher satisfaction in Search usage. Meanwhile, circle to search is now available on over 200 million android devices. In Cloud and YouTube, we said at the beginning of 2024 that we expected to exit the year at a combined annual revenue run rate of over $100 billion. We met that goal and ended the year at a run rate of $110 billion. We are set up well for continued growth. So today I'll provide an update on our AI progress and how it's improving our core consumer products. Then I'll touch on Cloud, YouTube, Platforms and Devices, and Waymo. Let's start with AI. Last quarter, I outlined the three areas of our differentiated full stack approach to AI innovation. Our leading AI infrastructure, our world-class research, including models and tooling, and our products and platforms that bring these innovations to people at scale. First, AI Infrastructure. Our sophisticated global network of cloud regions and data centers provides a powerful foundation for us and our customers, directly driving revenue. We have a unique advantage, because we develop every component of our technology stack, including hardware, compilers, models, and products. This approach allows us to drive efficiencies at every level, from training and serving, to developer productivity. In 2024, we broke ground on 11 new cloud regions and data center campuses in places like South Carolina, Indiana, Missouri, and around the world. We also announced plans for seven new subsea cable projects, strengthening global connectivity. Our leading infrastructure is also among the world's most efficient. Google data centers deliver nearly 4x more computing power per unit of electricity compared to just 5 years ago. These efficiencies, coupled with the scalability, cost and performance we offer, are why organizations increasingly choose Google cloud's platform. In fact, today, cloud customers consume more than 8x the compute capacity for training and inferencing compared to 18 months ago. We'll continue to invest in our cloud business to ensure we can address the increase in customer demand. Second, world class research including models. In December, we unveiled Gemini 2.0, our most capable AI model yet, built for the agentic era. We launched an experimental version of Gemini 2.0 flash, our workhorse model with low latency and enhanced performance. Flash has already rolled out to the Gemini app, and tomorrow we are making 2.0 flash generally available for developers and customers, along with other model updates. So stay tuned. Late last year, we also debuted our experimental Gemini 2.0 Flash Thinking Model. The progress-to-scale thinking has been super fast, and the reviews so far have been extremely positive. We are working on even better thinking models and look forward to sharing those with the developer community soon. Gemini 2.0 advances in multi modality and made of tool use enable us to build new agents that bring us closer to our vision of a universal assistant. One early example is Deep Research. It uses agentic capabilities to explore complex topics on your behalf, and give key findings, along with sources. It launched in Gemini Advanced in December, and is rolling out to android users all over the world. We are seeing great product momentum with our consumer Gemini app, which debuted on IOS last November. And we have opened up trusted tester access to a handful of research prototypes, including Project Mariner, which can understand and reason across information on a browser screen to complete tasks and Project Astra. We expect to bring features from both to the Gemini app later this year. We're also excited by the progress of our video and image generation models. Veo 2, our state-of-the-art video generation model, and Imagen 3, our highest quality text-to-image model. These generative media models, as well as Gemini, consistently top industry leader boards and score top marks across industry benchmarks. That's why more than 4.4 million developers are using our Gemini models today, double the number from just 6 months ago. And we continue to drive research breakthroughs in quantum computing. At the end of last year, we announced Willow, our new state-of-the-art quantum computing chip that can reduce errors exponentially as we scale up using more qubits. Willow is an important step in our journey to build a useful quantum computer with practical applications. This technology holds so much promise, which is why there was real excitement around this breakthrough. Third, our Products and Platforms put AI into the hands of billions of people around the world. We have seven Products and Platforms with over 2 billion users and all are using Gemini. That includes Search, where Gemini is powering our AI overviews. People use Search more with AI overviews and usage growth increases over time as people learn that they can ask new types of questions. This behavior is even more pronounced with younger users who really appreciate the speed and efficiency of this new format. We also are pleased to see how Circle to Search is driving additional Search use and opening up even more types of questions. This feature is also popular among younger users. Those who have tried Circle to Search before now use it to start more than 10% of their searches. As AI continues to expand the universe of queries that people can ask, 2025 is going to be one of the biggest years for Search innovation yet. Now, let me turn to key highlights from the quarter across Cloud, YouTube, Platforms and Devices, and Waymo. First, Google Cloud. Our AI-powered cloud offerings enabled us to win customers such as Mercedes-Benz, Mercado Libre and Servier. In 2024, the number of first-time commitments more than doubled, compared to 2023. We also deepened customer relationships. Last year, we closed several strategic deals over $1 billion, and the number of deals over $250 million doubled from the prior year. Our partners are further accelerating our growth, with customers purchasing billions of dollars of solutions through our Cloud marketplace. We continue to see strong growth across our broad portfolio of AI-powered Cloud solutions. It begins with our AI Hypercomputer, which delivers leading performance and cost, across both GPUs and TPUs. These advantages help Citadel with modeling markets and training, and enabled Wayfair to modernize its platform, improving performance and scalability by nearly 25%. In Q4, we saw strong uptake of Trillium, our sixth-generation TPU, which delivers 4x better training performance and 3x greater inference throughput compared to the previous generation. We also continue our strong relationship with NVIDIA. We recently delivered their H200-based platforms to customers and just last week, we were the first to announce a customer running on the highly-anticipated Blackwell platform. Our AI developer platform, Vertex AI, saw a 5x increase in customers year-over-year, with brands like Mondelez International and WPP building new applications and benefitting from our 200+ foundation models. Vertex usage increased 20x during 2024, with particularly strong developer adoption of Gemini Flash, Gemini 2.0, Imagen 3, and most recently, Veo. We are also seeing strong growth in our AI-powered databases, data analytics, and cybersecurity platforms. Customers including Radisson Hotels are now using Gemini to Search and analyze multi-modal data from across multiple Clouds. Our AI-powered Threat Intelligence and Security Operations products help customers, including Vodafone and AstraZeneca, identify, protect and defend against threats. Our growing portfolio of AI applications is also seeing strong customer adoption. In Q4, we introduced Google Agentspace, which helps enterprises synthesize data with Google-quality Search, create Gemini-powered agents, and automate transactions for employees. In addition, we recently gave all Google Workspace Business and Enterprise customers access to all of our powerful Gemini AI capabilities to help boost their productivity. Moving to YouTube. Nielsen data shows YouTube continues to be number one in streaming watchtime in the U.S., with our share of streaming now at a record high. On election day alone, over 45 million viewers across the U.S. watched election-related content on YouTube. Our early investment in podcasts is paying off. We integrated podcasts into the core YouTube experience, particularly with video. We are now the most frequently used service for consuming podcasts in the U.S., according to a recent Edison report. This success reflects our long-term approach of investing in emerging trends, from mobile to the living room. We now have over 250,000 creators in the YouTube Shopping affiliate program in the U.S. and Korea alone. We expanded YouTube Shopping at the end of last year to three additional countries, allowing even more creators to share their favorite products with fans and grow their businesses. Philipp will talk more about YouTube performance later in the call. Next, Platforms and Devices. Google One's performance has been outstanding, and is one of our fastest growing subscription products in terms of subscribers and revenue growth. Last month, we announced the first beta of Android 16, plus new Android updates, including a deeper Gemini integration coming to the new Samsung Galaxy S25 series. We also recently announced Android XR, the first Android platform built for the Gemini era. Created with Samsung and Qualcomm, Android XR is designed to power an ecosystem of next-generation extended reality devices, like headsets and glasses. Finally, a few words on Waymo which made tremendous progress last year, safely serving more than 4 million passenger trips. It is now averaging over 150,000 trips each week, and growing. Looking ahead, Waymo will be expanding its network and operations partnerships to open up new markets, including Austin and Atlanta this year and Miami next year. And in the coming weeks, Waymo One vehicles will arrive in Tokyo for their first international road trip. We're also developing the sixth-generation Waymo Driver, which will significantly lower hardware costs. I want to thank our employees around the world for another great quarter. 2025 is going to be exciting and we're all ready for it. Philipp, I will hand it over to you.

Philipp Schindler

Thanks, Sundar, and hello, everyone. I will quickly cover performance for the quarter, then frame the rest of my remarks around the progress we are delivering across Search, Ads, YouTube and Partnerships, highlighting the impact AI is having on our business and our customers. Google Services revenues were $84 billion for the quarter, up 10%, driven primarily by 11% year-on-year growth in advertising revenues. Strong growth in Search and YouTube advertising was partially offset by year-over-year decline in network revenues. In terms of vertical performance, the 13% increase in Search and other revenues was led by financial services followed by retail. The 14% growth in YouTube advertising revenues was driven by strong spend on U.S. election advertising, with combined spend from both parties almost doubling from what we saw in the 2020 elections. Now, in Q4, we saw continued strong growth in revenues from Search. We had lots of exciting updates in December, and we're rapidly integrating our AI innovation into our consumer experiences. We've already started testing Gemini 2.0 in AI overviews and plan to roll it out more broadly later in the year. In Search, we're seeing people increasingly ask entirely new questions using their voice, camera, or in ways that were not possible before, like with Circle to Search. We're making these benefits available to more consumers. Google is already present in over half of journeys where a new brand, product, or retailer are discovered. By offering new ways for people to Search, we're expanding commercial opportunities for our advertisers. Shoppers can now take a photo of a product and, using lens, quickly find information about the product, reviews, similar products, and where they can get it for a great price. Lens is used for over 20 billion visual search queries every month, and the majority of these searches are incremental. Retail was particularly strong this holiday season, especially on Black Friday and Cyber Monday, which each generated over $1 billion in ad revenue. Interestingly, despite the U.S. holiday shopping season being the shortest since 2019, retail sales began much earlier, in October, causing the season to extend longer than anticipated. People shop more than a billion times a day across Google. Last quarter, we introduced a reinvented Google shopping experience, rebuilt from the ground up with AI. This December saw roughly 13% more daily active users on Google shopping in the U.S. compared to the same period in 2023. Closing out on Search with travel, and sharing another interesting trend where we saw spend expand to travel Tuesday. This contributed to 20% year-on-year revenue growth for travel advertisers across Cyber Monday and Travel Tuesday. Moving to Ads. We continue investing in AI capabilities across media buying, creative and measurement. As I said before, we believe that AI will revolutionize every part of the marketing value chain, and over the past quarter, we've seen how our customers are increasingly focusing on optimizing their use of AI. As an example, Petco used DemandGen campaigns across targeting, creative generation, and bidding to find new pet parent audiences across YouTube. They achieved a 275% higher return on ad spend and a 74% higher click through rate than their social benchmarks. On media buying, we made YouTube select creator takeovers generally available in the U.S. and will be expanding to more markets this year. Creators know their audience the best and creator takeovers help businesses connect with consumers through authentic and relevant content. Looking at Creative, we introduced new controls and made reporting easier in PMAX, helping customers better understand and reinvest into their best performing assets. Using asset generation in PMAX. Event Ticket Center achieved a 5x increase in production of creative assets, saving time and effort. They also increased convergence by 300% compared to the previous period when they used manual assets. And finally, Measurement. Last week, we made Meridian, our marketing mix model, generally available for customers, helping more business reinvest into creative and media buying strategies that they know work. Based on a Nielsen meta-analysis of marketing mix models, on average, Google AI-powered video campaigns on YouTube deliver 17% higher return on advertising spend than manual campaigns. Turning to YouTube. We saw robust revenue growth backed by continued growth and watch time across ad supported and premium experiences. Our focus here remains on building a streaming platform that enables creators to thrive and unlock the full potential of AI. Expanding on our state-of-the-art video generation model, we announced Veo 2, which creates incredibly high quality video in a wide range of subjects and styles. It's been inspiring to see how people are experimenting with it. We'll make it available to creators on YouTube in the coming months. We continue to invest in helping YouTube creators work with brands. All advertisers globally can now promote YouTube creator videos and ad campaigns across all AI-powered campaign types in Google Ads. And creators can tag partners in their brand videos. Sephora used DemandGen's Shorts-Only channel to boost traffic and brand searches for the Holiday Gift Guide campaign and leverage creator collaborations to find the best gift. This drove an 82% relative uplift in searches for Sephora Holiday. Shorts continues its ascent and is closing the gap with long form. In 2024, the monetization rate of Shorts relative to in-stream viewing increased by more than 30 percentage points in the U.S., and we expect to make additional progress in 2025. We're making it easier for advertisers to benefit from Shorts on all screens. We're particularly excited by its success on connected TV, which now makes up 15% of shorts viewing in the U.S. Using a combination of ad formats, Louis Vuitton reached their overall objectives on both long-form and short-form content. Their shorts exceeded luxury goods benchmark for average view duration by 89% for equivalent video lengths, while their long-form content exceeded the benchmark by over 15%, with strong engagement from Gen Z and Millennials. Looking into the living room, we continue to be number one in streaming watch time in the U.S for nearly two years, according to Nielsen, and our share of streaming is at a record high. Viewers globally streamed over 1 billion hours of YouTube content daily on their TVs in 2024. YouTube makes multi-year investments to tap into shifting consumer behavior. The current surge in living room viewership directly reflects years of work to build the right products and partnerships. Creators are now prioritizing high-quality viewing experiences that truly shine on TV screens, inspiring even more viewers to tune in. In fact, the number of creators making majority of revenue from TV is up over 30% year-on-year. We have also invested in podcasts, where popular shows like Club Shay Shay and Lex Friedman are increasingly a visual format. YouTube creators and viewers are embracing this. In 2024, people watched over 400 million hours of podcasts each month on living room devices alone. YouTube is now the most popular service for podcast listening in the U.S., according to Edison. As always, let me wrap with the strong momentum we're seeing in partnerships, where the breadth of what Google has to offer is increasingly being recognized. Sundar mentioned our deepening partnership with Samsung. Another expanding partnership is with Citi, who is modernizing its technology infrastructure with Google Cloud to transform employee and customer experiences. Using Google Cloud, it will improve its digital products, streamline employee workflows, and use advanced high-performance computing to enable millions of daily computations. This partnership also fuels Citi's generative AI initiatives across customer service, document summarization, and search to reduce manual processing. With that, allow me a moment to thank Googlers everywhere for their extraordinary commitment and to our customers and partners for their continued trust. Anat, over to you.

Anat Ashkenazi

Thank you, Philipp. We're pleased with the continued momentum we're seeing across the business as Alphabet revenue for 2024 reached $350 billion, up 14% on a reported basis and 15% in constant currency versus 2023. My comments will focus on year-over-year comparisons for the fourth quarter, unless I state otherwise. I will start with the results at the Alphabet level and we'll then cover our segment results. I'll end with some commentary and expectations over the first quarter and full year 2025. We had another strong quarter in Q4 with robust momentum across the business. Consolidated revenue of $96.5 billion, increased by 12% in both reported and constant currency. Search remained the largest contributor to revenue growth, followed by Cloud. Total cost of revenue was $40.6 billion, up 8%. Tech was $14.8 billion, up 6%. We continue to see a revenue mix shift with Google Search growing at double-digit levels, while network revenues, which have a much higher tech rate, declined. Other cost of revenue was $25.8 billion, up 9%, with the increase primarily driven by content acquisitions costs, primarily for YouTube, followed by depreciation, due to increasing investments in our technical infrastructure. Growth in content acquisition and depreciation were partially offset by our year-over-year decline in hardware costs due to the shift in timing of our made-by-Google launches to the third quarter 2024 compared to the fourth quarter of 2023. In terms of total expenses, the year-over-year comparisons reflect $1.2 billion in exit charges that we took in the fourth quarter of 2023 in connection with actions to optimize our global office space. As previously disclosed, those charges were allocated across the expense lines in other costs of revenue and OpEx based on associated headcount. Total operating expenses decreased 1% to $24.9 billion. R&D investments increased by 8%, primarily driven by increase in compensation and depreciation expenses, partially offset by the impact of charges for office [ph] space optimization in the fourth quarter of 2023. Sales and marketing expenses decreased 5%, primarily reflecting the optimization charges last year, as well as declines in compensation and in ads and promotion expenses due to the timing shift of the Pixel launch from Q4 to Q3. G&A expenses declined by 15%, reflecting a shift of timing in our charitable contributions, as well as the optimization charges last year. Operating income increased 31%, the score [ph] to $31 billion, and operating margin increased to 32%, representing 4.6 points of margin expansion. Net income increased 28% to $26.5 billion, and earnings per share increased 31% to $2.15. We delivered free cash flow of $24.8 billion in the fourth quarter and $72.8 billion for the full year 2024. We ended the quarter with $96 billion in cash and marketable securities. Turning to segment results. Google Service revenues increased 10% to $84.1 billion, reflecting the strong momentum across Google Search and YouTube ads. Google Search and other advertising revenues increased by 13% to $54 billion. The robust performance of Search was once again broad-based across verticals, led by the financial service vertical due to strength in insurance, followed by retail. YouTube advertising revenue increased 14% to $10.5 billion, driven by brand, followed by direct response advertising. Network advertising revenue of $8 billion, were down 4%. In the fourth quarter, the year-over-year comparison in all of our advertising revenue lines was impacted by the increase in strength in advertising revenue in Q4 2023, in part from APAC-based retailers. Subscription platforms and device revenues increase 8% to $11.6 billion, primarily reflecting growth in subscription revenues, partially offset by the shift in timing of the launch of our made by Google devices to the third quarter, compared with the fourth quarter in 2023. We continue to have significant growth in our subscription products, primarily due to increase in the number of paid subscribers across YouTube TV, YouTube Music Premium, and Google One. With regards to Platform, we saw a slight increase in the growth rate in Play, primarily due to a strong increase in the number of buyers. Google's service operating income increased 23% to $32.8 billion, and operating margin increased from 35% to 39%, representing a meaningful margin expansion. Turning to the Google Cloud segment, which continued to deliver very strong results this quarter. Revenue increased by 30% to $12 billion in the fourth quarter, reflecting growth in GCP across core GCP products, AI infrastructure, and generative AI solutions. Once again, GCP grew at a rate that was much higher than cloud overall. Healthy Google Workspace growth was primarily driven by increase in average revenue per seat. Google Cloud operating income increased to $2.1 billion, and operating margin increased from 9.4% to 17.5%. We're pleased with the work the cloud team is doing to deliver valuable solutions to the customer and generate revenue growth, as well as its continued focus on driving efficiencies across the cloud business. As for Other Bets, for the fourth quarter, revenue were $400 million, and the operating loss was $1.2 billion. The year-over-year decline in revenue and increase in operating loss primarily reflect the milestone payment in the fourth quarter of 2023 for one of the Other Bets. Turning to Alphabet level activities, the largest component of this line is our investments in AI research and development activities which support all of Alphabet. As a reminder, Alphabet level activities have included nearly all severance charges from reductions in workforce and office space charges. In the fourth quarter of 2024, the biggest factor in year-over-year comparison is the $1.2 billion in charges in the fourth quarter of 2023, almost entirely in connection with office space optimization. With respect to CapEx, our reported CapEx in the fourth quarter was $14 billion, primarily reflecting investments in our technical infrastructure, with the largest component being investment in servers, followed by data centers, to support the growth of our business across Google Services, Google Cloud, and Google DeepMind. In Q4, we returned value to shareholders in the form of $15 billion in share purchases and $2.4 billion in dividend payments. Overall, we returned a total of nearly $70 billion to shareholders in 2024. Turning to 2025, I would like to provide some commentary on several factors that will impact our business performance in both the first quarter and the full year 2025. First in terms of revenue, I'll highlight two items that will have meaningful impact on Q1 revenue across the company. First in terms of revenues, I'll highlight two items that will have meaningful impact on Q1 revenues across the company. The first is the impact of foreign exchange rates. At the current spot rates, we expect a larger headwind to our revenues from the strengthening of the U.S dollar relative to key currencies in Q1 versus Q4 2024. Second is the impact of leap year. We expect a headwind from having one less day of revenue in Q1 2025 compared with leap year in the first quarter of 2024. As for our segments, Google Services, advertising revenue in 2025 will be impacted by lapping the strength we experience in the financial service vertical throughout 2024. And in Cloud, given the revenues are correlated with the timing of deployment of new capacity, we could see variability in cloud revenue growth rates depending on when new capacity comes online during 2025. Moving to investments, starting with our expectation for CapEx for the full year 2025. As we mentioned on the Q3 call, as we expand our AI efforts, we expect to increase our investments in capital expenditure for technical infrastructure, primarily for servers, followed by data centers and networking. We expect to invest approximately $75 billion in CapEx in 2025 with approximately $16 billion to $18 billion of debt in the first quarter. The expected total investment level may fluctuate from quarter-to-quarter, primarily due to timing of deliveries and construction schedules. In terms of expenses, first, the increase in our investment in CapEx over the past few years will increase pressure on the P&L, primarily in the form of higher depreciation. In 2024, we saw 28% year-over-year growth in depreciation as we put more technical infrastructure assets into service. Given the increase in CapEx investments over the past few years, we expect the growth rate and depreciation to accelerate in 2025. Second, we expect some headcount growth in 2025 in key investment areas such as AI and cloud. As you just heard from Sundar, we're delivering products and solutions to customer at a rapid pace, building, testing, and launching products faster than ever before. And as I mentioned on the Q3 call, we're doing that while also focusing on driving further efficiencies in how we operate the business. Before we take questions, I'd like to recap the financial results for the year. For the full year 2024, revenue grew by 14% or by $43 billion, reaching $350 billion. Google Services and Google Cloud each continue to see double-digit revenue growth coupled with margin extension. YouTube and cloud revenues combined, ended the year at $110 billion annual run rate. And in 2024, we generate total income of $112 billion, an increase of 33% from 2023. We're pleased with the momentum we're seeing in AI innovation and monetization. We've been using AI to improve the performance of our ads business for well over a decade, and Cloud is generating billions in annual revenue from AI infrastructure and generative AI solutions. We're also excited about the potential to bring new experiences to users that will provide additional opportunities for monetization. And I look forward to sharing more in our progress throughout the year. Sundar, Philipp, and I will now take your questions.

Operator

[Operator Instructions] And our first question comes from Brian Nowak with Morgan Stanley. Your line is now open.

Brian Nowak

Thanks for taking my questions. I have two, one for Sundar, one for Anat. Sundar, maybe kind of step back on Search, there's a -- it seems like there's a lot of advances to come with Gen AI and agentic possibilities with Search. Can you just sort of walk us through your big picture vision over the next few years of how you think about your Search product continue to evolve to stay at the top of the funnel and drive more engagement and monetization for your users and advertisers. And then the second one, Anat, I think about 90 days ago, you talked about sort of further efficiencies and areas of simplification on the OpEx space. Can you just sort of walk us through some examples of where you see the potential for further efficiencies to the OpEx space, excluding the DNA step-ups that we have to come in 2025. Thanks.

Sundar Pichai

Thanks, Brian. On Search, obviously, we view, I mean, this has been a long continual journey. AI overviews has been the next step. It's playing out positively, as we have indicated, the metrics look great, and we are obviously trading on that experience, bringing better and better models, expanding to the number of queries where it works and so on. But there's a lot more to come. I think we'll continue bringing AI in more powerful ways, in multi-modal ways. Things like what we've done with Lens Circle to Search, you can imagine the future with Project Astra. You can also imagine areas like we have done in Gemini Deep Research, possibilities where you are really dramatically expanding the types of use cases for which Search can work, things which don't always get answered instantaneously but can take some time to answer. Those are all areas of explorations and you will see us, putting newer experiences in front of users through the course of 2025. And so I do feel the opportunity space with AI, there's a lot of unlock ahead.

Anat Ashkenazi

Thanks. And for the question with regards to where do we see or where do I see leverage moving forward and some of the comments I've made on the previous call. I certainly see opportunities for further productivity and efficiency, and this is one of our priority areas. And we're going to do that so that we can make sure we continue to invest in areas such as AI and cloud where we see potential for continued growth. I'll remain focused on areas that I've mentioned before, which include the technical infrastructure, so the $75 billion in CapEx I mentioned for this year, the majority of that is going to go towards our technical infrastructure, which includes servers and data centers. So ensuring we do that in the most efficient way is critical. Second is managing headcount growth, and we're going to be investing in areas of growth, such as AI and cloud, but looking across the organization and moderating that growth will be important. Optimizing the real estate footprint is one of the areas I've mentioned. We're continuing to focus on that. As well as looking at how we simplify the organization, we've previously mentioned bringing like areas together. Sundar talked about bringing some of the AI research teams together so that we can operate with increased speed, but also how we operate within the organization. Using our own AI tools to how we run the business, whether it's the code that Sundar mentioned on the previous call, writing code with AI, or even running some of our key processes using AI tools. So we're looking at all that. It's going to -- it's -- this is not a one-quarter type of effort. It's going to continue throughout the year, and we're going to continue to focus on that so that we can support the growth in other areas.

Operator

Our next question comes from Doug Anmuth with JPMorgan. Your line is now open.

Doug Anmuth

Thanks for taking the questions, one for Philipp and one for Anat. Philipp, can you just talk more about the expanded rollout of ads on AI overviews and perhaps what additional things you may have learned in 4Q? And I guess in particular, just curious if you rolled out to a higher percentage of commercial queries and is it still fair to say that you're monetizing nearly on par with existing search? And then Anat, just on the -- on cloud growth, a little bit of decel 3Q to 4Q, but it sounded like you also suggested that your capacity constrained in the fourth quarter. I just wanted to push on that a little bit more. How -- is that accurate and is it fair to say that revenue growth could have been higher with much more capacity? Thanks.

Philipp Schindler

So, on your first question, first of all, AI overviews, which is really nice to see, continue to drive higher satisfaction and search usage. So, that's really good. And as you know, we recently launched the ads within AI overviews on mobile in the U.S., which builds on our previous rollout of ads above and below. And as I talked about before, for the AI overviews overall, we actually see monetization at approximately the same rate, which I think really gives us a strong base on which we can innovate even more.

Anat Ashkenazi

On the cloud questions, so first I'm excited that we ended the quarter at $12 billion and a 30% year-over-year growth, very impressive growth. And as I've mentioned, the prepared remark, GCP grew at a much higher rate than overall cloud. Two items to think about from a deceleration perspective. The first is we are lapping a very strong quarter of [indiscernible] deployment in Q4 of 2023. The second is the one you've alluded to. We do see and have been seeing very strong demand for AI products in the fourth quarter in 2024. We exited the year with more demand than we had available capacity. So we are in a tight supply demand situation, working very hard to bring more capacity online. As I mentioned, we've increased investment in CapEx in 2024, continue to increase in 2025, and will bring more capacity throughout the year.

Doug Anmuth

Thank you both.

Operator

Your next question comes from Eric Sheridan with Goldman Sachs. Your line is now open.

Eric Sheridan

Thank you for taking the question. I'll just ask one. Sundar, with the news that came out of China a little over 2 weeks ago, I think investors have been asking a lot of questions about the long-term cost curve for AI as AI moves from the infrastructure layer to the application layer or from training to inference and maybe even custom silicon becomes more dominant across the theme. I would love to get your perspective on your take on that news a couple of weeks ago and what it might mean for Alphabet longer term. Thank you.

Sundar Pichai

Thanks, Eric. Look, I think there's been a lot of observations on DeepSeek. First of all, I think a tremendous team. I think they've done very, very good work. Look, I think for us it's always been obvious, over time, there's frontier model development, but you can drive a lot of efficiency to serve these models really, really well. And if you look at one of the areas in which the Gemini model shine is the Pareto frontier of cost, performance, and latency and if you look at all three attributes, I think we lead this Pareto frontier. And I would say both are 2.0 flash models or 2.0 flash thinking models. They are some of the most efficient models out there, including comparing to DeepSeek's V3 and R1. And I think a lot of it is our strength of the full stack development into an optimization, our obsession with cost per query, all of that, I think, sets us up well for the workloads ahead, both to serve billions of users across our products and on the cloud side. Couple of things I would say are, if you look at the trajectory over the past 3 years, the proportion of the spend towards inference compared to training has been increasing, which is good because obviously inference is to support businesses with good ROIC. And so I think that trend is good. I think the reasoning models, if anything, accelerates that trend because it's obviously scaling upon inference dimension as well. And so I think, look, I think part of the reason we are so excited about the AI opportunity is we know we can drive extraordinary use cases because the cost of actually using it is going to keep coming down, which will make more use cases feasible. And that's the opportunity space. It's as big as it comes. And that's why you're seeing us invest to meet that moment.

Eric Sheridan

Thank you.

Operator

Your next question comes from Michael Nathanson with MoffettNathanson. Your line is now open.

Michael Nathanson

Thank you. I have two, one for Philipp and one for Anat. Philipp, question for you is we're starting to see more AI tools on eCommerce sites and something like research with AI recommendations is on Google Shopping. Can you talk about how that product and other AI tools are impacting shopping behavior? I've had that impact in modernization. And then I guess the $75 billion question Anat is, how do you think about long-term capital intensity for this business? It sounds like there's a bit of constraint in terms of getting things built, but how do you think about the modeling of capital intensity going forward, and what are the things that you're looking forward to in terms of whether or not this is the right level of spend? Thank you.

Philipp Schindler

Look, excellent question. We've been using our advances in AI to make search for products on Google even easier, obviously. And in Q4, we actually introduced quite a transformed Google Shopping experience, which we rebuilt from the ground up with AI. And people shop more than a billion times a day across Google. Last quarter, we introduced this fully reinvented Google Shopping experience. In December, we saw roughly, I mentioned, as 13% more daily active users in Google Shopping in the U.S compared to the same period last year. So that's a good development here. And the new Google Shopping experiences, specifically to your question, uses AI to really intelligently show the most relevant products, helping to speed up and simplify your research. You get an AI-generated brief with top things to consider for your search, plus maybe products that meet your needs. So shoppers very often want low prices. So the new page not only includes like deal finding tools like price comparison, price insights, price tracking throughout, but there's also a new and dedicated personalized deals page where you can browse deals for you, and all this is really built on the backbone of AI. So we think this is a very interesting opportunity.

Anat Ashkenazi

And on the question of capital expenditure, and I think you may have two questions in there, one is just the capital intensity and then how would we think about return on that invested capital. So on the first one, certainly we are looking ahead, but we are managing very responsibly with a very rigorous even internal governance process looking at how do we allocate the capacity and what would we need to support the customer demand externally, but also across the Google, the Alphabet business. And as you've seen in the comment I've just made on cloud, we do have demand that exceeds our available capacity, so we'll be working hard to address that and make sure we bring more capacity online. We do have the benefit of having a very broad business, and we can repurpose capacity, whether it's through Google Services or Google Cloud to support, as I said, whether it's Search or GDM or Google Cloud Customers, we can do that in a more efficient manner. We also look at every investment that we make to ensure that we're doing it in the most cost-efficient way to optimize our data center. As you know, our strategy is mostly to rely on our own self-design and build data centers. So they're industry-leading in terms of both cost and power efficiency at scale. We have our own customized TPUs. They're customized for our own workload. So, they do deliver outstanding superior performance and CapEx efficiency. So, we're going to be looking at all that when we make decisions as to how we're going to progress capital investments throughout the coming years.

Michael Nathanson

Thank you.

Operator

Your next question comes from Mark Shmulik with Bernstein. Your line is now open.

Mark Shmulik

Yes, thanks for taking the question. Sundar, I just wanted to follow-up to Brian's question from earlier. With your own Project Mariner efforts and a competitor's recent launch, it seems there's suddenly really strong momentum on AI consumer agents and kind of catching up to that old Google duplex vision. I think when you look a few years ahead, where do you see consumer agents going and really, what does it mean to Google Search outside of Lens? Is there room for both to flourish or will they eventually be in conflict with each other? Thank you.

Sundar Pichai

Look, I think first of all, we are definitely seeing a lot of progress in the underlying capabilities of these models. Gemini 2.0 was definitely built with the view of enabling more agentic use cases. And so, I actually, and we are definitely seeing progress inside, and I think we'll be able to do more agentic experiences for our users. Look, I actually think all of this expands the opportunity space. I think historically we've had information use cases, but now you can have, you can act on your information needs in a much deeper way. It's always been our vision when we have talked about Google Assistant, et cetera. So I think the opportunity space expands. I think there's plenty of, it feels very far from a zero-sum game. There's plenty of room, I think, for many new types of use cases to flourish. And I think for us, we have a clear sense of additional use cases we can start to tackle for our users in Google Search. And all the early work with AI overview shows that users will react positively to that. So, I'm pretty excited about what's ahead.

Operator

Your next question comes from Ross Sandler with Barclays. Your line is now open.

Ross Sandler

Great. One on infrastructure and then one on the guidance on revenue. So Sundar, if we look at the inference cost per 1 million tokens and not the API pricing we see for Gemini versus something like GPT-4, but the raw cost of generating inference tokens on your TPU stack. How much more efficient do you think you guys are in terms of generating 1 million tokens compared to inference costs running on your cloud peers? Do you see this as an advantage as everything shifts to inference? And then Anat, you call out lapping the financial services category strength in 2025 as a bit of a problem for Search. Could you just quantify a little bit of that? Is that kind of the same as when you guys talk about lapping the Asia outbound advertiser channel? Any numbers you could put around that headwind? Thank you very much.

Sundar Pichai

Ross, like, look, the whole TPU project started, V1 was effectively an inferencing chip, so we've always, part of the reason we have taken the end-to-end stack approach is that, so that we can definitely drive a strong differentiation in end-to-end optimizing, and not only on a cost, but on a latency basis and a performance basis. We have the Pareto frontier we mentioned, and I think our full-stack approach and our TPU efforts all play -- give a meaningful advantage, and we plan, you already see that. I know you asked about the cost, but it's effectively captured. When we price outside, we pass on the differentiation. It's partly why we've been able to bring forward flash models at very attractive value props, which is what is driving both developer growth. We've doubled our developers to 4.4 million in just about 6 months. Vertex usage is up 20x last year. And so all of that is a direct result of that approach and so we will continue doing that.

Anat Ashkenazi

And on the question regarding my comment on lapping the strength in financial services, this is primarily related to the structural changes with regards to insurance, it is more specifically within financial services, it was the insurance segment and we saw that continue, but it was a one-time kind of a step up and then we saw it throughout the year. I am not going to give any specific numbers as to what we expect to see in 2025, but I am pleased with the fact that we are seeing and continue to see strength across really all verticals including retail and exiting the year in a position of strength. If anything, I would highlight as you think about the year, the comments I have made about the impact of FX, as well as the fact that we have one less day of revenue in Q1.

Operator

Your next question comes from Justin Post with Bank of America. Your line is now open.

Justin Post

Great. Thank you. A couple for Philipp. You mentioned higher search usage with overviews. Just wondering how you're feeling about overall Search usage. Is it accelerating as you integrate more AI? I know there's a lot of traffic growth at competitors with AI, but just wondering if you're seeing a real increase in total volumes of information gathering. And then second, on YouTube, thinking about kind of maybe a move from more professional content to user-generated content, what is that doing for your usage, and how do you think about the margin impact from that? Thank you.

Philipp Schindler

Maybe on -- Justin on Search usage overall, our metrics are healthy. We are continuing to see growth in Search on a year-on-year basis in terms of overall usage. Of course, within that, AI overviews has seen stronger growth, particularly across all segments of users, including in younger users. So it's being well received. But overall, I think through this AI moment, I think Search is continuing to perform well. And as I said earlier, we have a lot more innovations to come this year. I think the product will evolve even more. And I think as you make Search, as you give, as you make it more easy for people to interact and ask follow-up questions, et cetera, I think we have an opportunity to drive further growth.

Sundar Pichai

And some color on your YouTube question. Look, YouTube ads overall had a very healthy growth in Q4 driven by brand followed by direct response. The U.S. election advertising led brand revenue growth, and we saw nearly double the spending from 2020. I mentioned that we also had a strong contribution from finance, retail, and the tech verticals. So overall, strong operating metrics. Watch time growth remains robust, particularly in key monetization opportunity areas, such as Shorts and Living Room, just to set the stage here one more time. On your question specifically on the UGC side, look, we have a very strong position with creators and we've always said creators are at the center of YouTube success here. They're the number one most important thing we care about and this strong position really gives us a lot of confidence here. Today we have more than 3 million channels that are actually in the YouTube Partner Program, which is really an incredible program. So we're very confident with the position we have and where this goes.

Justin Post

Thank you.

Operator

And our last question comes from Ken Gawrelski with Wells Fargo. Your line is now open.

Ken Gawrelski

Thank you very much. Two, please. I want to focus on Gemini and the consumer agent side. Sundar, there was some press reports that suggested you have an ambitious goal on growing the usage by the end of '25. Two questions on that, please. First, what’s -- how should we think about the approach that you're going to employ to achieve this goal? Is it more aggressive marketing for Gemini as a standalone product, or is it tighter integration into existing experiences, whether it be search mail, maps, et cetera. And then the second one is, how should we think about the future monetization opportunity of Gemini? Today, it's really a premium subscription offering or a free offering. Over time, do you see an ad component and anything you can share on that, please? Thank you.

Sundar Pichai

First of all, we've had strong momentum for Gemini on the app side, particularly through the second half of 2024. And some of it is we've made it more easily accessible. We've brought it to, for example, with a dedicated app on iOS, which has been super positively received and definitely getting a lot of traction there. So definitely driving organic growth by putting the product out. We just last week rolled out with our 2.0 series of models. So 2.0 Flash, I mean, I think that's one of the most capable models you can access at the free tier. So that's definitely contributing as well. And so we are rapidly trading. We've had a couple of key innovations there. Gemini Live, I think, has been definitely a hit with users, as well as for advanced users, Gemini Deep Research. So I think a combination of innovation, continually trading on the product and making it better, is driving a lot of usage. And we'll have a lot more to come as we go this year. And we obviously have a partnership with Samsung, so there are other things which will contribute to it as well. On the monetization side, obviously for now we are focused on a free tier and subscriptions, but obviously as you've seen in Google over time, we always want to lead with user experience, We do have very good ideas for native ad concepts, but you will see us lead with the user experience. But I do think we're always committed to making the products work and reach billions of users at scale. And advertising has been a great aspect of that strategy. And so just like you've seen with YouTube, we'll give people options over time. But for this year, I think you will see us be focused on the subscription direction.

Ken Gawrelski

Thank you.

Operator

Thank you and that concludes our question-and-answer session for today. I would like to turn the conference back over to Jim Friedland for any further remarks.

Jim Friedland

Thanks everyone for joining us today. We look forward to speaking with you again on our first quarter 2025 call. Thank you and have a good evening.

Operator

Thank you everyone. This concludes today's conference call. Thank you for participating. You may now disconnect.