BERNSTEIN
SOCIETE GENERALE GROUP

3 November 2026

Global Software and US Communications Infrastructure

What are the AI Lab datacenter / compute commitments and what does it mean for hyperscalers and neoclouds?



Mark L. Moerdler, Ph.D.
+1 917 344 8506
mark.moerdler@bernsteinsg.com



Madison Rezaei
+1 917 344 8622
madison.rezaei@bernsteinsg.com



Firoz Valliji, CFA
+1 917 344 8316
firoz.valliji@bernsteinsg.com



Shelly Tang, CFA
+1 917 344 8342
shelly.tang@bernsteinsg.com



Nancy Wu
+1 917 344 8545
nancy.wu@bernsteinsg.com

With the AI Labs contracted to use a significant portion of the AI computing capacity being built, concerns are high of customer concentration risk and what the ramifications could be in a worst case condition. Today's note estimates OpenAI's annual contracted commitment as a proxy for all the AI Labs. We show that while there are substantial commitments they are layered over time and not all "commitments" are non-cancelable or have a specific time frame. While the capacity and the associated costs are growing so is their ability to leverage and pay. In other words, while concerns will continue, we believe the downside risks to Oracle and Microsoft and the other primary suppliers is low. The risks to the industry, we would argue, are not from the leading AI labs especially as they build out their platform (Link) and consumer / search businesses.

The wind came out of the sails of the AI trade over the last few weeks, spurred by neocloud credit fears (Link) and increased concerns around delivery delays. The biggest questions circulating are around the demand for data center and capacity: how much appetite is there, who funds it, and who are the winners and losers if there is a mismatch? Whether looking at hyperscalers, data center providers, or neoclouds, answering any of these questions comes down to a debate about the AI labs and their respective commitments.

Today's note estimates OpenAI's annual contracted commitment as a proxy for all AI Labs. While there are substantial commitments across the stack, the commitments are layered and time-sequenced. Not all commitments are non-cancellable. There is often fungibility in timing and sometimes location. And while the capacity and the associated costs are growing, so is OpenAI's ability to leverage and pay.

Given what we know today, we do not see risk to Oracle, Microsoft, or CoreWeave's existing deals with OpenAI. We believe existing systemic risk is less from the scale of lab commitments, especially as they build out their platform (Link) and consumer / search businesses, and more from an easing of the capacity scarcity that has led to more speculative builds and lower-quality business models. We also think risk varies by provider - contracts and commitments are not monolithic across suppliers, as we discussed in our recent interview with a hyperscaler leasing expert - Microsoft and Oracle especially have strong negotiating positions and likely tight contracts both with suppliers and customers.

We continue to be skeptics of the neocloud business model and believe labs will increasingly rely on large-scale cloud partnerships with hyperscalers, co-location deals, and even more self-builds as they mature. So while we do question the renewal likelihood of any OpenAI deals and the durability of long-term neocloud demand, we have reasonable confidence in the existing backlog and OpenAI's ability to pay.

See the Disclosure Appendix of this report for required disclosures, analyst certifications and other important information. Alternatively, visit our Global Research Disclosure Website.

First Published: 03 Nov 2026 11:37 UTC Completion Date: 03 Nov 2026 11:37 UTC

BERNSTEIN TICKER TABLE

			2 Nov 2026		TTM	Adjusted EPS				Adjusted P/E (x)		
			Closing	Price	Rel.							
Ticker	Rating	Cur	Price	Target	Perf.	Cur	2026A	2027E	2028E	2026A	2027E	2028E
MSFT (Microsoft)	O	USD	507.12	722.00	(12.2)%	USD	17.28	20.63	25.85	29.3	24.6	19.6
ORCL (Oracle)	O	USD	119.22	325.00	(64.0)%	USD	7.63	8.37	12.19	15.6	14.2	9.8
CRWV (CoreWeave)	U	USD	50.15	54.00	(70.6)%	USD	(1.20)	(5.14)	(6.60)	N/M	235	29.8
SPX			7,541.40									

O - Outperform, M - Market-Perform, U - Underperform, NR - Not Rated, CS - Coverage Suspended

CRWV valuation is EV/EBIT (x); CRWV base year is 2025;

Source: Bloomberg, Bernstein estimates and analysis.

INVESTMENT IMPLICATIONS

Overall we believe that investors should keep an eye on growth and spending at OpenAI (which will get easier once they file their S-1 and then go public) but not be overly concerned about the quality of the contracts between OpenAI and their partners, especially Microsoft and Oracle. We are not modeling in contract renewals after the initial contract duration, and we expect the smaller / shorter contracts may not also be renewed. But the contracts with Oracle and Microsoft will most likely be renewed at least one more time.

Oracle: Initially the announcement of the $300B of RPO from OpenAI drove the stock up, but then investor concerns kicked in surrounding not just the OpenAI agreements but all the AI datacenter agreements : 1) how much Oracle would need to raise; 2) the ROI for Oracle; 3) execution risk and 4) Customer concentration risk. Oracle at their FAM put to bed concern #1 and gave lots of details on how profitable their AI datacenter business will be (concern #2). It also showed how the company is managing the diverse execution risks (#3). Today's note while it does not completely answer the OpenAI exposure question should show that OpenAI management is ramping capacity over time and in-line to slower than revenue (especially if we expect some capacity to delay). Please note that OpenAI accounts for < 50% of Oracle RPO which likely decreases over time as Oracle's businesses in general and OCI AI specifically continue to grow.

Microsoft has a complex relationship with OpenAI as an investor, leveraging OpenAI models, right to use OpenAI IP; and having a commitment to deliver $250B in capacity (CPU and GPU) to OpenAI over time. Each tranche of capacity is a six year contract which equates to the useful life of the hardware and similar to Oracle ramps over time. Given the ramping nature of OpenAI's capacity additions, the likelihood some contracts will delay / be canceled and Microsoft's large global footprint we suspect that Microsoft is less likely to see delays impacting revenue than the neoclouds and co-location vendors and with the data in this note we feel confident that Micosoft will see that OpenAI revenue on-time and at good unit economics.

CoreWeave has signed up to ~$22.4B worth of contracts with OpenAI: an $11.9B deal signed in March 2025 (+$350M equity stake), another $4.0B in May 2025, and up to another $6.5B in September 2025. Based on the language, these are caps, not committed revenue - the contracts have commenced, but it's not yet clear how far along the contracts are, the exact expiration dates, or the power capacity. We have written that we don't have concerns about CRWV's backlog, though we are skeptical of their ability to renew these deals or sign incremental deals at strong economics given the proliferation of competitors.

Table Of Contents

Portfolio Managers Summary 4
What does this mean for our covered companies? 6
Key Takeaways 7
What has OpenAI contractually commitment to? 8
Setting the stage for the discussion 8
How much revenue can a MW generate? 9
How can we break apart OpenAI's compute spending into categories? 10
How much revenue do the Hyperscalers capture? 11
Oracle 12
Microsoft 13
Amazon 14
What do the Neoclouds capture in revenue? 14
What is the revenue / costs for the Co-location and "self-built" data centers? 15
What is the impact of the Server hardware agreements? 16

DETAILS

Generative AI has created one of the most visible and emotional roller coasters in the tech industry, possibly since the inception of computers. Almost on a daily basis there is news flow, articles, or expert commentary that drives re-evaluation as to where, how quick / impactful Gen AI will be on different parts of the technology ecosystem as well as on employment and every day life. The reaction is swings in how Gen AI is perceived (will it drive nirvana or destruction of humanity) as well as which stocks will be winners and which losers.

One of the more consistent areas of controversies is whether all the investment in compute is sustainable and economically valuable or a complete waste. For the hyperscalers which includes Microsoft and Oracle within our joint coverage (as well as Amazon and Google) as well as the neoclouds (e.g. CoreWeave), one of the questions overhanging these stocks is whether their largest customers, and more specifically the AI Labs, will live up to their contractual commitments. When said that boldly it sounds harsh and unreasonable but understandable given:

1. AI still in the early days of adoption and with limited proof points of guaranteed success;
2. A race is ongoing for who builds the biggest / best models and everyone is focused on the models;
3. Most recently concerns surrounding the impact of a slowdown in frontier model developments is adding further concern.

These concerns are negatively impacting the valuations of hyperscalers within our coverage such as Oracle and Microsoft; neocloud vendors within our joint coverage such CoreWeave; as well as a broad range of other tech companies including the other hyperscalers (Amazon and Google[1] Semiconductor vendors such as NVIDIA[2], Broadcom, memory manufacturers; and even investors such as SoftBank[3] (all of which are outside our direct coverage but in turn impactful on our coverage as they are suppliers and partners to all of our covered companies).

[1] Amazon and Alphabet are covered by Bernstein's US Internet team.

[2] NVIDIA and broadcom are covered by Bernstein's US Semiconductor team.

[3] SoftBank is covered by Bernstein's Japan Semiconductors team.

The simple question this note will try to answer that investors are trying to understand is 1st how large / real are the datacenter commitments and 2nd over what period of time. The question we can not completely answer at this time is the financial impact of those commitments on the AI Labs. We believe that one is able to see that:

- The non-cancelable dollar commitments are more reasonable than many believe
- They scale up over time which is natural given the opportunities (Link) and revenues should also scale.
- OpenAI's revenue should be growing faster than their AI capacity commitments over the next year which bodes well given how capacity and the cost of AI capacity scale.

Given the covered companies exposure is predominantly toward OpenAI we will focus in this note on OpenAI's non-cancelable contracts for capacity. Today's note, leverages data we have gathered on OpenAI, the contract commitments and estimates of go-live data and pricing to show that in fact OpenAI, and likely the AI labs in general, are in fact acting reasonable given their expectations of revenue growth and their ability to raises funds both in the private and public markets.

Please note that we have already done exercises of this form looking at Microsoft's leases, including not yet started leases, and future procurement commitments and shown that Microsoft is not acting unreasonable especially considering their historical rate of data build / lease, and they have not committed to a meaningful amount of hardware beyond the current FY.

PORTFOLIO MANAGERS SUMMARY

We are in the midst of a massive investment cycle in data center capacity to meet the current and future needs of Generative AI. At this time most of the investment is still going to training models, but that is starting to change, and we are seeing the early signs of increasing inference usage. This is a very good sign, but there is a long way to go as global annual inferencing spending must be, we estimate reach 5x to 20x the global annual spending on training. Consumer adoption is strong and leading the charge, but enterprise inferencing is still relatively early as legal, regulatory, compliance, risk, business and business process re-engineering moderate significantly adoption.

One of the numerous and constantly changing concerns impacting the tech industry and overhanging the hyperscalers (Microsoft, Amazon, Google, Oracle) and the neoclouds (e.g. CoreWeave) is how much compute capacity is generally being constructed and especially how much is being built to support the AI. In previous notes we looked at Azure (Link) and Oracle (Link).

At least part of the problem is that there are numerous conflicting numbers, both in GW and dollars, floating around. Many have used numbers upward of $1.5T in OpenAI commitments. But those numbers, we believe are more aspirational and very long term and not tied to specific contractual commitments. The more exact data we have come up with are:

- OpenAI stated in October 2025 that "The combined capacity from these five new sites … brings Stargate to nearly 7 gigawatts of planned capacity and over $400 billion in investment over the next three years. This puts us on a clear path to securing the full $500 billion, 10-gigawatt commitment we announced in January by the end of 2025, ahead of schedule."
- Datacenter.com in a July 2026 article said OpenAI "… now expects to spend approximately $750 billion on cloud computing infrastructure through 2030,"

While these two data points are helpful they don't tell the full story, at least as it relates to how investors in the AI Lab's datacenter providers should feel about the stability and longevity of their businesses. Today's note unpacks the available data from numerous sources to try to lay out the spending and map that to the revenue projections. We start with OpenAI and unpacks what we believe are the real contracts (versus announcements of partnerships or discussions that may or may not translate into contracted capacity agreements), contracted datacenter capacity, and estimating the annual spending for each.

One of the key points we argue is that while the company may have signed contracts adding up to $750B through 2030 they may not spend $750B in cash through 2030. The reasons are as follows:

1. We believe that the chips from NVIDIA, Broadcom, and AMD are likely vendor financed and will be paid over a useful life of 6 years. In addition, those contracts are not fixed either in the time horizon in which the chips will be delivered or in fact the quantity to be delivered. OpenAI can simply not buy from one partner and instead leverage another route to capacity.

2. The agreement with AMD and Cerebras include warrants which discount the real cost of these contracts. Depending on the price of AMD stock this could translate into those chips being almost free to OpenAI.

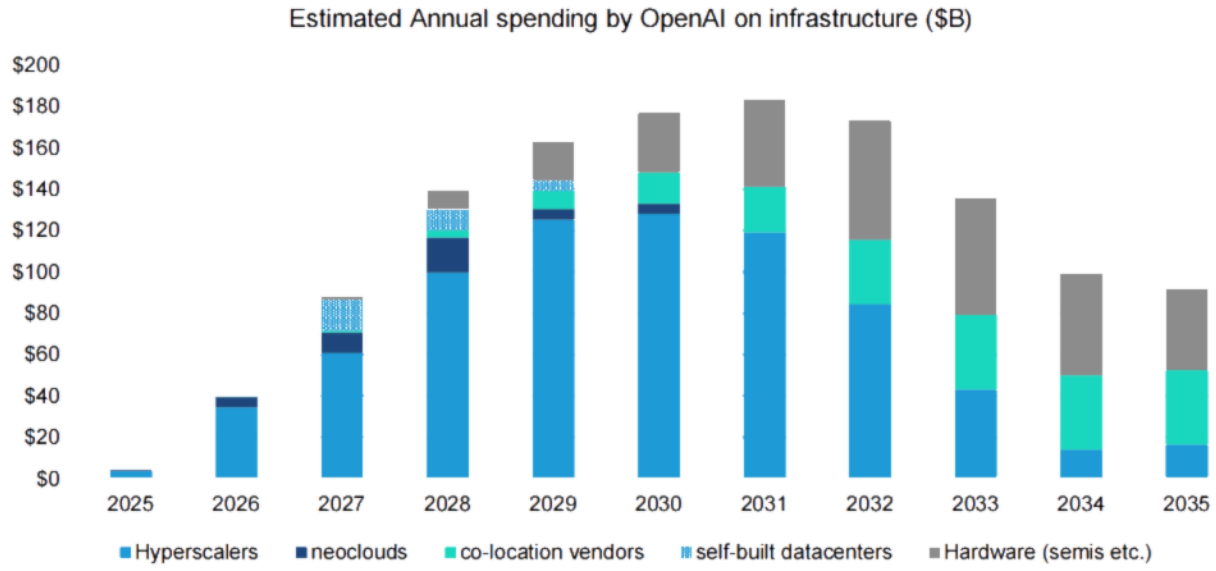
3. The cost of building the Project Camellia datacenter shell is being front-end loaded in our model but since this is a long-lived asset it could potentially be both recognized over time and funded in other ways.

We model (Exhibit 1), if all contracts come in on time and are not delayed or canceled - which is highly unlikely, OpenAI's spending on Infrastructure is estimated by us to be $614B through 2030. Again, the number could increase depending on the timing of expense and contract recognition to meet or beat the company's $750B.

On an annual basis we edge-of-the-envelope estimate that the maximum currently committed spending is $184B in 2031 (and our numbers could be incrementally off in either direction). In addition, delivery timing that many years out can move significantly. Mahesh Thiagarajan, President of Oracle Cloud Infrastructure, stated during Oracle's recent Financial Analyst Meeting "In my mind, infrastructure is effectively on time if it's a quarter or a couple of quarters around what you anticipated, again". Timing is difficult and likely as not some of this will push out to 2032.

Thus, the question investors need to consider is whether OpenAI, or any AI Lab, can ramp revenue, and we believe they can (but not necessarily everyone) given the breadth of their AI monetization strategies (Link).

EXHIBIT 1: **Estimated annual contracted spending by OpenAI on infrastructure tops out at $184B in CY31 and includes both leases and assets**



Source: Bernstein estimates

The total size of the commitments are important and show how much of global AI spending is being allocated to not just OpenAI but the AI Labs in general (if we extrapolate to their peers). But for the purposes of this note the important point is how the spending will grow and likely relatively in-line with their revenue.

OpenAI has said that "Looking back on the past three years, our ability to serve customers—as measured by revenue—directly tracks available compute..." "... we firmly believe that more compute in these periods would have led to faster customer adoption and monetization." (OpenAI Link)

While the data is sparse on OpenAI's revenue and revenue growth rate, two recent, and supposedly independent articles have stated that:

- OpenAI has told investors its annualized revenue for September was almost $50 billion (Reuters Link) with total revenue growing 77% YoY

- OpenAI told investors that it hit roughly $50 billion in annualized revenue at the end of September (CNBC)

A $50B revenue run rate at the end of September and total revenue growth 77% YoY would argue that next year's revenue would be > $80B which would exceed our estimate of CY27 infrastructure spending if we take out the asset investment of $15B (which we are front-end loading) building the Project Camellia datacenter shell. Note that OpenAI has beaten their revenue targets to date.

With infrastructure spending growing ~58% YoY in CY28, based on our modeling to $139B or $129B excluding the project Camellia datacenter shell build-out if OpenAI can grow 60% or faster revenue will continue to outgrow infrastructure spending and the trend accelerates, baring more datacenter contracts.

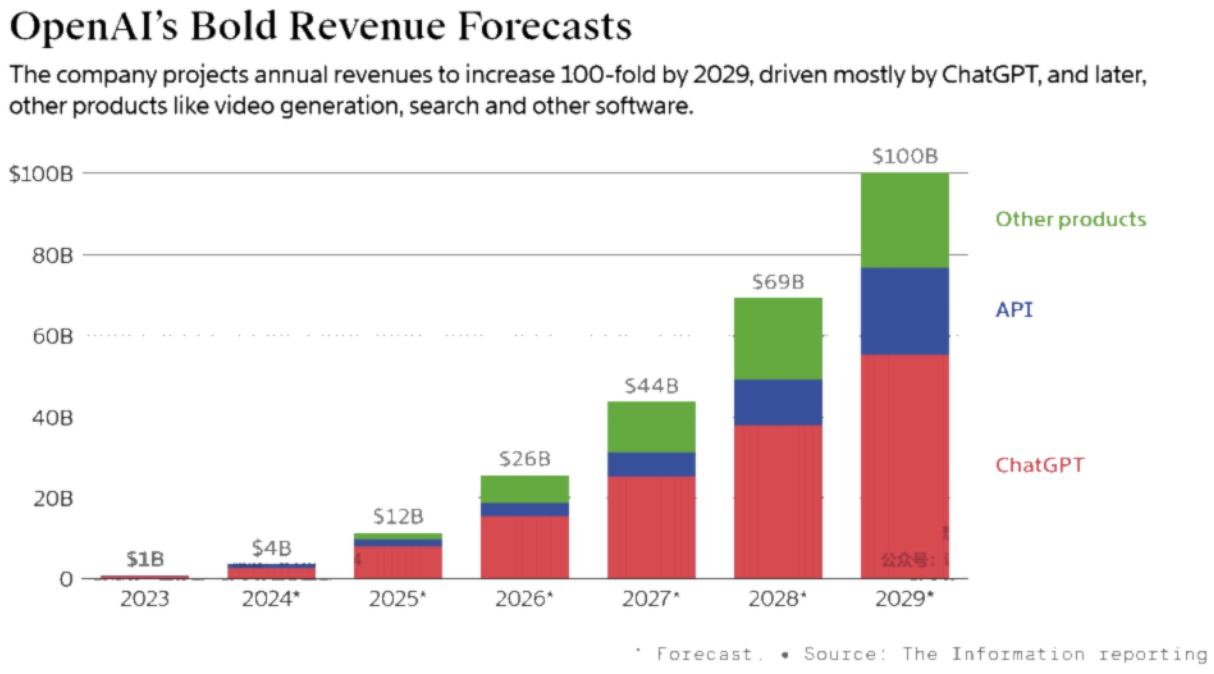
To put this in perspective we can look back to the data that was published in The Information in 2024 (Exhibit 2) where they estimated CY26 revenue of $26B and $44B in CY27. OpenAI beat their targets and could deliver double the revenue in CY27 that they expected in CY24.

In other words, while one has to believe they can sustain strong double-digit growth even excluding the company's fund-raising to date and the future IPO it seems very feasible that the company will not only be well positioned to pay all their infrastructure costs but improve gross margins.

Now it is important to remember that a good portion of the current AI infrastructure is used to build models and those expenses could be capitalized and expensed over time and thus margins would be better than these numbers would suggest.

While we are not focused on OpenAI's business economics we believe that as it relates to our Oracle, Microsoft and CoreWeave the OpenAI contracts are solid, and investor concerns should decrease. Again, this very rough numbers that have driven our infrastructure cost build out and revenue discussion, but directionally this makes sense given management and partner comments.

EXHIBIT 2: **In CY24 OpenAI projected revenue of $26B in CY26 and $100B**



Source: The information

WHAT DOES THIS MEAN FOR OUR COVERED COMPANIES?

Overall we believe that investors should watch but not be overly concerned about the quality of the contracts between OpenAI and their partners, especially Microsoft and Oracle. We are not modeling in contract renewals after the initial contract duration and we expect the smaller / shorter contracts may not also be renewed. But the contracts with Oracle and Microsoft will most

likely be renewed at least one more time.

- **Oracle:** Initially the announcement of the $300B of RPO from OpenAI drove the stock up, but then investor concerns kicked in surrounding not just the OpenAI agreements but all the AI datacenter agreements : 1) how much Oracle would need to raise; 2) the ROI for Oracle; 3) Customer exposure and 4)execution risk. Oracle at their FAM put to bed concern #1 and gave lots of details on how profitable their AI datacenter business will be. Today's note while it does not completely answer the OpenAI exposure question should show that OpenAI management is ramping capacity over time and in-line to slower than revenue (especially if we expect some capacity to delay).

- **Microsoft** has a complex relationship with OpenAI as an investor, leveraging OpenAI models, right to use OpenAI IP; and having a commitment to deliver $250B in capacity (CPU and GPU) to OpenAI over time. Each tranche of capacity is a six year contract which equates to the useful life of the hardware and similar to Oracle ramps over time. Given the ramping nature of OpenAI's capacity additions, the likelihood some contracts will delay / be canceled and Microsoft's large global footprint we suspect that Microsoft is less likely to see delays impacting revenue than the neoclouds and co-location vendors and with the data in this note we feel confident that Microsoft will see that OpenAI revenue on-time and at good unit economics.

- **CoreWeave** has signed up to ~$22.4B worth of contracts with OpenAI: an $11.9B deal signed in March 2025 (+$350M equity stake), another $4.0B in May 2025, and up to another $6.5B in September 2025. Based on the language, these are caps, not committed revenue - the contracts have commenced, but it's not yet clear how far along the contracts are, the exact expiration dates, or the power capacity. We have written that we don't have concerns about CRWV's backlog, though we are skeptical of their ability to renew these deals or sign incremental deals at strong economics given the proliferation of competitors.

KEY TAKEAWAYS

Our key takeaways from the work are as follows:

- The total size of the commitments are important and show how much of global AI spending is being allocated to not just OpenAI but the AI Labs in general (if we extrapolate to their peers). But for the purposes of this note the important point is how the spending will grow and likely relatively in-line with their revenue.

- OpenAI may have committed by 2030 to $750B but that does not mean they have paid $750B by 2030. We are modeling $614B in payments, but that is based on every single contract delivering on time, which is highly unlikely.

- In addition, the final dollars spent can be moved up or down by OpenAI based on real requirements. For example, the real number could be lower, either because of decisions by OpenAI((e.g. operating efficiency and the shift to platform drives margins up and capacity requirements down).

- **Hardware:** The agreements with NVIDIA, AMD, and Broadcom do not have, we understand, fixed commitments either in terms of quantity or timing which could meaningfully decrease spending in CY28 and beyond. The delivery dates seems to be quite flexible and even the quantity that will be delivered.

    - There have been discussions of Google TPUs being used by OpenAI but there is no definitive data or commentary as to the agreement but there was an announcement that Oracle will be supporting TPUs and Oracle also announced at their FAM (link) that they will be leveraging bring-your-own-hardware to lower their capex requirements. It is possible that the OpenAI / Google agreement is via Oracle datacenters.

- **Project Camellia** is OpenAI's self built / managed datacenter which we would argue is a test-bed for OpenAI rather than the start of a massive build out by OpenAI of their own datacenters. Note that the go-live of Project Camellia could be delayed if so desired by OpenAI to time capacity with demand.

- **Capacity deliveries could be delayed:** We have seen projects delay and this could delay both capacity and spending by OpenAI. Construction delays are normal in general (e.g. office buildings, apartment complexes, manufacturing facilities and datacenters). We would not be surprised to see a portion of the later datacenters delayed by months or quarters and some canceled. Thus, the real dollar spend, based on current commitments is likely smaller than either the company's $750B or our $614B and that is fine since we know that OpenAI and its major partners such as Microsoft and Oracle have built this into their models.

Overall, we believe that this very edge-of-the-envelop analysis would argue that so long as one believes that OpenAI will be successful in growing their business, and we believe that these revenue streams for our customers are relatively secure.

WHAT HAS OPENAI CONTRACTUALLY COMMITMENT TO?

As mentioned in the Portfolio Managers Summary investors do not have a have clean picture of what the AI Labs have contractually committed which is creating overhangs not just on the valuation of their datacenter providers but also creating concerns in the debt market which is impact the interest rates and availability of debt financing to build datacenters.

The fact that the data is sparse is not surprising, The AI Labs are growing fast, and their needs are changing and not just increasing capacity requirements but also the availability of lower cost Inferencing chips and higher cost components. Fundamentally, we believe that the AI Labs are driving the build-out of more capacity than they will need which will force supply and demand to normalize something we expect could happen in 2–3 years.

Herein lies one of the key areas, we believe, of misunderstanding. Numerous dollars amounts have been thrown around in the press about how much the AI Labs have committed to as well as over what period of time.

There are in fact 3 main factors that must be considered:

1. **Cancelable or non-cancelable:** What is the contractual commitments and more specifically what are the non-cancelable commitments or in some cases what are guaranteed vs potential (which we see with both data center and hardware) . We have seen numerous cases of "contracted agreements" changing and presumed contracts in fact being canceled (see below for examples)

2. **Contract duration** is also very important (e.g. it is a 3 or 5 or 6 year contract for cloud? A 15-20 year contract for colocation capacity?) as it spreads the available capacity and the "comittment dollars over time.

3. **Go-live data**: Data centers go-live over a period of time (e.g. the OpenAI agreement with Oracle was announced as a contract with capacity going live between FY26 and FY30) . In addition, within many data center projects there are multiple phases that can span months or years.

In this note we take a 2-pronged approach to modeling how OpenAI's data center costs will layer in:

- **Bottom up:** Leveraging our datacenter data (US Industrials & Tech: The Data Center Project Pipeline - Capacity, Construction & Cancellations) and the data we have gathered from the companies that we cover to build a bottom-up analysis of OpenAI's commitments.

- **Top down validation:** Comparing our data to management commentary including the recent statement by OpenAI allows to sanity check our estimates

SETTING THE STAGE FOR THE DISCUSSION

Before going into the specifics of the work we want to step back and put certain ground rules in place.

- There is a belief that every announcement and every news article about a potential datacenter build is another non-cancelable contractual commitment but in fact we know that some are guaranteed non-cancelable while others can and in fact have been canceled or never turned from some form of a letter of intent into a non-cancelable contract. Two examples come to mind, and we think help explain why there is so much confusion about what both the AI Labs and the hyperscalers are doing.

    - There was an agreement that OpenAI had with Oracle for a 2.1 GW expansion of the Oracle facility in Abilene Texas (adjacent to existing Oracle / OpenAI facilities). Some in the news media included this information as contracted / non-cancelable but in fact was canceled, and the builder (Crusoe) had to find other tenants, which took ~2 weeks. OpenAI

reversed its planned 2.1 GW Abilene expansion, redirecting that capacity elsewhere and ceding the adjacent 900 MW site to Microsoft/Crusoe — a useful data point on contract flexibility[4].

- There was an agreement, originally announced in July between Nscale and OpenAI and industrial firm Aker ASA for a datacenter in Norway (Stargate Norway). When OpenAI pulled out of the agreement (see CNBC article) Microsoft again stepped in and will be using that facility.

- Please note that we are deliberately separating pure chip-supply deals such as those OpenAI has signed with NVIDIA (up to 10 GW / $100B), AMD (up to 6 GW warrant deal), and Broadcom (up to 10 GW custom accelerators) from the AI data center agreements. The chip deals are silicon commitments rather than data center facility contracts and may either be:

  1. Used as bring-your-own-chips and rolled into an AI data center agreement already counted (e.g. Oracle has announced recently agreements that have increased their RPO without requiring meaningful cash commitments from Oracle some of which leverage likely bring-your-own-chip) and we don't want to double count.

  2. Used at some later date since the contracts do not seem to have contracted delivery schedules, at least based on the news flow to date.

  3. Used in the co-location facilities and in Project Camellia as we are modeling in this note.

- Note that while SQ FT and number of data centers are interesting to help triangulate they are less relevant when we are trying to figure what are their contractual dollar commitments and whether they can reasonably meet those requirements

- There are a number of small (100MB or smaller) agreements in India and elsewhere that we are leaving off as the quality of the data on these sites is low and the relative size is small

HOW MUCH REVENUE CAN A MW GENERATE?

Unfortunately the data is sparse for most data center contracts with many discussing megawatts but often not the revenue generated from the data center or contract durations. Therefore, many use megawatts as a proxy for revenue generation potential of a data center facility. In reality while MW is a good data point it is not a great proxy for either fully built out cost or revenue generating capacity, and we would argue that as we shift to inferencing which leverages more CPUs and use more inferencing chips the range of results will vary even more.

The variability of revenue per active MW already varies significantly (Exhibit 3). At the high end of the range are the more mature neocloud vendors and Oracle (which has said they can capture > $10B of revenue per GW). As the offerings / capabilities / operational efficiencies change once one shift from full AI data centers to co-location vendors to emerging AI infrastructure vendors with more limited capabilities and performance characteristics the revenue per MW decreases substantially.

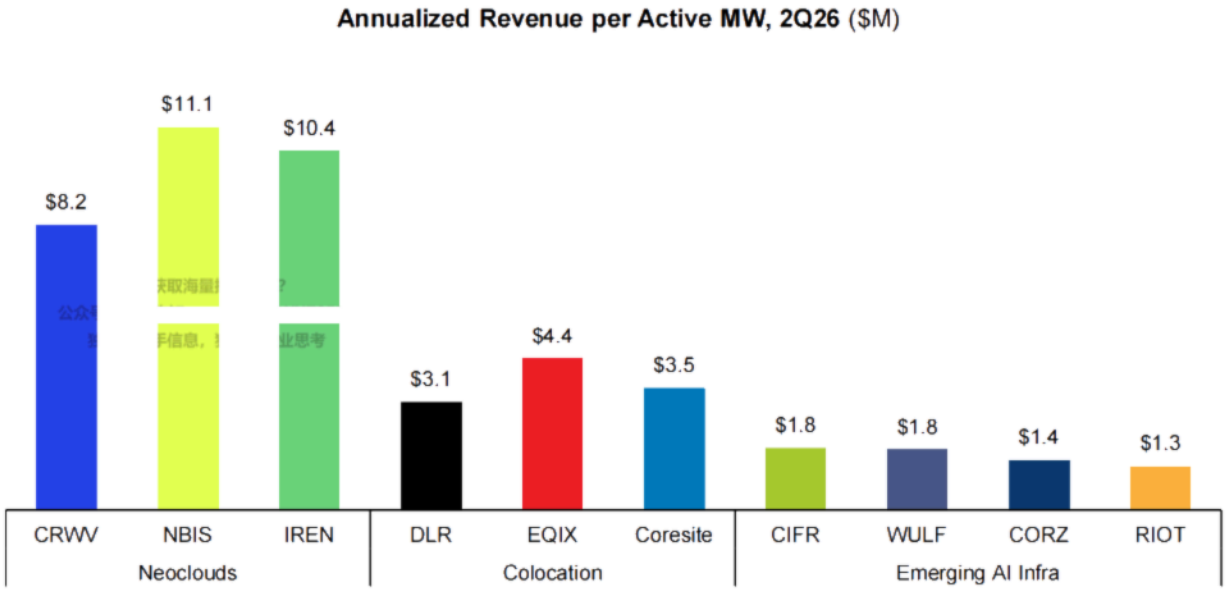
Please note that while hyperscalers and neoclouds are renting out fully built-out data center capacity including servers the co-location vendors are renting out, in many instances, data center shells often with power and cooling but no servers. The servers are separately arranged for by the client as discussed below.

We would argue that as one moves from pure AI Infrastructure-as-a-Service and the rental of compute to offering more platform functionality the revenue per GW will increase substantially, the moat surrounding the revenue for the hyperscaler / customer (and not necessarily the data center provider) will increase, and the long term stability / stickiness will increase for both the provider and the customer.

While we have some data on the economics of the data center we unfortunately do not have for all contracts, and thus we are leveraging this data in some instances.

[4] It took about 2 weeks for Cruseo to find a replacement customer for the facility.

EXHIBIT 3: **Annualized Revenue per Active MW ($M)**



Annualized 2Q26 figures; emerging AI & IREN figures are calculated on a per-contract basis.
Source: Company filings, Bernstein analysis and estimates

HOW CAN WE BREAK APART OPENAI'S COMPUTE SPENDING INTO CATEGORIES?

OpenAI's approach to meeting their compute capacity requirements has changed as the business has grown and matured.

- Initially they leveraged Microsoft for both training and inferencing compute with the training spending being paid via Microsoft's investment in OpenAI and all the OpenAI models running exclusively at Microsoft. Note that while inferencing was recognized within Azure revenue, training was paid for with credits and delivered by 3rd party datacenter providers
- Over time OpenAI has expanded substantially the breadth of methods they use to gain access to compute resources initially for training and more recently for inferencing (Quick Take: Initial thoughts on the newest update to Microsoft / OpenAI partnership). These now include:
  - Hyperscalers: Microsoft, Oracle, Google[5], and Amazon
  - Neoclouds: CoreWeave, Crusoe, Nscale, Cerebras
  - Co-location arrangements: Crusoe, Vantage, and a couple of large, announced U.S. campuses (Ports, Milam)
  - Server hardware procurement contracts: NVIDIA, Broadcom, and AMD[6]
  - Own data centers

Each of these has a different impact on both capacity (both in terms of MW and compute) and cost.

[5] Alphabet and Amazon are covered by Bernstein's US Internet team.

[6] NVIDIA, Broadcom and AMD are covered by Bernstein's US Semiconductor team

HOW MUCH REVENUE DO THE HYPERSCALERS CAPTURE?

OpenAI has multiple relationships with the 4 largest hyperscaler and each relationship is a bit different. In today's note we will focus only on the known contracted commitments for use of the hyperscaler's datacenter facilities for inferencing and training. Note with partnership agreements in which OpenAI platform software (Global Software: How do the AI Labs and neoclouds impact Cloud (IaaS / PaaS)) runs at the hyperscalers, we do not know that whether that usage is included in these commitments or not.

We have assembled our estimates of the revenue that OpenAI will be generating for the 3 hyperscalers (Microsoft, Oracle, Amazon) in Exhibit 4. As can be seen in Exhibit 4 spending with the hyperscalers ramps overall through 2030 and then naturally decreases unless the contracts are renewed. We would expect that they will be renewed since the hyperscalers are both key go-to-market partners as development platforms (AI and non-AI) require the full breadth of development, security, monitoring, and management tools and when considering the enterprise opportunity at a minimum AI is going to run where the data and applications reside. This can be seen in the OpenAI Amazon partnership announcement:

1. Amazon will run Frontier, OpenAI's "stateful" development runtime environment, on AWS. For more information on the development platform opportunity for the AI Labs see Global Software: How do the AI Labs and neoclouds impact Cloud (IaaS / PaaS). Note we do not know if the capacity that Amazon will use for OpenAI models, OpenAI Frontier and other future platform services will be in addition to or included within the capacity we are modeling. If in addition, then OpenAI will be capturing that revenue with no capacity / CAPEX similar to how they were capturing revenue from Microsoft Azure usage of OpenAI models.

2. This required changes of the Microsoft /OpenAI agreement to allow OpenAI to run their models on AWS (and elsewhere).

3. It also included a commitment for OpenAI to use 2GW of Amazon Trainium chips and Amazon to invest $50B in OpenAI.

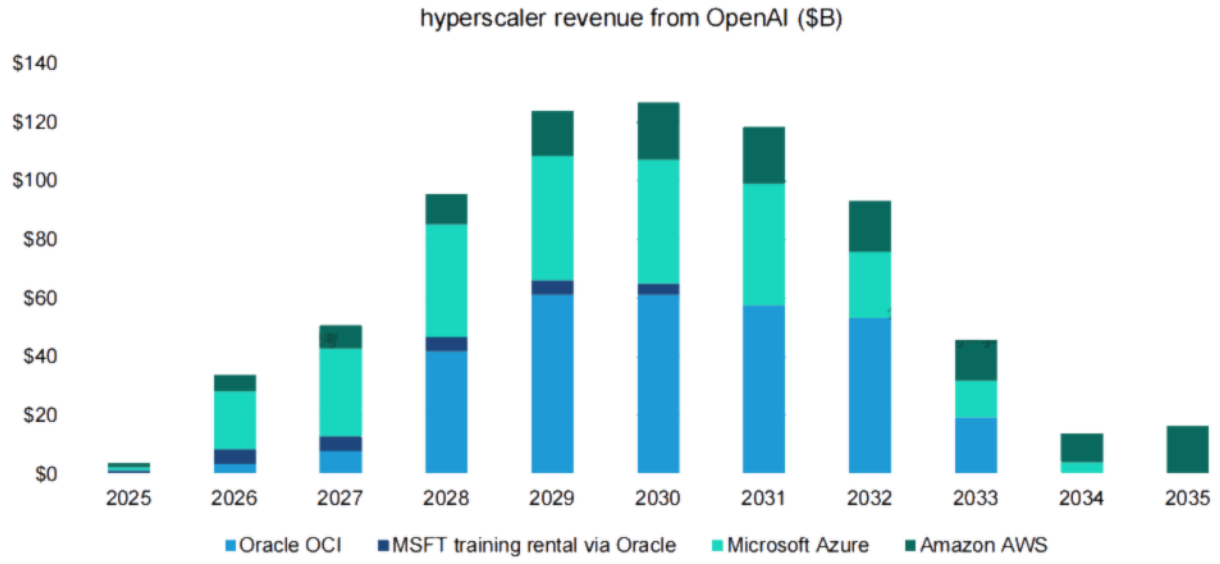
In summary, to make Frontier successful at Amazon OpenAI models needed to run on Amazon (where the data and apps reside).

Data on the arrangements between Google and OpenAI is spare. There does not seem to be a contract with OpenAI for capacity, even though this has been discussed numerous times in across news services and there was supposedly an agreement for capacity signed in May 2025. With the change in the agreement between Microsoft and OpenAI as to were inferencing could run and the announcement that OpenAI models would run at Google it is possible that capacity is being used by Google to operating OpenAI models for Google's customers and OpenAI has no dollar commitments involved.

Similarly, there does not seem to be an agreement for TPUs, even though that has been discussed in the news and OpenAI has said they are testing Google's TPUs.

For modeling purposes, given the lack of data, which, if the deals were significant would have been mentioned in Google's financial filings, we are excluding Google from the list.

EXHIBIT 4: **The revenue generated by OpenAI for the hyperscalers increased through CY30 and then declines**



* The MSFT training is paid by MSFT to Oracle leveraging the OAI investment credits. This is revenue to Oracle and may or may not be included in the $300B
Source: Bernstein estimates

ORACLE

The vast majority of Oracle's relationship with OpenAI is via a $300B agreement that is composed of multiple tranches of capacity agreements. Each agreement is 5–6 years in length signed in 2025 and starting, we believe, in CY26 that go live through 2030. Based on the data we have compiled on the specific datacenters would argue that the bulk of the capacity could come on-line in CY28 (Oracle FY28/FY29).

Note that these agreements are not the only agreement under which Oracle is supplying capacity to OpenAI. There is a prior agreement between Microsoft and Oracle under which Oracle built out their Abilene Texas datacenter for OpenAI use. Microsoft has been paying Oracle using the Microsoft /OpenAI investment credits, and we expect that will continue at least through this year. Thereafter, there are multiple scenarios, but we would expect that OpenAI will be using those facilities for a full 6-year period of time, possibly under the $250B OpenAI/Microsoft commitment.

It is important to note that the initial training capacity between OpenAI and Microsoft, under Microsoft's investment in OpenAI have no cash requirements from OpenAI, they use the investment credits. Therefore, when considering these facilities (including Abilene 1 and likely a prior agreement of Microsoft with CoreWeave) required no cash from OpenAI until they have used up the original 13B+ investment.

All of these agreements, we understand, including the Abilene 1 facility with Microsoft are non-cancelable.

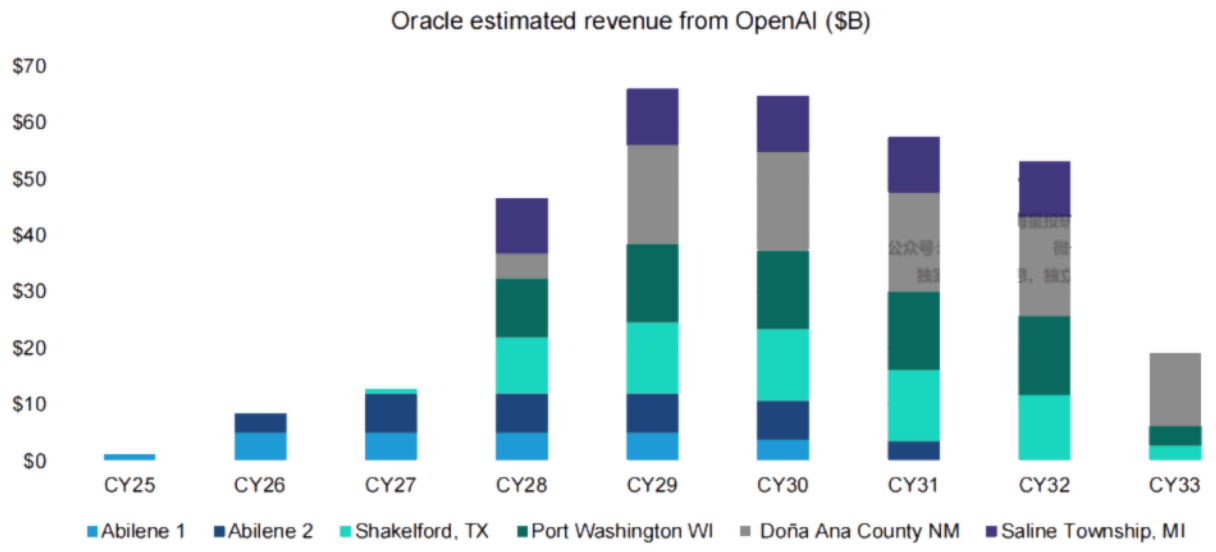
We attempt to layout Oracle's revenue stream from OpenAI through CY33 (Exhibit 5). For the purposes of this note we are only assuming the current agreements and no renewals, even though we expect that all of these facilities will continue to be used by OpenAI for at least another 5-6 re-commitment, likely with new hardware (as the existing hardware will have roughly hit their 6-year useful life).

For those that follow Oracle, and we recommend anyone reading this note should be following and owning Oracle stock, Oracle will be capturing slightly more than the $300B they added to RPO.

Our estimates of Oracle's revenue stream spread over time (Exhibit 6) from the $300B in RPO announced last year. The reason for the $312B in our estimates and the company's disclosure is that Oracle, like Microsoft discounts revenue streams years out when adding to RPO. These estimates have Oracle capturing 57% of their OpenAI contract value through their FY30 which is a bit less than we thought they might.

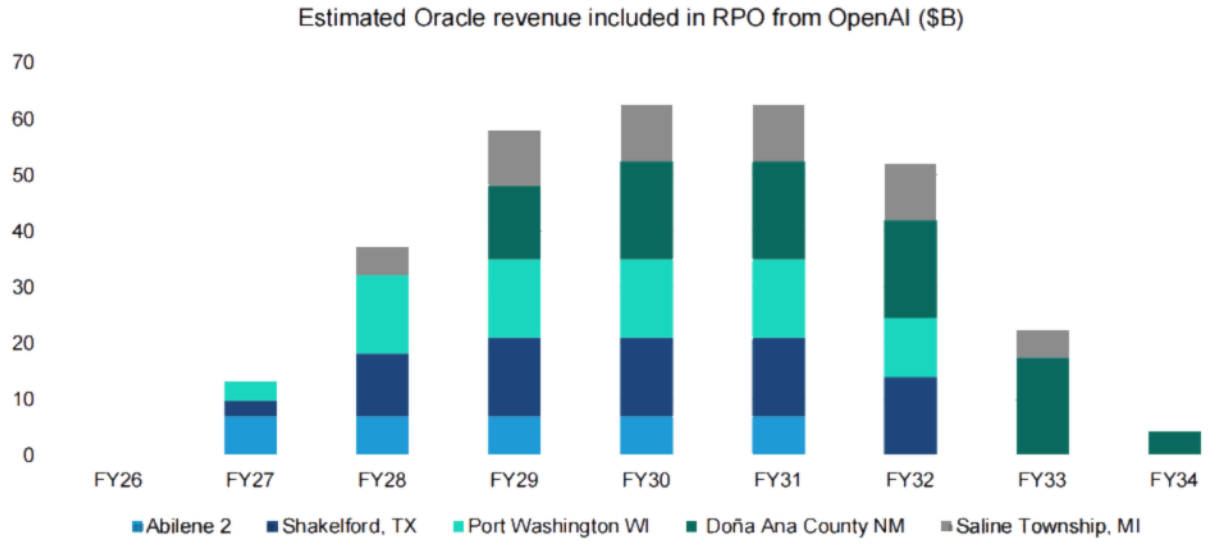
EXHIBIT 5: **Oracle's revenue from OpenAI, including the prior MSFT and OAI contracts in Abeline TX, layer in through CY33 (Oracle FY34)**

Note that Abilene 1 is a contract between Microsoft and Oracle for the benefit of OpenAI.



Source: Bernstein estimates and analysis

EXHIBIT 6: **Spreading Oracle's revenue from OpenAI we get ~$300B in revenue of which 57% is captured through FY30**



Source: Company reports, Bernstein analysis & estimates

MICROSOFT

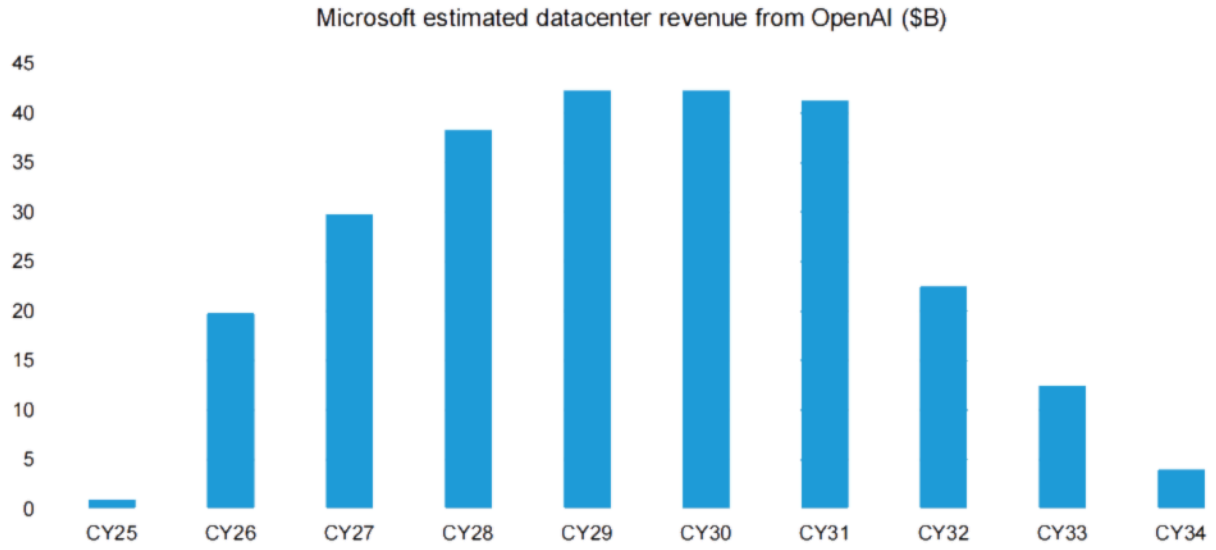
The OpenAI / Microsoft relationship is complex. Microsoft is an investor; a user of OpenAI models both in their apps and offered on Azure (now at no charge through 2032), and a supplier of Cloud capacity. Microsoft, in their investments in OpenAI supplied OpenAI with credits which are used for datacenter capacity for OpenAI training. Microsoft does not include that in their revenue or CAPEX. OpenAI initially was only supplied of compute capacity, but that has change over the last year as seen in all these other agreements.

In Microsoft's Oct 2025 Blog they stated that "OpenAI has contracted to purchase an incremental $250B of Azure services, and Microsoft will no longer have a right of first refusal to be OpenAI's compute provider." Note that the $250B is incremental to prior commitments. We understand that Microsoft's contracts with very large AI customers, including the AI labs, are in tranches of capacity running for the useful life of the hardware which is 6 years.

While the datacenter in Abilene Texas datacenter that Microsoft contracted with Oracle for was initially solely for training, we are assuming that today it could be used for training and inferencing and are including the spending in our calculations to cover spending > $250B.

EXHIBIT 7: **Microsoft's estimated revenue per year from datacenter capacity from OpenAI**

This does not include revenue share income



Source: Bernstein analysis & estimates

AMAZON

In addition to the partnership with OpenAI to run OpenAI's development platform on AWS discussed earlier, there are two datacenter capacity commitment contracts:

- A $38B 7-year agreement signed in 2025
- An additional $100B 8-year agreement signed in 2026 as part of the OpenAI / Amazon partnership. This agreement included the use of as much as 2GW of Trainium chips at AWS.

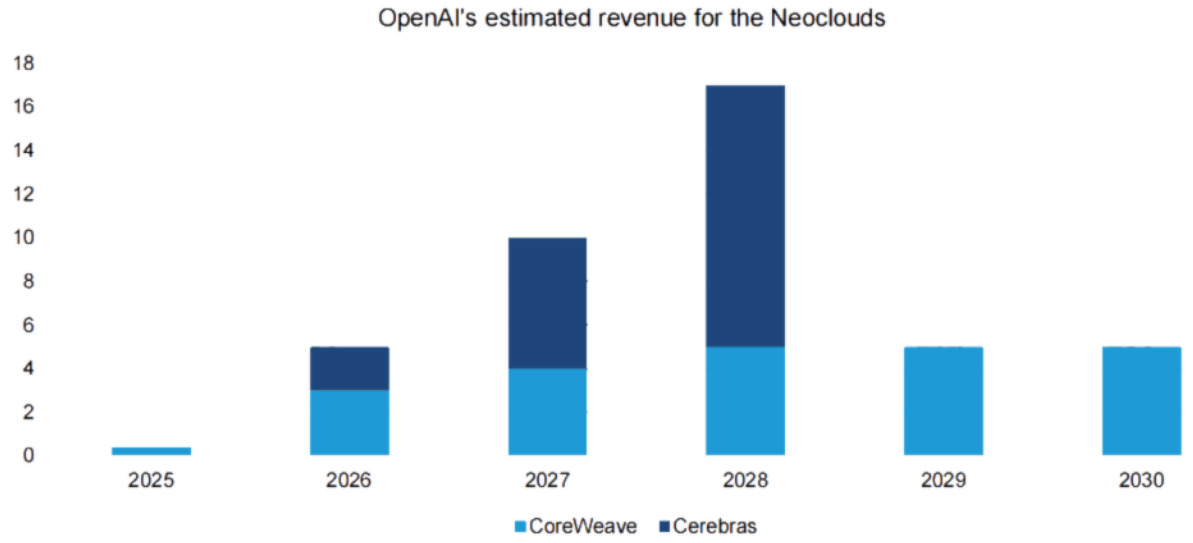
WHAT DO THE NEOCLOUDS CAPTURE IN REVENUE?

OpenAI has multiple agreements with both CoreWeave and Cerebras for datacenter capacity (Exhibit 8). We believe those agreements extend for CoreWeave through 2030 and have a ~5 year duration, while the Cerebras agreement is only for 3-years and would seem to us to be bridging capacity until the capacity comes on-line from Oracle, co-location vendors, etc.

These contracts are similar to the hyperscalers in that they include not just data center shells but also the requisite hardware for training and inferencing. The difference, we would argue is the neoclouds lack a wide range of functionality and capabilities that are necessary to sell more than raw compute to a single customer per datacenter.

人在草木间

EXHIBIT 8: **OpenAI's contracts with the neoclouds (CoreWeave and Creebras) extended through CY30**



Source: Company reports, Bernstein analysis and estimates

WHAT IS THE REVENUE / COSTS FOR THE CO-LOCATION AND "SELF-BUILT" DATA CENTERS?

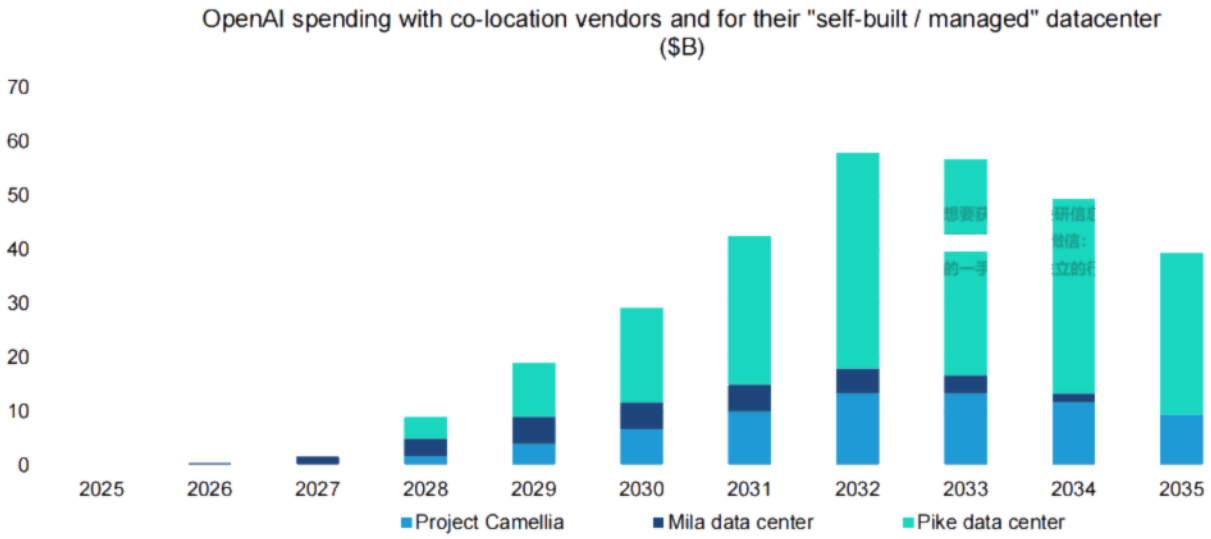
OpenAI has 2 data center agreements with a large provider building data centers specifically for OpenAI. Under those agreements, the developer is building the shells and arranging the power, but OpenAI is likely going to supply the hardware (see below). We expect the hardware will come from the hardware agreements with NVIDIA, Broadcom and AMD plus likely other agreements for the full hardware stack required for AI datacenters (AI and CPU servers, storage, networking equipment, racks). These data center facilities will come on line starting in CY27 but predominantly in CY28 and beyond.

Please note that we believe that the Cerebras agreement includes some form of Cerebras warrants given to OpenAI as part of their arrangements. While exact details are sparse we would expect that the value of these warrants could meaningfully offset the cost of Cerebras data centers.

In addition, to the two co-location data centers OpenAI is building their own data center. Project Camellia is the first OpenAI managed data center of size. OpenAI, we understand, will be project managing the construction, funding the build-out including the infrastructure and power, and then operating the data center. Project Camellia is a planned $20 billion, 4.4 million sq ft AI data center campus in Effingham County, GA, with 3.2 GW of power from Georgia Power, slated for phased delivery 2028–2032. While the company has not yet discussed the hardware that will be going into the data center we are assuming that it will be filled with GPUs or ASICs from NVIDIA, Broadcom or AMD (see below).

EXHIBIT 9: **OpenAI's estimated spending on co-location facilities and "Project Camellia" which they are coordinating the building and then managing**

The economics in this chart are based on our estimate of the total cost of these datacenters including the hardware discussed later in hte note



Source: Company reports, Bernstein analysis & estimates

WHAT IS THE IMPACT OF THE SERVER HARDWARE AGREEMENTS?

The agreements that have been verified are with NVIDIA, Broadcom and AMD. As with all the agreements broadly we discuss in this note the data is sparse on these agreements, but we are assuming that the delivery time horizon is not necessarily fixed and that the semiconductor companies are vendor financing these purchases or in some cases using the chips as their investment in OpenAI. There are partnerships with memory vendors, but we suspect that similar to Microsoft and Oracle, there is no fixed comment on OpenAI's part for quality and delivery dates.

We understand that the hardware agreements have the following rough terms:

- NVIDIA has a 10GW semiconductor agreement with OpenAI with the 1st GW committed and the rest coming over time (likely a 5-year period of time)but not necessarily contractually non-cancelable commitments.
    - NVIDIA is also back-stopping neocloud datacenter agreements which include those listed above and used by OpenAI. They are also leasing datacenter capacity from Oracle and others.
- Broadcom has more of a co-development agreement for 10GW again over time (likely a 5-year period of time) but there are questions whether OpenAI has comitted to these chips directly. We believe that they are contractually non-cancelable commitments to purchase. We understand that OpenAI is expected to take 1.3GW of Broadcom's 1st generation AI ASIC (Jalapeño) in 2027, and 5GW of the 2nd gen (Serrano) in 2028 but how hard and fast these agreements are is a question.
- AMD agreement is for 6GW of semiconductor chips with zero-cost warrants that vest based on AMD's stock price, in tiers, and the actual purchase of the chips. Note that given the pricing of the warrants and the chips, one could argue that effective cash requirement from OpenAI is small. We believe that OpenAI can purchase or not the chips over the period of time, but the warrants are a very strong incentive to purchase.

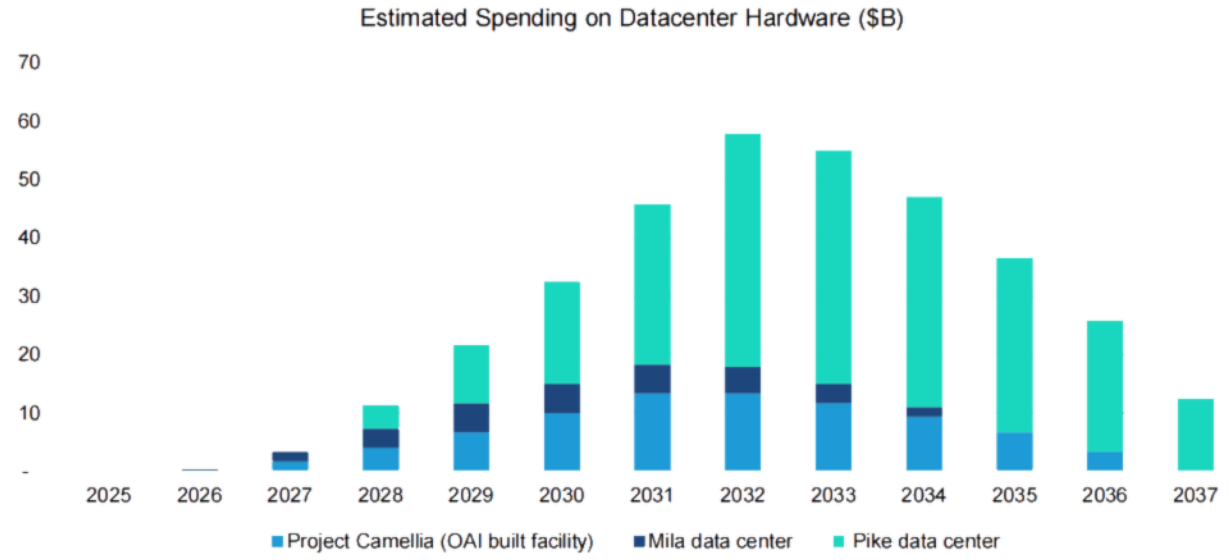
We are modeling for completeness' sake:

- That these commitments translate into hardware spending to fill the co-location and Camellia datacenters (Exhibit 10)
- That OpenAI has arranged vendor financing for at least the GPUs and AI ASICs spread over a useful life of 5–6 years.

- If one were to consider OpenAI's minimum cash commitments then one could exclude most of these since they are cancelable and could include the AMD agreement with a very low cash value given the value of the warrants.

Periodically we have heard about an agreement (e.g. The Information - Google and OpenAI) between OpenAI and Google under which OpenAI would purchase TPUs but there are no details available to date. Whether the TPUs would be used as part of one of the existing agreements is not known. One possibility is Oracle, which announced recently they would be offering TPUs would be for OpenAI or even via a Bring-Your-Own-Hardware agreement (Oracle AI World and FAM: Was it everything one could want and more?). Given the lack of details we are not including any Google TPUs in our modeling at this time.

EXHIBIT 10: **Our estimates for OpenAI's hardware spending CY25 to CY27 for the 2 co-location datacenters and the 1 datacenter OpenAI is leading the building of**



Source: Company reports, Bernstein analysis & estimates