

# TOP MARK CAPITAL MANAGEMENT LLC

Report For the Quarter Ended

June 30, 2023

416 13th Street

Suite 200

San Diego, CA 92101

E: [contact@topmarkcapital.com](mailto:contact@topmarkcapital.com)

T: +1 (646) 470-3016

# INDEX

|  |    |
|--|----|
| INDEX.....   | 2  |
| LETTER TO OUR PARTNERS AND PROSPECTIVE PARTNERS.....           | 3  |
| PARTNERSHIP LETTER: TOP MARK CAPITAL PARTNERS LP.....          | 5  |
| PARTNERSHIP LETTER: OP MARK HEALTH PARTNERS LP.....            | 13 |
| SELECT ARTICLES, OBSERVATIONS & COMMENTARY.....                | 18 |
| NVIDIA’s CORNERSTONE for UNMATCHED DOMINANCE in AI (CUDA)..... | 19 |
| FORGET BLOCKCHAIN: GENERATIVE AI IS THE REAL Web3.....         | 25 |
| IMPORTANT DISCLAIMER.....                                      | 28 |
| REFERENCES.....  | 29 |



We are pleased to present this update on the performance of our Partnerships: Top Mark Capital Partners (TMCP) and Top Mark Health Partners (TMCM). TMHP, our newest partnership, made its first investments in Q2. You can rest assured that both partnerships follow the same rigorous research process (detailed in our [2022 Annual Letter](#)). The only distinction between the two partnerships is that TMHP is exclusively investing in businesses that relate to health and health care.

In the pages that follow you will find performance and commentary for each partnership as well as a few essays that offer unique and contrarian perspectives (we try not to waste your time or ours by regurgitating what news and research outlets have already produced). We trust this will offer valuable insights into our strategic thinking and investment approach in the current economic environment.

For those seeking long-term outperformance, our approach has demonstrated quite satisfactory returns (TMCP has compounded at an annualized rate of 23%, net of fees since inception). Feel free to reach out to us directly to discuss the partnerships in more detail, and as always we welcome referrals to like-minded individuals.

As a partner, you can expect from us quarterly letters like this one in which we discuss our investment thoughts. In our opinion these documents, once digested, provide the information needed to form a judgment about what and how we are doing. From Top Mark's administrator (Formidium Fund Management Services), you can expect a quarterly statement of your account as well as annual audited financial statements. Let us know if you are not receiving all you should.

Top Mark's focus is genuinely long-term, and more regular reports, daily, weekly, monthly or otherwise, are likely to be of little value to you, and may even be counterproductive. With that said, Jason and I can be found every week on the Telltales

podcast if you'd like to hear from us more frequently. You can sign up for the Telltales weekly newsletter at [telltales.us](http://telltales.us) to get a notification when new episodes are released.

If we are to have a sustainable comparative advantage, it will come from two sources... 1) your manager's capital allocation skills, and 2) from the patience of our partners.

When it comes to patience, we are fortunate to be surrounded by like-minded, long-term investors. Only by looking further out than the short-term crowd can we expect to beat them. We are an investment partnership, not a hedge fund, and the relationship we seek with our partners is different.

**In the event that you share our temperament and perspective, we invite you to consider subscribing to one of our Partnerships.**

As always, we thank you for your confidence and value your support.



Mike Nicoletti

Jason Wallace

# TOP MARK CAPITAL PARTNERS LP

Partnership Letter

For the Quarter ended June 30, 2023



## PARTNERSHIP LETTER & PERFORMANCE UPDATE<sup>1</sup>

| To June 30, 2023                      | TMCP % Net<br>Return <sup>2</sup> | S&P 500 Total<br>Return <sup>3</sup> |
|---------------------------------------|-----------------------------------|--------------------------------------|
| One year                              | 24%                               | 20%                                  |
| Two Years                             | -1%                               | 7%                                   |
| Three Years                           | 37%                               | 51%                                  |
| Four Years                            | 127%                              | 62%                                  |
| Five Years                            | <b>90%</b>                        | 79%                                  |
| Six Years                             | <b>316%</b>                       | 104%                                 |
| Seven Years                           | <b>946%</b>                       | 141%                                 |
| Eight Years                           | <b>900%</b>                       | 151%                                 |
| Nine Years                            | <b>867%</b>                       | 169%                                 |
| Since Inception <sup>4</sup>          | <b>845%</b>                       | 290%                                 |
| <i>Annualized Since<br/>Inception</i> | <b>23%</b>                        | <i>13%</i>                           |

Partnership interests in Top Mark Capital Partners increased in price by 25% year-to-date and 24% over the last 12 months which compares favorably to the S&P 500 Total Return Index (17% year-to-date and 20% over the last 12 months). With that said, we encourage you to assess performance on a cumulative basis, rather than focusing on

---

<sup>1</sup>Please review the Important Disclaimers on Page 28

<sup>2</sup> Performance is net of all fees (i.e., management fees, costs, *and* performance fees) and rounded to the nearest whole percentage.

<sup>3</sup> S&P 500 total return, i.e. dividends reinvested.

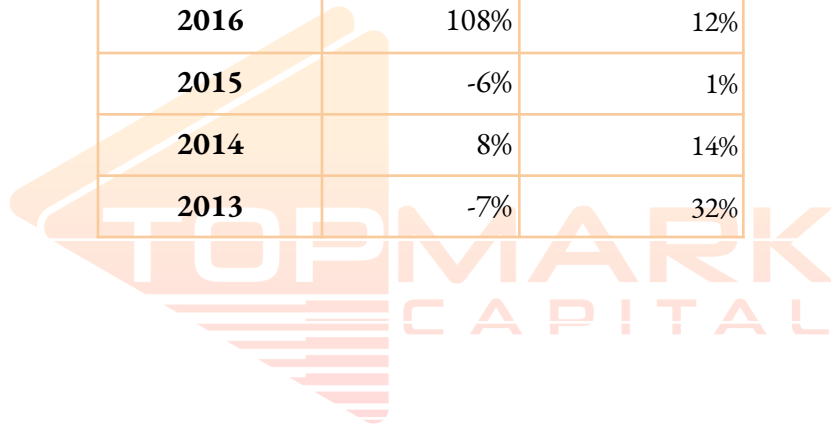
<sup>4</sup> September 1, 2012

discrete, short-term, increments. This approach offers a more accurate reflection of the investment's long-term value creation. The table above shows cumulative returns from various start dates, reinforcing the fact that results are best compared over long time horizons. You will notice that our returns are presented net of all fees and expenses.

The S&P 500 Total Return index is included, not as an explicit benchmark, but as a simple gauge against another investment you could have made, with returns presented the right way (cumulatively). We expect that our process will yield returns over the long-run that surpass those of a broad based index like the S&P 500. We know you have many options when it comes to placing long-term investments and we take seriously the responsibility that comes with stewarding your capital. As a reminder, your managers collective holdings represent the majority of our liquid net worths. We are truly in the same boat, for the long-haul.

For those who insist on examining annual returns data, we have included a table on the following page for your review. This information serves to provide additional context for our performance, but we reiterate the importance of evaluating returns cumulatively for a more accurate understanding. While evaluating returns in discrete periods, such as monthly, quarterly, or yearly, may seem like a natural and convenient way to measure performance, it can also lead to misleading and inaccurate conclusions. We wrote about evaluating returns in our 2022 Annual Letter (if you'd like a copy you can download it [here](#) or request it from us directly). By emphasizing the cumulative nature of returns, we aim to cultivate a more holistic and accurate understanding of our partnership's performance over time.

| <b>Year</b>     | <b>TMCP % Net Return<sup>5</sup></b> | <b>S&amp;P 500 Total Return<sup>6</sup></b> |
|-----------------|--------------------------------------|---|
| <b>2023 YTD</b> | 25%                                  | 17%   |
| <b>2022</b>     | -31%                                 | -18%  |
| <b>2021</b>     | 34%                                  | 29%   |
| <b>2020</b>     | 53%                                  | 18%   |
| <b>2019</b>     | 64%                                  | 31%   |
| <b>2018</b>     | 14%                                  | -4%   |
| <b>2017</b>     | 48%                                  | 22%   |
| <b>2016</b>     | 108%                                 | 12%   |
| <b>2015</b>     | -6%                                  | 1%  |
| <b>2014</b>     | 8%                                   | 14%   |
| <b>2013</b>     | -7%                                  | 32%   |




---

<sup>5</sup> Performance is net of all fees (i.e., management fees, costs, *and* performance fees) and rounded to the nearest whole percentage.

<sup>6</sup> S&P 500 total return, i.e. dividends reinvested.

## PORTFOLIO UPDATE & COMMENTARY

*Most new technologies foster improved product performance. I call these sustaining technologies...*

*... What all sustaining technologies have in common is that they improve the performance of established products, along the dimensions of performance that mainstream customers in major markets have historically valued. Most technological advances in a given industry are sustaining in character...*

*Disruptive technologies bring to a market a very different value proposition than had been available previously. Generally, disruptive technologies underperform established products in mainstream markets. But they have other features that a few fringe (and generally new) customers value. Products based on disruptive technologies are typically cheaper, simpler, smaller, and frequently, more convenient to use.*

*- Clayton Christensen  
The Innovator's Dilemma*

Your managers were busy turning over rocks in Q2, but we did not make any new investments, nor did we sell any of our existing positions during the quarter. Our goal is to identify one good investment per quarter. Given the fact that markets moved directionally up and the fact that we made more than one investment in Q1, we are not surprised to have no activity to report this quarter. The table below provides an overview of our Partnership's allocation by theme (as of June 30, 2023). Brief descriptions of each of the themes are included below, with the exception of the "Software is Eating the World" theme for which considerable developments have

occurred and additional insights are warranted and provided below. It is important to note that these themes are not mutually exclusive (i.e. some names in “Software is Eating the World” could also be in “Semiconductors are the Teeth”), nor are they always homogeneous (“Open RAN, 5G, & the Future of Wireless Networks” includes both convertible bonds and stocks), so they should not be considered ‘asset classes’ in the traditional sense of the term. Nevertheless, given our investment process<sup>7</sup>, we believe the table below and the descriptions that follow provide a useful snapshot of our investment activities.

| <i><u>Percent</u></i> | <i><u>Theme</u></i>                             | <i><u>Dividend /<br/>YTM</u></i> |
|-----------------------|---|----------------------------------|
| 28%                   | Software is Eating the World                    | 0.3%                             |
| 23%                   | TIAA  | 9.3%                             |
| 18%                   | Health  | 0.0%                             |
| 8%                    | Semiconductors are the Teeth                    | 1.1%                             |
| 5%                    | Metals Powered Economy                          | 1.0%                             |
| 4%                    | Open RAN, 5G, & the Future of Wireless Networks | 22.0%                            |
| 86%                   | Stocks, Convertible Bonds, & Private Credit     | 3.9%                             |
| 14%                   | Cash  |                                  |
| 100%                  | Total Partnership                               |                                  |

### Software is Eating the World (28%)

The phrase “software is eating the world” was first coined by Marc Andreessen, a prominent venture capitalist, in a 2011 Wall Street Journal article<sup>8</sup>. He argued that software was transforming industries and becoming an integral part of our daily lives, with companies that embraced this shift positioned to thrive while others faced obsolescence. This theme captures the idea that software-driven innovation is

---

<sup>7</sup> If you are unfamiliar with our process and philosophy, please request from us our [2022 Annual letter](#), or reach out to us directly to request a copy

<sup>8</sup> Andreessen, “Why Software Is Eating The World.”

permeating every aspect of the global economy, revolutionizing industries, and creating new opportunities for growth and value creation.

The most distinctive thing about software, from a competitive strategy perspective, is that it has zero marginal cost and zero distribution cost. This has particularly interesting implications that both investors and regulators struggle to wrap their heads around. Today, software includes ‘Artificial Intelligence’, because, at the end of the day, **when AI actually works, we just call it software.**

In the context of the excerpts from Clayton Christensen’s Innovator’s Dilemma provided at the beginning of this section, an important question for investors to ponder is whether or not AI is a disruptive or a sustaining innovation. Ben Thompson, a writer and analyst that has expanded on Christensen’s corpus over the years points out that it is much easier in hindsight to determine if an innovation was sustaining or disruptive. He also draws the conclusion that AI is likely to be sustaining<sup>9</sup>. We generally agree with his sentiment but reserve the fact that this applies to large technology companies as it relates to recommender systems and large language models (LLMs). We are also cognizant of the fact that these AI technologies are likely to be disruptive to some industries, particularly services industries, not generally considered competitive to the large technology companies.

The launch of ChatGPT, a LLM, woke the general public up to the fact that artificial intelligence is not just a fuzzy sci-fi theory, but rather here and now (though a bit wonky at times<sup>10</sup>). The general availability of these models and the increasing ease of incorporating them into products will trickle down to improvements for users, and for some businesses, an increase in cash flows. A few weeks ago, we did a short hand

---

<sup>9</sup> Thompson, “Google I/O and the Coming AI Battles.”

<sup>10</sup> “Hallucinations Could Blunt ChatGPT’s Success - IEEE Spectrum.”

calculation of the impact to cash flow of Microsoft's strategic investment in OpenAI and subsequent announcement of the AI infused Microsoft 365 Copilot (at \$30 per user per month<sup>11</sup>) on the [Telldata Podcast](#). Moving forward, expect to see more user facing 'AI' built into B2B software (and other areas), but don't expect all AI implementations to generate \$30B+ in annual free cash flow<sup>12</sup>.

As you likely know, we've been beating the drum on AI, and specifically Nvidia, for a while now (dating back to 2016). Nvidia is the most important 'pick-and-shovel' in the AI industry. Counterintuitively, however, its competitive advantage has been made possible by a relentless pursuit of differentiation by software. For those reasons we classify it under this theme rather than the "Semiconductors are the Teeth" theme. For a thorough analysis of the Nvidia moat, and a rebuttal to recent articles that have contested the company's moat, please see the article, "Nvidia's Cornerstone for Unmatched Dominance in AI (CUDA)" (found in a later section of this quarterly letter).

Many a value investor is seething over Nvidia's sky-high PE multiple, but our analysis suggests that the multiple is likely warranted given the unprecedented demand for the company's products. We cannot say the same for Nvidia's supposed competitors, AMD and Intel, however.

### TIAA (23%)

"TIAA" is an abbreviation for "There Is An Alternative". The name is both a counterbalance and an extension of the now defunct "TINA" ideology ("There Is No Alternative"), which gained traction in the wake of the Global Financial Crisis, suggesting that low bond yields left investors with no other choice but to hold equities.

---

<sup>11</sup> "Microsoft to Charge Premium for Generative AI Features | Financial Ti..."

<sup>12</sup> Mike and Jason, "2023.29 Breaking Down AI Innovations, Energy Trends, and Healthcare Insights."

Amidst the circumstances of 2020 and 2021, when equities were broadly overvalued and bond yields were barely above zero, our persistent challenge was to uncover value. Some of the investments we have made in this theme include ‘unloved equity’, credit, and workouts.

#### Health (18%)

Our health theme contains select positions that otherwise are in our newly launched partnership: Top Mark Health Partners (TMHP). Capital deployed in this ‘theme’ to date was mainly a result of the fact that our research turned up exceptional opportunities prior to the launch date of TMHP.

#### Semiconductors are the Teeth (8%)

The "Semiconductors are the Teeth" theme highlights the crucial role that semiconductor technology plays in enabling the "Software is Eating the World" phenomenon. As software continues to drive innovation and transform industries, semiconductors act as the "teeth" that facilitate this process.

#### Metals Powered Economy (5%)

The "Metals Powered Economy" theme centers around the transition to renewable energy solutions and the growing need for innovative energy storage and transportation methods. One of the key technologies emerging in this space is battery technology, and electric vehicles (EVs) are a prime, early application. Our [2022 Annual Letter](#)<sup>13</sup> included more details on this theme.

#### Open RAN, 5G, & the Future of Wireless Networks (4%)

Our “5G & the Future of Wireless Networks” theme aims to capitalize on the transformative potential of 5G technology and the rapidly evolving wireless network

---

<sup>13</sup> The 2022 Annual Letter is available [here](#), or reach out directly to request a copy

landscape. For a comprehensive analysis of this theme, please refer to our [2022 Q2 Letter](#)<sup>14</sup>.

# TOP MARK HEALTH PARTNERS LP

Partnership Letter

For the Quarter ended June 30, 2023



---

<sup>14</sup> The 2022 Q2 Letter is available [here](#), or reach out directly to request a copy

## PARTNERSHIP LETTER

It is an honor to be providing you the first update to Top Mark Health Partners (TMHP), our new healthcare focused partnership. We look forward to this introduction being the first of many such updates.

Total healthcare expenditures as a percentage of GDP have grown at an alarming rate in the developed world, especially in the United States, where the figure is nearly 20%. Meanwhile life expectancy in the United States is lower than in other developed countries who spend a far lower percentage of GDP on healthcare. The reasons for the disparity are varied and complex and highlight a growing opportunity for disruption in both the United States and other countries. We expect that technology will have an increasingly outsized effect on the healthcare industry in the coming decades. In other words, **healthcare is likely to change at an increasing rate**. This change creates an opportunity set that we believe will include the types of asymmetric risk/reward profiles that are well suited to the type of research and analysis we specialize in at Top Mark Capital. TMHP will be similar to our flagship fund (Top Mark Capital Partners), but will focus specifically on health and healthcare related investments.

One unique feature of this partnership is our relationship with Cold Spring Harbor Lab (CSHL). CSHL, home to eight Nobel Prize winners, is a private, not-for-profit laboratory that employs more than 1,000 people including 600 scientists, students and technicians. Our management company, the General Partner, has committed to a multi-year relationship with the lab where  $\frac{1}{4}$  of the performance fees generated from managing Top Mark Health will be donated to the lab. As a result, CSHL and the important research they foster will benefit from the growth of our partnership.

## PORTFOLIO UPDATE, PERFORMANCE & COMMENTARY<sup>15</sup>

The great Charlie Munger recently said “If you’re trying to do better than average, you’re lucky if you have four things to buy. And to ask for 20 is really asking for egg in your beer.” Mike and I are firm believers in that math. The research process we’ve developed managing Top Mark Capital Partners typically generates one good idea per quarter. A good idea that we feel has the upside opportunity to evolve into a great idea. As a result, it will take us some amount of time to build a full portfolio of investments even while employing a concentrated strategy.

One criterion that Mike and I look for in a company is owning a durable competitive advantage. That is something that certain verticals within the healthcare industry uniquely supply. In most industries, being a successful first mover normally puts a target on your back. In pharmaceuticals, regulatory agencies offer a huge moat in a number of manners. First, there is patent protection. Secondly, a would-be competitor must prove they have a safe and effective treatment. Third, that treatment must statistically outperform the current standard of care. And fourth, the treatment must undergo a lengthy and costly clinical trial process that is prohibitively expensive to test any treatments that garner less than the highest confidence.

In its first quarter of operation, the partnership has made three investments, all of which we were exploring to some extent prior to the official launch date of Top Mark Health Partners. The fund has risen 4.1%, which compares to a gain of 3.4% by the S&P 500 Healthcare Sector index. Due to the brief period since inception, and lack of fully deployed capital, our results to date are rather meaningless in the long term.

---

<sup>15</sup>Please review the Important Disclaimers on Page 28

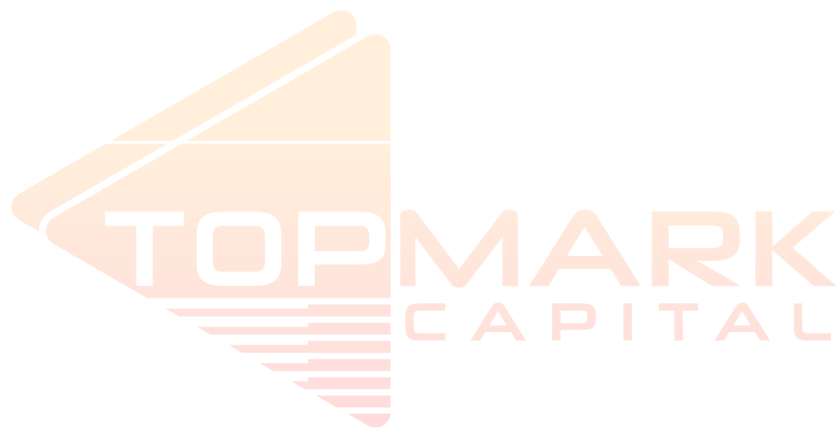
We will leave you with a thought on an advantage that TMHP holds which we plan to exploit: **patience**. We invest for the long term and even the most promising new treatments with regulatory fast-track status take years to bring to market. The average holding period for equity shares is a shockingly low 5.5 months<sup>16</sup>. This short-termism creates a persistent opportunity for investors with a long-term outlook to buy great companies where the value of their exceptional science has not yet been realized.



---

<sup>16</sup> “Buy, Sell, Repeat! No Room for ‘hold’ in Whipsawing Markets.”

SELECT ARTICLES, OBSERVATIONS &  
COMMENTARY



# NVIDIA's CORNERSTONE for UNMATCHED DOMINANCE in AI (CUDA)

*"We Have No Moat, And Neither Does OpenAI" proclaimed a leaked paper circulated internally at Google<sup>17</sup>. We believe that this is true for most businesses. However, one company does have a moat, a quite large moat, and that company is Nvidia.*

Nvidia has been the ultimate "pick and shovel" investment - playing a key role, dating back to 2012, in enabling the application of Artificial Intelligence (AI) technologies. This strategy has catapulted it from a \$30bn gaming hardware company to a big tech juggernaut now valued at more than \$1T. This article will not only show how Nvidia turned CUDA into their AI moat, but also show why PyTorch and Triton are not competitors. It will also detail CUDA's true competitor (AMD's ROCm) and show why it is vastly inferior.

## CUDA

The enabling force behind Nvidia's rise is a software library named Compute Unified Device Architecture, or CUDA. CUDA was originally released in 2007 as a software development kit allowing for general-purpose computing on Graphical Processing Units (GPUs), or a General-Purpose Graphical Processing Unit (GPGPU). Nvidia made this software free to use - enabling any engineer who wished to take advantage of the speed with which a GPU could perform parallelized floating point math operations to do so. Early uses of GPGPU processing were for cryptography, sorting algorithms, and building physics simulators similar to those used in virtual

---

<sup>17</sup> Patel, "Google 'We Have No Moat, And Neither Does OpenAI.'"

gaming worlds. Founder Jensen Huang and Nvidia hoped that by enabling the use of GPUs for accelerated computing tasks, they would sell more GPU cards. Today, we are witnessing the manifestation of this vision.

The catalyzing event that ushered in this new age of AI and cemented Nvidia as one of its dominant players occurred in 2012. In that year a computer vision model named AlexNet won the annual ImageNet competition. This competition consisted of using computers to identify objects in images from a set of 12 million pictures spanning 22,000 different objects.

At the time, the best algorithms programmed on CPUs could do no better than 74% accuracy, with most doing far worse. AlexNet, an algorithm trained with CUDA on GPUs, achieved 85% accuracy, trouncing the competition in 2012<sup>18</sup>. AlexNet used a Convolutional Neural Network (CNN) architecture, which at the time was thought to be impractical due to the amount of computation needed to train it. **AlexNet was Nvidia's coming out party.**

The following year every competitor was using a CNN. And every CNN was using CUDA on a Nvidia GPU<sup>19</sup>. The software development world took notice, building nearly all of the subsequent machine learning libraries upon CUDA.

### PyTorch & Triton: Challengers to the CUDA Moat?

Two machine learning libraries, PyTorch and Triton, that have achieved broad industry adoption, have been touted by popular media outlets as proof that Nvidia's

---

<sup>18</sup> "AlexNet and ImageNet."

<sup>19</sup> "ImageNet Large Scale Visual Recognition Competition 2013 (ILSVRC2013)."

CUDA moat is eroding<sup>20</sup>. However, due to the technical nature and some easily glossed over details, we believe these outlets are over emphasizing the threat to Nvidia.

PyTorch is a machine learning software framework originally created by Meta (Facebook) in 2016. Software engineers write code in either Python or C++ (traditional programming languages), which both use PyTorch to leverage the built-in machine learning tools. Meta released PyTorch as open-source, meaning it is freely available for anyone to use, contribute to, or modify for a specific use. However, as alluded to above, **the numeric operations that PyTorch performs are accelerated by CUDA, on Nvidia GPUs.**

PyTorch has since seen wide adoption by the industry - from small-scale hobbyist projects to large-scale projects like the Tesla Autopilot system. Members of the PyTorch community had the reasonable desire to remove vendor lock-in and provide the option to utilize PyTorch on AMD GPUs as well as CPUs. Adding this capability required a major revision to the framework, which resulted in PyTorch 2.0<sup>21</sup> (released March 15th, 2023) ultimately enabling AI applications to be compiled for use on either AMD or Nvidia hardware. However, it is not so simple as to 'switch' from Nvidia to AMD GPUs with pre-existing code. This is because CUDA and ROCm do not have the same functionality. Therefore a transition from Nvidia to AMD may require re-architecting the codebase and potentially eliminating functionality. The process by which PyTorch can execute code on an AMD GPU will be discussed in more detail below.

Similarly, Triton is a programming language developed by OpenAI that obfuscates CUDA so that an engineer does not have to program with CUDA directly. A common misconception is to call Triton a competitor to CUDA. This misunderstanding stems

---

<sup>20</sup> "Josh Wolfe on Where Investors Will Make Money in AI."

<sup>21</sup> "PyTorch 2.0."

from the conflation of the fact that just because a Triton programmer does not write CUDA code, it does not also mean that Triton does not use CUDA. In fact, Triton, in its current form, only executes on Nvidia GPUs (via CUDA). That said, the OpenAI team has left the door open to support other hardware configurations, not just AMD GPUs but also CPUs (listed as "under development")<sup>22</sup>.

*We're releasing Triton 1.0, an open-source Python-like programming language which enables researchers with no CUDA experience to write highly efficient GPU code—most of the time on par with what an expert would be able to produce.*

*- OpenAI Website*

In the two years subsequent to Triton's release, OpenAI has not implemented the support for AMD GPUs or CPUs. The general belief, potentially arising from the idea that Triton is a CUDA competitor, is that OpenAI's intention was to remove CUDA from the software stack. **We find this view naive, and lacking technical understanding.** A more likely explanation is that finding a workforce of engineers experienced programming to the CUDA interface was more difficult than developing a software abstraction, removing the need for engineers to directly use, or learn CUDA.

## AMD's ROCm

Meta and OpenAI are not natural competitors with Nvidia; rather, they are Nvidia's customers. As large customers, they reasonably don't like having a sole supplier of a critical component of their businesses. They would prefer to have AMD as well as Nvidia GPUs available for AI workloads. The Radeon Open Compute platform

---

<sup>22</sup> "Triton."

(ROCm), first released in 2016 (nearly a decade after CUDA), is AMD's attempt to make that possible and ultimately compete with Nvidia in this rapidly growing market.

ROCm is directly analogous to CUDA, it is a software stack for programming AMD GPGPU processors. As with CUDA, PyTorch 2.0 uses ROCm when an engineer wishes to develop their software for an AMD processor. While there is much to be said for learning from a prior example, ROCm has been playing catchup ever since it was launched. Nvidia recognizes that CUDA is their most important moat and they spend lavishly to maintain that advantage. This is supported by the fact that Nvidia now employs more software engineers than hardware engineers<sup>23</sup>. ROCm has always been a step or two behind as far as introducing new capabilities, supporting its entire line of GPUs<sup>24</sup>, or even running on Windows<sup>25</sup>. Additionally, the supporting infrastructure desired for a software library just isn't there for ROCm. For example, the software documentation, tutorials, and support is nowhere near what is offered by the CUDA community. Even some of the best engineers have had a difficult time finding a list of which AMD GPUs even properly run ROCm<sup>26</sup>.

### In Summary...

For the non-programmers reading this, the following analogy is more relatable (at least to the traditional finance crowd). **Just as you would build a financial model in Excel, you would build your new, clever AI chatbot using PyTorch (or Triton). PyTorch relies on either CUDA or ROCm and similarly, Excel relies on an**

---

<sup>23</sup> "NVIDIA Corporation Presents at Citi 2019 Global Technology Conference."

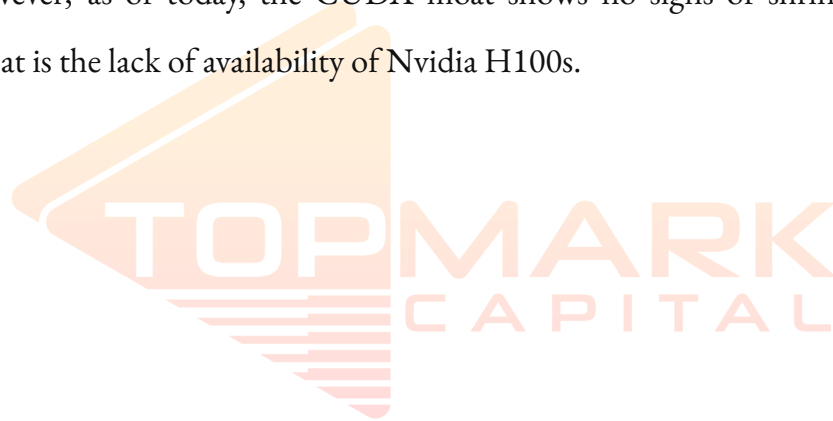
<sup>24</sup> Mujtaba, "AMD CEO Teases ROCm Support Coming To Radeon Consumer GPUs Soon."

<sup>25</sup> Ansari, "Why NVIDIA Keeps Winning.... And AMD Does Not."

<sup>26</sup> *This AMD Documentation Is Basically Unusable.*

**operating system, either Windows or MacOS. The catch, as there always is, is that PyTorch on ROCm is limited in just the same way that Excel on MacOS is annoyingly inferior to Excel on Windows.**

With how fast artificial intelligence applications are being developed, and certain aspects of it have been categorized as a "race", no one other than those running the largest stable AI workloads have felt enough pain to go through the effort of migrating away from an Nvidia GPU architecture. It's probable that one day, these two GPGPU software interfaces will reach parity. An investor in any of these companies will need to pay attention to this topic specifically, looking for signs that Nvidia's moat is starting to erode. However, as of today, the CUDA moat shows no signs of shrinking and its biggest threat is the lack of availability of Nvidia H100s.



## FORGET BLOCKCHAIN: GENERATIVE AI IS THE REAL Web3

Tech execs and other prominent industry leaders have all voiced opinions on generative AI being revolutionary, an iPhone moment, scary, and even “more capable than anything we’ve seen before”. Even the chairwoman of the FTC recently wrote that AI is bringing about the greatest social change that we’ve seen since the Web 2.0. The term ‘Web3’, referring to the third generation of the internet, is often misunderstood as a synonym for blockchain, the technology that powers cryptocurrencies like Bitcoin and Ethereum. However, blockchain is nothing more than a database technology that stores information. It does not represent a fundamental shift in how we interact with the internet. Rather, these generational shifts are defined by its contributors, and the third iteration is all about generative Artificial Intelligence.

Web 1.0 is thought of as the static web. This era was defined by static web pages produced by relatively few content creators. Think newspapers that have been put online. AOL and Yahoo provided you with stories to read, and if you were knowledgeable enough, you could create your own personal web page. This was a very passive web experience for most people.

Web2 ushered in the era where everyone became a participant on the internet. The term Web 2.0 was coined in 1999 when new ideas for incorporating information sharing directly into the internet were taking root. This ultimately led to the social platforms such as Facebook and Tik Tok that we are all familiar with, allowing every participant to be a direct contributor. Web2 is often called “the social web”. Today, anyone can participate in the public discussion, create YouTube videos or podcasts, or even become an influencer to promote products.

The term 'Web3' was co-opted by the crypto and blockchain community early on in an effort to legitimize the technology. You can't blame them, they were a solution looking for a problem and the hype allowed them to raise a tremendous amount of money to continue their development. However, nine years after that moment, we're still waiting for that one blockchain product that changes our daily lives. I surmise we will forever be waiting.

Given the first two generations of the internet were defined not by how computing got done, but rather by who contributed to the content, it simply does not make sense to define Web3 as one of blockchain technology.

There have been dramatic shifts in computing along every step of our online experience, from mainframes, to PCs, the cloud, and mobile, and blockchain more closely resembles a database innovation along this journey. But as users, we define the internet based on how we consume and interact with content online. As such, we do not view the very nature of the internet as changing when the method of computation is swapped.

The real shift to Web3 comes with the introduction of generative AI as a direct contributor. Bill Gates recently stated “I knew I had just seen the most important advance in technology since the graphical user interface” while referring to generative AI.

AI has already been playing a significant role in determining what advertising is displayed to whom. For example, we encounter this daily when Meta uses AI algorithms to determine the best targeted ad placement. However, ad placement was just the warm-up. The announcements of ChatGPT, Google Bard, and Microsoft Copilot have only thrust AI into the spotlight and made clear that this next generation of the

internet is upon us today. With the rapid rise of ChatGPT, we are no doubt already seeing tangible contributions by AI language models whether we notice it or not.

In contrast to the empty promises of the blockchain, what you will be doing in the near future is watching movies where AI was a contributor to the casting, editing, post-production, or even screenwriting. You will view AI curated sports highlights. You will be reading AI generated summaries of Zoom meetings, State of the Union Addresses, or that conference you couldn't attend. You will start following AI personalities on social media to both entertain and inform you. You are going to subscribe to podcasts entirely created and hosted by AI. Everyone will have a personal editor and assistant. And we will all likely have the pleasure of viewing advertisements crafted individually for us. History will show this as the transition to web3, marked by the rise of generative AI.

Did we really think the next evolution of a technology as transformative as the internet was monkey pictures and Shiba Inu coins? Absolutely not. Web3 is the introduction of artificial intelligence as a contributor.

## IMPORTANT DISCLAIMER

**Past performance is no guarantee of future results. Investing in equities and fixed income involves risk, including the possible loss of principal.** The investment performance presentation contains historical data and information relating to the performance of certain investments. These figures should not be considered as a guarantee or a reliable indicator of future performance. Investment returns and the value of an investment can go up or down, and there is no assurance that any investment strategy will achieve its objectives, generate profits, or avoid losses. Investing in financial markets inherently involves a certain degree of risk and speculation. The value of investments, and the income generated from them, can fluctuate due to various factors, including but not limited to market conditions, economic changes, interest rates, and political events. As such, there is always the potential for loss, and you should only invest funds that you can afford to lose. The S&P 500 Index is included to allow you to compare your returns against an unmanaged capitalization-weighted index of 500 stocks designed to measure performance of the broad domestic economy through changes in the aggregate market value of the 500 stocks representing all major industries. The NASDAQ 100 Index measures the performance of the largest 100 stocks on the NASDAQ.

**IMPORTANT DISCLOSURES.** The information herein is provided by Top Mark Capital Management LLC ("Top Mark Capital") and: (a) is for general, informational purposes only; (b) is not tailored to the specific investment needs of any specific person or entity; and (c) should not be construed as investment advice. Top Mark Capital makes no representation with respect to the accuracy, completeness or timeliness of the information herein. Top Mark Capital assumes no obligation to update or revise such information. In addition, certain information herein has been provided by and/or is based on third party sources, and, although Top Mark Capital believes this information to be reliable, Top Mark Capital has not independently verified such information and is not responsible for third-party errors. You should not assume that any investment discussed herein will be profitable or that any investment decisions in the future will be profitable. Investing in securities involves risk, including the possible loss of principal.

## REFERENCES

- “AlexNet and ImageNet: The Birth of Deep Learning | Pinecone.” Accessed July 31, 2023.  
[https://www.pinecone.io/learn/series/\[object%20Object\]/imagenet/](https://www.pinecone.io/learn/series/[object%20Object]/imagenet/).
- Andreessen, Marc. “Why Software Is Eating The World.” *Wall Street Journal*, August 20, 2011, sec. Life and Style.  
<https://online.wsj.com/article/SB10001424053111903480904576512250915629460.html>.
- Ansari, Tasmia. “Why NVIDIA Keeps Winning.... And AMD Does Not.” *Analytics India Magazine*, October 12, 2022.  
<https://analyticsindiamag.com/why-nvidia-keeps-winning-and-amd-does-not/>.
- Bloomberg.com*. “Josh Wolfe on Where Investors Will Make Money in AI.” July 17, 2023.  
<https://www.bloomberg.com/news/articles/2023-07-17/josh-wolfe-of-lux-capital-on-investing-on-ai-and-computing>.
- Financial Times*. “Microsoft to Charge Premium for Generative AI Features | Financial Ti....” July 18, 2023. <https://archive.li/OU05i>.
- “Hallucinations Could Blunt ChatGPT’s Success - IEEE Spectrum.” Accessed August 1, 2023.  
<https://spectrum.ieee.org/ai-hallucination>.
- “ImageNet Large Scale Visual Recognition Competition 2013 (ILSVRC2013).” Accessed July 31, 2023. <https://image-net.org/challenges/LSVRC/2013/results.php>.
- Mike, and Jason. “2023.29 Breaking Down AI Innovations, Energy Trends, and Healthcare Insights.” Substack newsletter. *Telltales* (blog), July 19, 2023.  
<https://telldales.substack.com/p/202329-cashflow-impact-of-microsofts>.
- Mujtaba, Hassan. “AMD CEO Teases ROCm Support Coming To Radeon Consumer GPUs Soon.” *Wccftech*, June 19, 2023.  
<https://wccftech.com/amd-ceo-teases-rocm-support-coming-to-radeon-consumer-gpus-soon/>.

“NVIDIA Corporation Presents at Citi 2019 Global Technology Conference,” September 5, 2019.

Patel, Dylan. “Google ‘We Have No Moat, And Neither Does OpenAI,’” January 16, 2023. <https://www.semianalysis.com/p/google-we-have-no-moat-and-neither>.

“PyTorch 2.0.” Accessed July 31, 2023. <https://pytorch.org/get-started/pytorch-2.0/>.  
*Reuters*. “Buy, Sell, Repeat! No Room for ‘hold’ in Whipsawing Markets.” August 3, 2020, sec. Business News.  
<https://www.reuters.com/article/us-health-coronavirus-short-termism-anal-idUSKBN24Z0XZ>.

*This AMD Documentation Is Basically Unusable*, 2023.  
<https://www.youtube.com/watch?v=Zsh6lPqvAcw>.

Thompson, Ben. “Google I/O and the Coming AI Battles.” *Stratechery* by Ben Thompson, May 15, 2023. <https://stratechery.com/2023/google-i-o-and-the-coming-ai-battles/>.

“Triton.” C++. 2014. Reprint, OpenAI, July 31, 2023. <https://github.com/openai/triton>.

