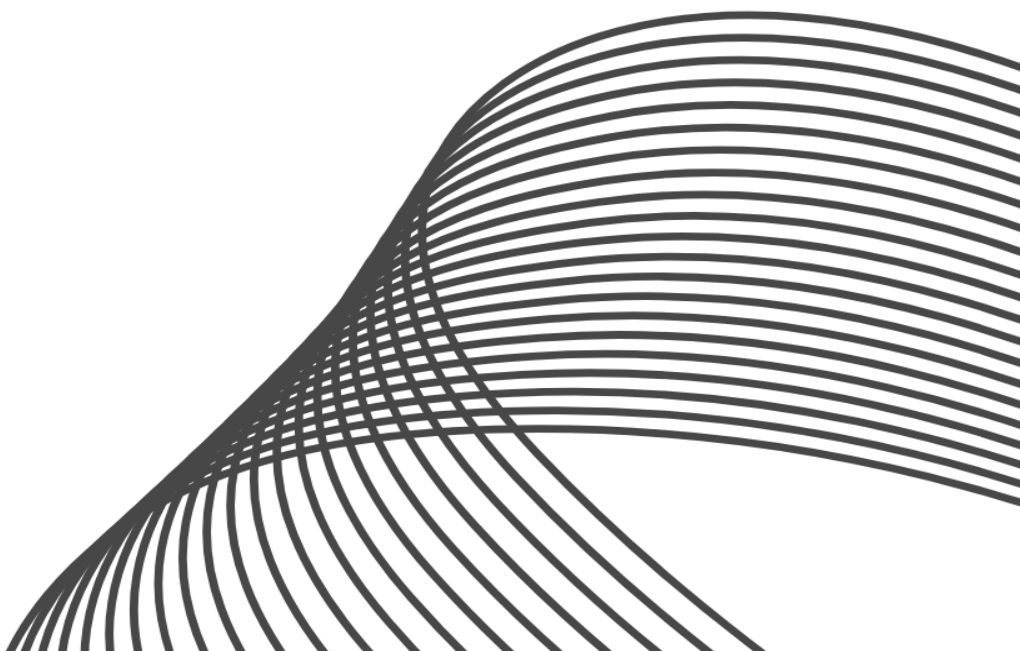




# Groq

## Company Primer



Contents

Quick snapshot ..... 2

Executive summary ..... 3

Funding overview ..... 4

Recent strategic moves ..... 5

Insights on company strategy ..... 9

Management profiles <sup>Beta</sup> ..... 11

Competitive landscape ..... 13

Market insights ..... 24

Annexure ..... 30

References ..... 32



## Quick snapshot

- **Company firmographics** (Source: Crunchbase)
  - Company Type: Startup; Operating Status: Active
  - Year founded: 2016; HQ: Mountain View, United States
  - No. of employees: 101-250
  - Revenue range: \$10M to \$50M
  - Website: <http://groq.com/>
- **Funding** (Source: Crunchbase)
  - Total Funding: USD 362.6M
  - Current Stage: Series C
  - Key investors: Tiger Global Management, D1 Capital Partners, Social Capital, Infinitum Partners
- **Product details** (Source: Wokelo Synthesis)
  - Product Category: Artificial Intelligence Computing
  - Industry: Electronics, Machine Learning, Semiconductor
  - Summary: Groq is a company that significantly simplifies computing processes to boost the performance of workloads in artificial intelligence (AI), machine learning (ML), and high-performance computing (HPC). Their core offering is centered around enhancing the speed and efficiency of these advanced computational tasks, making them more accessible and effective.

## Executive summary

- Groq radically simplifies compute to accelerate workloads in artificial intelligence, machine learning, and high-performance computing. Groq's revolutionary approach to simplifying complex computations is poised to advance the AI and HPC arenas through their innovative solutions.
- Groq's primary product offering includes its tensor streaming processor (TSP), which specializes in real-time AI inference and high-performance computing. It boasts 1 PetaOp performance, and is designed for low latency and high throughput. The TSP stands out for its predictable performance and ease of integration with machine learning models. Groq demonstrates technological capabilities with its Language Processing Unit (LPU), advancing natural language processing applications by enabling natural conversations with AI bots at speeds surpassing GPU solutions. Strategic partnerships, like with Argonne National Laboratory, and independent benchmarking highlight Groq's contribution to the industry.
- Groq's strategic advancements include forging partnerships with cloud service providers like Nimble, technology leaders like Samsung for next-generation AI chips production, and acquiring Maxeler Technologies for HPC and ML. These moves, underscored by a significant Series C fundraising of \$300 million co-led by Tiger Global Management and D1 Capital, demonstrate Groq's commitment to innovation, scalability, and entering new markets while enhancing financial health and positioning for market leadership.
- Groq aims to tailor its offerings to a segment of customers consisting of consumer electronics manufacturers, financial technology companies, high-performance computing centers, cybersecurity solutions providers, and AI and ML platforms. The strategy to target large clients for substantial revenue demonstrates a focused approach towards deep integration and capitalizing on the significant needs of companies in domains requiring real-time computing, like transportation and security.
- Groq positions itself for future growth by targeting large-scale data centers and niche markets with its superior processing capabilities to capitalize on the booming AI chip industry. The company's roadmap involves attracting performance-constrained customers in the AI hardware market while navigating competitive challenges from industry leaders like Nvidia and startups alike. Groq looks to differentiate with its product offerings and strategic customer focus, alongside industry insider appointments to the board to guide its cloud-focused evolution.

## Funding overview

### Funding Details

#	Date	Round	Amount Raised	Select Investors
1	2021-04-14	Series C - Groq	\$ 300.0M	Tiger Global Management, D1 Capital Partners
2	2020-08-12	Venture Round - Groq	-	D1 Capital Partners, Infinitum Partners
3	2018-09-05	Venture Round - Groq	\$ 52.3M	Social Capital
4	2018-07-10	Series B - Groq	-	-
5	2017-04-21	Venture Round - Groq	\$ 10.3M	Social Capital

Source: Crunchbase

## Recent strategic moves

### Insights from recent news

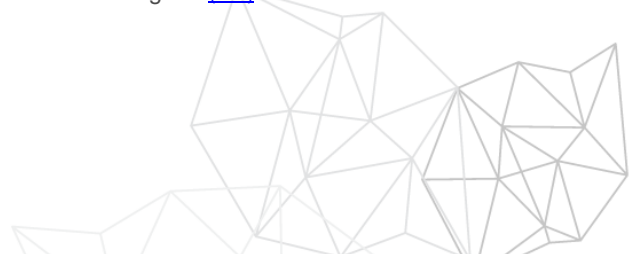
- **Accelerated AI Inference** <sup>[1-25]</sup>: Groq has cemented its position as an innovator in AI inference with its Language Processing Unit (LPU) designed for running Large Language Models (LLMs). The LPU provides faster responses and higher throughput for AI applications, particularly in natural language processing. Groq's technology facilitates natural conversations with AI bots at speeds significantly surpassing existing GPU alternatives, with claims of up to 75 times faster than human typing capabilities. Its custom chips, known as LPUs, have been independently benchmarked, with Groq's offering outperforming competitors in latency, throughput, and overall performance. Partnering with entities like Argonne Leadership Computing Facility (ALCF) and aiXplain, Groq demonstrates its hardware's capacity to accelerate complex scientific and enterprise applications. The company also achieves milestones by running foundational LLMs like Meta AI's Llama-2 70B with record-breaking token speeds. Groq's approach prioritizes speed, efficiency, and scalability, aiming to democratize AI advancement and application across various industries.
- **Cloud and HPC Expansion** <sup>[9,14,16,23-24]</sup>: Groq's strategic focus on expanding into cloud services and high-performance computing (HPC) is exemplified by its partnerships with cloud providers like Nimble and its integration into prestigious HPC facilities like Argonne National Laboratory. GroqRack systems offer robust solutions that empower researchers and developers to tackle demanding AI, ML, and HPC tasks with unprecedented performance levels. With Groq's TSP architecture, customers gain access to accelerators that double the performance of GPU systems without batching dependency, making real-time inference more feasible. The inclusion of Groq hardware on the Nimble Cloud and AI Testbeds at ALCF significantly enhances computational capabilities, potentially revolutionizing research fields such as material science, drug discovery, and other data-intensive domains.
- **Foundry Partnerships** <sup>[26-27]</sup>: Groq has partnered with technology giant Samsung to produce its next-generation AI chips. This strategic move leverages Samsung's newly invested \$17 billion Texas plant, tapping into advanced semiconductor manufacturing processes. This collaboration signifies Groq's commitment to fostering innovation within the AI acceleration market and strengthens its product roadmap with a US-based silicon partner. Groq's reliance on Samsung's Foundry Design Service team and manufacturing on the 4nm SF4X process ensures access to cutting-edge technological advancements and reinforces its North American-based engineering and manufacturing operations.
- **Financial Health and Investment** <sup>[28-30]</sup>: Groq's financial trajectory has been positively influenced by substantial funding rounds, specifically the Series C fundraising round that secured \$300 million. Co-led by prominent investment firms like Tiger Global Management and D1 Capital, Groq's total funding reaches \$367 million. This financial injection supports the company's ambitious growth plans across diverse industries, fosters talent acquisition, and propels the development of innovative products. With strong investor confidence, Groq

aims to reduce the cost of compute to zero, unlocking the immense potential of AI. The semiconductor industry's fertile landscape positions Groq to capture a significant share of the market, which is projected to range from \$65 billion to \$100 billion by 2025.

- **Strategic Acquisitions**<sup>[22]</sup>: Groq's acquisition of dataflow systems pioneer Maxeler Technologies aligns with its mission to deliver converged HPC and ML solutions. This move not only enhances Groq's architectural capabilities but also opens doors to the European market. By integrating Maxeler's dataflow computing proficiency into Groq's Tensor Streaming Processor architecture, the company aims to catalyze developer velocity, scalability, and cost efficiency in advanced computing applications. Leveraging the leadership and expertise from Maxeler promises streamlined innovation and positioning within the rapidly growing compute landscape.

Sources:

- [1] Tomsguide (22nd Feb 2024) - Groq lets you use multiple AI models quickly — here's how [\(link\)](#)
- [2] Gigazine (21st Feb 2024) - ``Groq'', which develops ``Language Processing Unit (LPU)'' that operates large-scale language models (LLM) at ... [\(link\)](#)
- [3] Tomsguide (20th Feb 2024) - Forget ChatGPT — Groq is the new AI platform to beat with blistering computation speed [\(link\)](#)
- [4] Cointelegraph (19th Feb 2024) - Groq AI chip system goes viral and rivals ChatGPT, challenges Elon Musk's Grok [\(link\)](#)
- [5] PR Newswire (13th Feb 2024) - Groq® LPU™ Inference Engine Leads in First Independent LLM Benchmark [\(link\)](#)
- [6] dzone.com (23rd Jan 2024) - Making AI Real: How Groq Simplifies Machine Learning for Developers [\(link\)](#)
- [7] PR Newswire (09th Jan 2024) - Groq® Opens API Access to Real-time Inference, the Magic Behind Instant Responses from Generative AI Products [\(link\)](#)
- [8] Hpcwire (09th Jan 2024) - Groq Opens API Access to Real-time Inference, Fueling Consumer AI Applications with Unprecedented Speed [\(link\)](#)
- [9] Insidehpc (21st Nov 2023) - ALCF: Groq AI Online Workshop, Dec. 6-7 - High-Performance Computing News Analysis [\(link\)](#)
- [10] HPCwire (19th Nov 2023) - Groq Achieves New Benchmark in Large Language Model ... [\(link\)](#)
- [11] PR Newswire (07th Nov 2023) - Groq Sets New Large Language Model Performance Record of 300 Tokens per Second per User on Meta AI Foundational LLM, Llama-2 70B [\(link\)](#)
- [12] Datanami (26th Oct 2023) - Groq Unveils New Low Latency LPU System at SC23, Paving the Way for Future AI Applications [\(link\)](#)
- [13] Pnewswire (26th Oct 2023) - Groq to Feature World's Fastest GenAI Inference Performance for Foundational LLMs at Supercomputing '23 on Its LPU ... [\(link\)](#)
- [14] Datacenterdynamics (19th Oct 2023) - Argonne Lab deploys Groq AI hardware - DCD [\(link\)](#)
- [15] Pnewswire (20th Sep 2023) - Groq to Showcase World's Fastest Large Language Model Performance, Powered by Its LPU™ System, at the Global ... [\(link\)](#)
- [16] EE Times (12th Sep 2023) - Groq Demonstrates Fast LLMs on 4-Year-Old Silicon [\(link\)](#)
- [17] Yahoo (31st Aug 2023) - Groq Smashes LLM Performance Record Again Using an LPU™ System With No Response From GPU Companies [\(link\)](#)
- [18] Pnewswire (08th Aug 2023) - Groq™ First to Achieve 100 Tokens Per Second Per User on Meta AI's Llama-2 70B, Leading AI Artificial Intelligence ... [\(link\)](#)
- [19] Pnewswire (15th Mar 2023) - AI Accelerator Groq™ Adapts and Runs LLaMA, the Meta™ Chatbot Model and Competitor to ChatGPT, for Its Systems [\(link\)](#)
- [20] PRNewswire Asia (16th Jun 2022) - Frost & Sullivan Recognizes Groq with the Technology Innovation Leadership Award for Its Sophisticated Processor Architecture Technology [\(link\)](#)
- [21] Allaboutcircuits (29th Mar 2022) - Ep. 42 | Groq CEO and Ex-Google Jonathan Ross on the Petaflop AI Chip and First Ever TPU - Engineering Podcast [\(link\)](#)
- [22] Hpcwire (01st Mar 2022) - Groq Acquires Dataflow Systems Firm Maxeler Technologies [\(link\)](#)
- [23] Eetasia - Groq Rolls out AI Chip on the Cloud [\(link\)](#)




- [24] Insidehpc - Groq Hardware Now Available on Nimbix Cloud for AI Workloads - High-Performance Computing News Analysis [\(link\)](#)
- [25] Fiercееlectronics - Groq TSP blows away single chip record at 1 PetaOp/s [\(link\)](#)
- [26] Ajudaily (16th Aug 2023) - Samsung to produce AI chips for California-based AI semiconductor startup Groq [\(link\)](#)
- [27] Pnewswire (15th Aug 2023) - Groq Selects Samsung Foundry to Bring Next-gen LPU™ to the AI Acceleration Market [\(link\)](#)
- [28] Unite - AI Chip Startup Groq Closes \$300 Million in Series C Fundraising [\(link\)](#)
- [29] Hpcwire - Groq Closes \$300 Million Fundraising Round [\(link\)](#)
- [30] Yahoo - Tiger Global, D1 Lead \$300 Million Round in AI Chip Startup Groq [\(link\)](#)

## M&A strategy

Source: Crunchbase, Wokelo synthesis, News articles (sourced from 8k+ publishers)

Groq, a generative artificial intelligence solutions company, has acquired Maxeler Technologies, a pioneer in dataflow computing. Maxeler specializes in converged high-performance computing solutions for various domains such as financial services, climate, brain, and quantum computer simulations. This acquisition will enhance Groq's product capabilities, expand their presence in Europe, and strengthen their position in the AI, machine learning, and high-performance computing markets.

**Table: M&A Timeline**

#	Year	Logo	Name	Product Category	HQ
1	2022		Maxeler Technologies	Accelerated High Performance Computing	Palo Alto, United States

## Partnerships & GTM

Source: Crunchbase, Wokelo synthesis, News articles (sourced from 8k+ publishers)

Groq has announced a partnership with OneNano, a next-generation cryptocurrency exchange platform, aiming to revolutionize cryptocurrency exchanges. OneNano plans to leverage Groq's ultra-low latency AI acceleration solutions to develop a secure and reliable platform with an 8,000x speedup compared to current exchanges. The platform will offer real-time tracking of positions and a stable exchange in all market conditions, with testing starting next year and an official launch planned for the second quarter of 2023.

**Table: Partnerships Timeline**

#	Year	Partnership
1	2022	Groq and OneNano Partner to Revolutionize Cryptocurrency Exchanges <a href="#">(link)</a>





## Appointments and other hires

- **Strategic Onboarding of Cloud Expertise to Groq's Board<sup>[1]</sup>:** Groq Inc. has enhanced its Board of Directors with the appointment of Raju Gulabani, a leader with a solid trajectory of scaling cloud businesses. His prior role as Vice President of AWS's Database and Analytics business, where he escalated the division from inception to market dominance, positions him as a strategic asset for guiding Groq through its next phase of cloud-focused product evolution.
- **Infusion of Semiconductor Insight and Strategy on Groq's Board<sup>[2]</sup>:** The appointment of Andy Rappaport to Groq's Board of Directors melds deep semiconductor industry understanding with Groq's innovative trajectory. His recognition of computational architecture shifts aligns with Groq's vision for AI and machine learning workloads, underlining his future-facing strategic counsel expected to fortify Groq's commercial trajectory.

Sources:

[1] Prnewswire (22nd Feb 2022) - Groq Adds Former Amazon Leader Raju Gulabani to Board of Directors [\(link\)](#)

[2] Prnewswire (25th Feb 2021) - Groq Appoints Andy Rappaport to Its Board of Directors [\(link\)](#)

## Industry recognition and certification

- **Technology Innovation Leadership Award<sup>[1]</sup>:** Groq has been recognized with the 2022 North America Technology Innovation Leadership Award by Frost & Sullivan for its advancements in AI processors for data centers. This accolade underscores the company's successful implementation of innovative features and functionalities in its product, leading to swift market adoption and enhanced customer value.

Sources:

[1] PRNewswire Asia (16th Jun 2022) - Frost & Sullivan Recognizes Groq with the Technology Innovation Leadership Award for Its Sophisticated Processor Architecture Technology [\(link\)](#)

## Insights on company strategy

### Takeaways from leadership interviews and case studies

- **Products<sup>[1-7]</sup>:** Groq's unique selling point lies in its tensor streaming processor (TSP) which specializes in inference processing for real-time AI and high-performance computing (HPC), claiming speeds of 1 PetaOp or one quadrillion operations per second. The TSP architecture is designed to provide predictable and repeatable performance with low latency, high throughput, and significantly reduced complexity compared to traditional CPU and GPU systems. By starting with software and creating a simple but fast chip, Groq ensures data flows in at the right time and place, resulting in immediate calculations without delay. This approach offers a compelling alternative to traditional multi-core chip designs, aiming for orders of magnitude performance improvements. Groq's software-defined hardware coupled with their advanced compiler technology simplifies deployment allowing easy integration with machine learning models without proprietary interfaces, which prevents hardware lock-in issues and creates potential for rapid, real-time processing for various applications including autonomous vehicles and data center operations.
- **Roadmap<sup>[6-7]</sup>:** Groq's future strategy involves targeting large data centers and niche markets such as autonomous vehicles, where reliability, predictability, and low-power consumption are crucial. Their strategic focus is on securing a small number of large-scale customers to widely deploy their chips, which would suffice for the company to prosper independently. Groq plans on using its superior processing capabilities to attract customers who currently face performance bottlenecks with existing AI hardware solutions. By positioning itself as an alternative to large industry players, Groq aims to gain a significant share in the rapidly growing AI chip market, which is especially important considering its decision against seeking acquisition and instead aiming for sustainable, standalone growth.
- **Customers<sup>[6]</sup>:** Groq is looking to work closely with a select set of customers, akin to "elephant hunting," where each sizable client provides substantial revenue, ensuring the company's sustainability. This approach indicates a focus on quality over quantity, emphasizing deep integration and high value for each client. Groq is currently sending out chip samples, signaling that the company is in the late stages of product testing and close to securing customer commitments. The target customer base includes operators of large data centers and industries that require instant processing capabilities such as transportation, security, and robotics.
- **Industry outlook<sup>[7]</sup>:** The industry outlook for AI chips is incredibly promising, with a projected market growth to nearly \$129 billion by 2025, a threefold increase from the 2018 market size of approximately \$43 billion. This booming market is driven by the need for specialized AI processors in an array of applications from data centers to edge devices. Companies like Groq are contributing to the emergence of new architectures that challenge traditional devices used in AI processing such as CPUs and GPUs. The trend towards integrating memory closer to computational cores and enabling processing parallelism is expected to permeate across various industries, fostering innovations in chip technology that accommodate the burgeoning demands of AI applications.

- **Competition and challenges<sup>[5,7]</sup>:** As a new entrant in the AI chip market, Groq faces the challenge of breaking through a field dominated by established giants like Intel, NVIDIA, and Qualcomm. It also contends with a multitude of other startups vying for market share in the growing AI processing segment. The challenges consist of proving the practical benefits of their hardware, managing production costs, and establishing a foothold in a market that traditionally has high barriers to entry. Competition comes from both large corporations developing their own silicon and numerous innovative startups providing unique approaches to AI chip design, thus creating a highly dynamic and competitive landscape. Groq's product differentiation and strategic customer focus are essential as it navigates the intricacies of manufacturing, selling, and scaling within this competitive market.

Sources:

- [1] Tomsguide (22nd Feb 2024) - Groq lets you use multiple AI models quickly — here's how [\(link\)](#)
- [2] dzone.com (23rd Jan 2024) - Making AI Real: How Groq Simplifies Machine Learning for Developers [\(link\)](#)
- [3] Forbes (10th Nov 2022) - Groq – Reimagining High Performance Computing [\(link\)](#)
- [4] Allaboutcircuits (29th Mar 2022) - Ep. 42 | Groq CEO and Ex-Googler Jonathan Ross on the Petaflop AI Chip and First Ever TPU - Engineering Podcast [\(link\)](#)
- [5] Cnbc - Ex-Googlers left secretive AI unit to form Groq with Palihapitiya [\(link\)](#)
- [6] Datacenterknowledge - Former Google Chip Guru Takes Novel Approach to AI at Groq [\(link\)](#)
- [7] Fiercееlectronics - How AI chips will explode 3x by 2025 with startups like Hailo, Syntiant and Groq [\(link\)](#)



## Management profiles <sup>Beta</sup>

Management	Background
<b>Michelle Donnelly</b> <sup>[1-2]</sup> Chief Revenue Officer <a href="#">LinkedIn Profile</a>	<ul style="list-style-type: none"> <li>Michelle Donnelly currently serves as the Chief Revenue Officer at Groq. In her role, she exhibits a customer-obsessed mindset and applies her experience as a growth hacker to lead the company's revenue generation strategies. Her approach emphasizes trust, customer success, and innovation, integrating these pillars to exceed revenue targets and drive performance gains. Michelle is committed to solving machine learning challenges with Groq's solutions, ensuring customer success through transparent and strategic partnerships.</li> <li><b>Previous Experience:</b> Sales and go-to-market leader at Salesforce.</li> </ul>
<b>Tobi Crabtree</b> <sup>[3]</sup> VP of People & Culture	<ul style="list-style-type: none"> <li>Tobi Crabtree is the current VP of People &amp; Culture at Groq. His role involves overseeing and nurturing the company's culture, empowering collaboration among dispersed teams, and focusing on talent development and team building. He puts emphasis on groqster camaraderie through intellectually and interpersonally satisfying engagement. Tobi also ensures open communication and employee inclusion in the company's growth trajectory.</li> <li><b>Previous Experience:</b> Formerly employed at Intel before joining Groq.</li> </ul>
<b>Jim Miller</b> <sup>[3-4]</sup> VP, Hardware Engineering <a href="#">LinkedIn Profile</a>	<ul style="list-style-type: none"> <li>Jim Miller serves as the VP of Hardware Engineering at Groq. In his role, he is dedicated to fostering innovation in machine learning (ML) through the GroqChip, based on the Tensor Streaming Processor architecture. His focus lies in ensuring simplicity, performance, predictability, and power efficiency to facilitate ML advancements. With a strong leadership background, he contributes to Groq's vision and culture of ownership among a talented and nimble team.</li> <li><b>Previous Experience:</b> Formerly at Amazon, Jim Miller was responsible for the hardware development, delivery, and support for AWS' compute platforms. His prominent role at Intel earlier in his career involved designing the i486 microprocessor.</li> </ul>
<b>Mark Heaps</b> <sup>[3,5]</sup> VP of Brand & Creative	<ul style="list-style-type: none"> <li>Mark Heaps serves as the VP of Brand &amp; Creative at Groq, where he leads efforts related to brand building and creative processes. He is recognized for fostering a high concentration of talented teams and having a significant impact on the company's creative direction. His views on the intersection of AI and human creativity, particularly regarding GenAI and Large Language Models (LLMs), highlight his expertise in integrating AI into creative workflows, enhancing productivity, and driving innovation.</li> <li><b>Previous Experience:</b> Prior to joining Groq, Mark Heaps has accumulated experience at Duarte, Google, and Apple, holding roles related to creative and brand development.</li> </ul>



**Jonathan Ross** <sup>[6-10]</sup>  
CEO & Founder  
[LinkedIn Profile](#)

- Jonathan Ross is the CEO & Founder of Groq, a company that specializes in making AI accessible and providing innovative compute technologies for AI, ML, and high performance computing solutions. He leads the company with a vision to advance the full promise of AI through technology like the GroqChip, offering unparalleled deterministic execution, low latency, and power efficiency. With a leadership style that emphasizes confidence, continuous learning, and empowering smarter individuals, Jonathan guides Groq in its mission while fostering talent density and a collaborative culture. He's involved in pushing the boundaries of generative AI and Large Language Model-powered solutions, driving the company's growth and technological advancements.
- **Previous Experience:** Prior to founding Groq, Jonathan Ross initiated Google's TPU effort as a 20% project and designed and implemented the core elements of the original chip. He also served on Google X's Rapid Eval Team, where he incubated new Units for Alphabet. Ross has had significant influence in the technology and leadership spheres, evidenced by his collaborative experiences with other industry leaders.
- **Education background:** Jonathan Ross studied mathematics and computer science at NYU's Courant Institute, where, during his second year, he was the first computer science undergraduate to complete courses typically reserved for Ph.D. students.

- [1] Groq (2022-06-10) - Why Groq - Michelle Donnelly - Groq [\(link\)](#)  
[2] Groq - Press Release - Groq [\(link\)](#)  
[3] Groq (2024-02-20) - About - Groq [\(link\)](#)  
[4] Groq (2021-10-28) - Come Join Our Elite Team of AI Innovators - Groq [\(link\)](#)  
[5] Groq (2023-05-18) - Groq Receives Comparably Award for Best Leadership Team Based ... [\(link\)](#)  
[6] Groq (2022-06-10) - Jonathan Ross: 'Every. Word. Matters.' - Groq [\(link\)](#)  
[7] Groq (2021-10-28) - Honoring Women of the Past, Hearing Women of the Present - Groq [\(link\)](#)  
[8] Groq (2024-02-20) - Groq® Welcomes Harvard Professor Youngme Moon to Board of ... [\(link\)](#)  
[9] Groq (2024-02-15) - ArtificialAnalysis.ai LLM Benchmark Doubles Axis To Fit New Groq ... [\(link\)](#)  
[10] Groq (2023-11-09) - HumanPlus: The Co-evolution of AI & Human Potential - Groq [\(link\)](#)












## Competitive landscape

### Overall categories of players

Source: Wokelo synthesis, Crunchbase firmographics

Note: Please refer this section for sample profiles adjacent to Groq

Categories	Example Players	
Market Leaders		<b>NVIDIA:</b> NVIDIA is a global leader in AI and computing, known for its GPUs that accelerate a wide range of AI, machine learning, and high-performance computing tasks.
Broader Incumbents		<b>Intel:</b> Intel offers a broad portfolio of computing products and technologies, including CPUs and accelerators for AI and high-performance computing applications.
		<b>AMD:</b> AMD produces CPUs and GPUs that support AI, machine learning, and high-performance computing, competing directly with NVIDIA and Intel.
		<b>IBM:</b> IBM offers advanced AI and machine learning solutions, including IBM Watson, as well as high-performance computing systems.
Adjacent Players		<b>Google Cloud:</b> Google Cloud provides powerful cloud computing services with specialized AI and machine learning capabilities through its machine learning engine and AI platform.
		<b>Amazon Web Services (AWS):</b> AWS offers extensive cloud computing capabilities with specific services and hardware tailored for AI and machine learning, including AWS SageMaker.
		<b>Microsoft Azure:</b> Microsoft Azure provides a comprehensive suite of cloud services, including AI and machine learning tools and platforms for developers and businesses.
Niche Players		<b>Arm Holdings:</b> Arm Holdings designs processors used in a wide range of devices; its technology is increasingly being used for AI and machine learning applications, particularly in mobile and IoT devices.
		<b>Qualcomm:</b> Qualcomm is known for its Snapdragon processors, which power smartphones and are increasingly being utilized for AI and machine learning applications within mobile and edge computing.



**Synopsys:** Synopsys offers electronic design automation software tools and IP, crucial for designing chips used in AI and machine learning applications.



**Numenta:** Numenta is focused on machine intelligence research and is developing unique computing frameworks inspired by the human brain to advance AI and machine learning.

## Emerging Startups



**Graphcore:** Graphcore is developing innovative IPU (Intelligence Processing Unit) hardware specifically designed for accelerating AI and machine learning computations.



**Cerebras Systems:** Cerebras Systems stands out for its development of the world's largest computer chip, designed to dramatically accelerate AI computations.



**Cambricon:** Cambricon is a Chinese startup focusing on creating AI chips for cloud computing and edge devices, contributing to the acceleration of AI workloads.






**SambaNova Systems:** SambaNova Systems is developing next-generation computing platforms to support AI applications, leveraging novel dataflow architecture for machine learning and data analytics.








## Select startups

Source: Wokelo proprietary algo for company prioritization, Crunchbase firmographics






Note: This list may not be exhaustive and includes a sample set of companies in this space.

	Company	Details	Funding	Product description
	<b>Cerebras Systems</b> AI Acceleration Chips	Founded: 2016 HQ: Sunnyvale, United States	Funding: USD 715.0M Last Round: Nov'21 Stage: Series F - Cerebras Systems	Cerebras Systems develops computing chips with the sole purpose of accelerating AI.
	<b>SmartMore Corporation Limited</b> Artificial Intelligence Solutions	Founded: 2019 HQ: Shenzhen, China	Funding: USD 300.0M Last Round: Jun'21 Stage: Series B - Smartmore Corporation Limited	SmartMore Technology provides artificial intelligence solutions.
	<b>MIT-IBM Watson AI Lab</b> Artificial	Founded: 2017 HQ: Cambridge, United States	Funding: USD 240.0M Last Round: Sep'17 Stage: Venture	MIT-IBM Watson AI Lab focuses on fundamental artificial intelligence research.



	Intelligence Research Facility		Round - Mit-Ibm Watson Ai Lab	
	<b>SiMa.ai</b> Machine Learning, Computer Vision	Founded: 2018 HQ: San Jose, United States	Funding: USD 190.0M Last Round: Jun'23 Stage: Venture Round - Sima.Ai	SiMa.ai is a machine learning startup that aims to deliver the industry's first software-centric platform.
	<b>Run:AI</b> AI Cluster Management Platform	Founded: 2018 HQ: Tel Aviv, Israel	Funding: USD 118.0M Last Round: Mar'22 Stage: Series C - Run:AI	Run:ai AI Cluster Management Platform helps organizations squeeze more from their GPUs and streamline AI/ML development
	<b>Defined.ai</b> AI Data Marketplace Services	Founded: 2015 HQ: Seattle, United States	Funding: USD 78.6M Last Round: Jan'22 Stage: Venture Round - Defined.Ai	Defined.ai (former DefinedCrowd) enabling AI creators of the future.
	<b>Recogni</b> Autonomous Driving Technology	Founded: 2017 HQ: San Jose, United States	Funding: USD 73.9M Last Round: Feb'21 Stage: Series B - Recogni	System solutions company that specializes in the design of high-performance, low-power inference processors for AI
	<b>Edgecortix</b> AI Semiconductor Design	Founded: 2019 HQ: Tokyo, Japan	Funding: USD 35.6M Last Round: Oct'23 Stage: Venture Round - Edgecortix	Semiconductor design company headquartered in Japan, that develops accelerators for AI and machine learning.
	<b>Algo</b> AI Supply Chain Planning	Founded: 2016 HQ: Troy, United States	Funding: USD 35.0M Last Round: Jul'21 Stage: Venture Round - Algo	Enterprise AI powered Supply Chain Planning platform delivered by a Virtual Business Analyst named Algo.
	<b>Intelligence Indeed</b> Artificial Intelligence Enterprise	Founded: 2018 HQ: Hangzhou, China	Funding: USD 31.1M Last Round: Dec'23 Stage: Series C - Intelligence Indeed	Intelligence Indeed is an artificial intelligence enterprise for decision-making field of large-scale and complex problems.
	<b>Preligens</b> Artificial Intelligence Geospatial Analysis	Founded: 2016 HQ: Paris, France	Funding: USD 28.3M Last Round: Nov'20 Stage: Series A - Preligens	We develop pioneering technology to empower analysts to take smart actions for a safer world
	<b>Evolv AI</b> Artificial Intelligence Platform	Founded: 2019 HQ: San Francisco, United States	Funding: USD 23.3M Last Round: Nov'23 Stage: Venture Round - Evolv Ai	Artificial intelligence platform to automatically discover, personalize, and serve better journeys by continuously adapting to live behavior
	<b>Ask-AI</b> AI Analytics	Founded: 2021 HQ: Tel Aviv, Israel	Funding: USD 20.0M Last Round: Jan'24 Stage: Series A - Ask-Ai	Ask-AI mines text-heavy corporate knowledge and customer communications for precise answers and actionable insights.

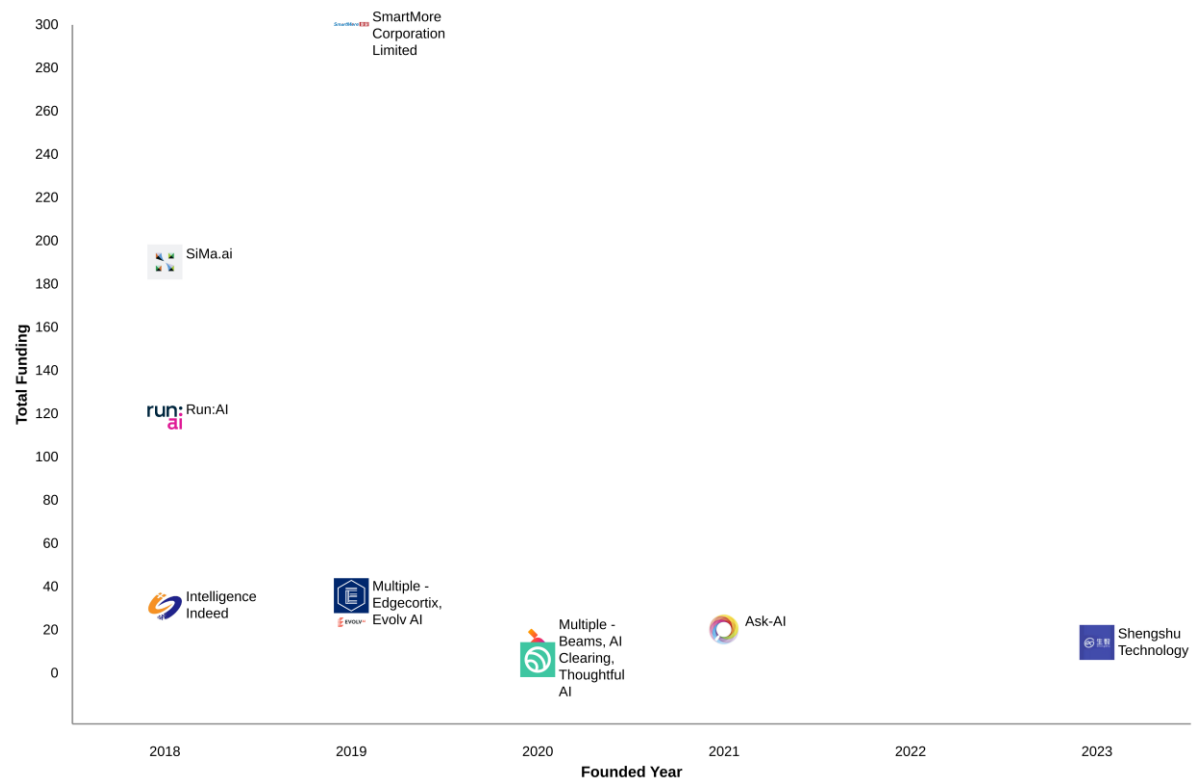


	<b>AI Clearing</b> Construction Tracking	Founded: 2020 HQ: Austin, United States	Funding: USD 17.5M Last Round: Oct'23 Stage: Series A - AI Clearing	Full digital field construction progress tracking. To decrease re-work cost. To mitigate litigation risk. 100% databased.
	<b>Thoughtful AI</b> Healthcare Automation Technology	Founded: 2020 HQ: Chicago, United States	Funding: USD 15.8M Last Round: Mar'22 Stage: Seed Round - Thoughtful Ai	AI-powered Healthcare Automation
	<b>Shengshu Technology</b> Artificial Intelligence Solutions	Founded: 2023 HQ: Haidian, China	Funding: USD 14.0M Last Round: Aug'23 Stage: Angel Round - Shengshu Technology	Shengshu Technology is an artificial intelligence solution provider that focuses on the creation of multi-modal application products.
	<b>Quartic.ai</b> Industrial AI/IIoT	Founded: 2017 HQ: San Jose, United States	Funding: USD 12.0M Last Round: May'22 Stage: Venture Round - Quartic.Ai	Quartic.ai is an industrial AI and IIoT software provider.
	<b>Beams</b> AI Insights Platform	Founded: 2020 HQ: Berlin, Germany	Funding: USD 6.0M Last Round: Nov'21 Stage: Seed Round - Beams	AI Insights Platform for Highly Regulated Industries.



## Chart: Select startups by total funding

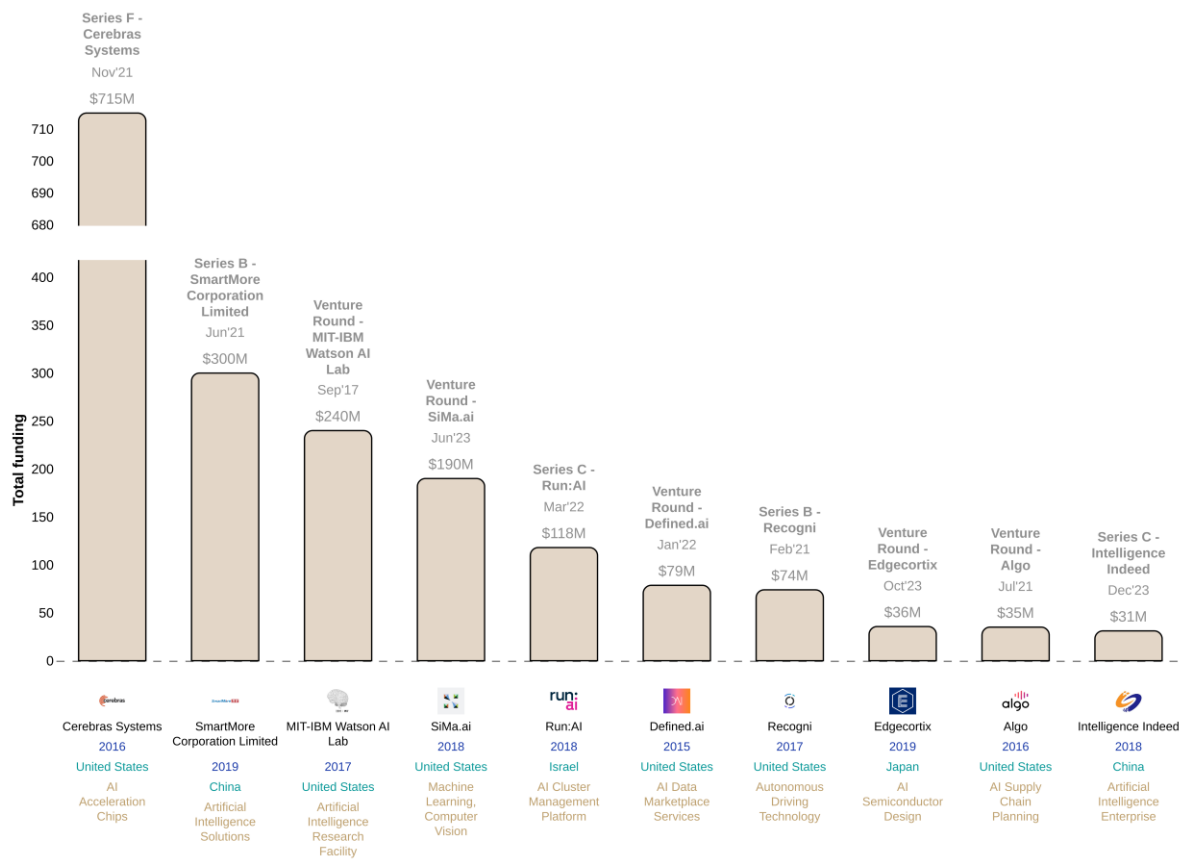
(All numbers in \$M)



Source: Wokelo generated, Crunchbase data

## Chart: Funding by startups

(All numbers in \$M)






Source: Wokelo generated, Crunchbase data

## Select public companies

Source: Wokelo proprietary algo for company prioritization, Crunchbase firmographics, Finnhub financial data

Note: This list may include companies in related or adjacent segments that may not necessarily align. These include closest matched public companies for benchmarking.

			
Company	Google	Intel	NVIDIA
Details	Founded: 1998 HQ: Mountain View, United States	Founded: 1968 HQ: Santa Clara, United States	Founded: 1993 HQ: Santa Clara, United States
Overview	Google is a multinational corporation that specializes	Intel designs, manufactures, and sells	NVIDIA is a computing platform company operating at the

	in Internet-related services and products.	integrated digital technology platforms worldwide.	intersection of graphics, HPC, and AI.
<b>Year</b>	2023	2023	2023
<b>Sales</b>	USD 307.4B	USD 54.2B	USD 27.0B
<b>EBITDA</b>	USD 100.2B	USD 9.6B	USD 7.8B
<b>EBIT</b>	USD 88.2B	USD 31.0M	USD 6.3B
<b>Market Cap</b>	USD 1.8T	USD 188.6B	USD 1.8T
<b>EV</b>	USD 1.7T	USD 254.0B	USD 478.9B
<b>EV/EBIT</b>	19.78	8.2K	76.34
<b>EV/EBITDA</b>	17.42	26.37	61.26
<b>Gross Margin (%)</b>	56.94	40.04	59.51
<b>EBIT Margin (%)</b>	28.7	0.06	23.26
<b>Net Margin (%)</b>	24.01	3.11	16.19

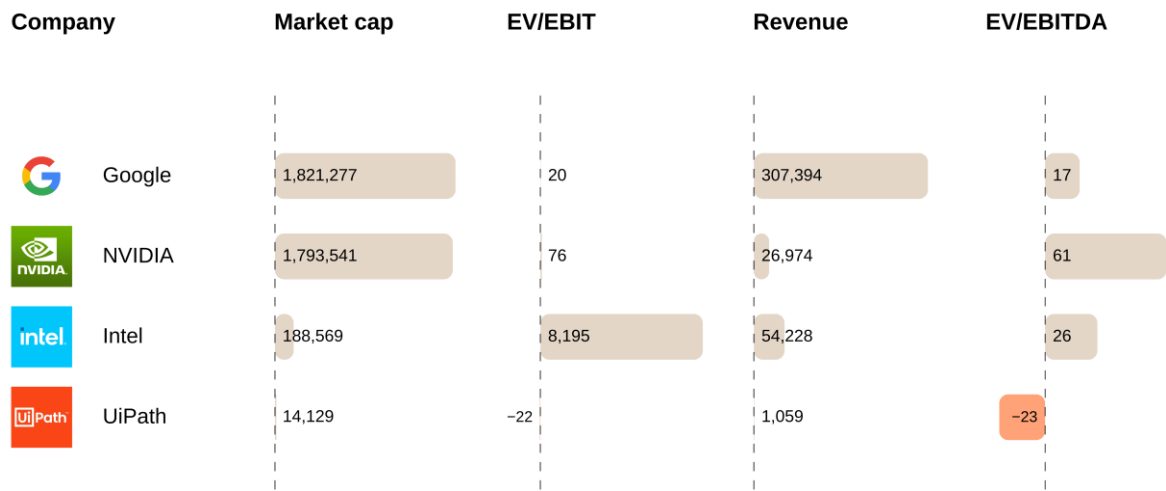


<b>Company</b>	UiPath
<b>Details</b>	Founded: 2005 HQ: New York, United States
<b>Overview</b>	UiPath is a software company that develops robotic process automation and artificial intelligence software.
<b>Year</b>	2023
<b>Sales</b>	USD 1.1B
<b>EBITDA</b>	USD -306.2M
<b>EBIT</b>	USD -324.9M
<b>Market Cap</b>	USD 14.1B
<b>EV</b>	USD 7.1B
<b>EV/EBIT</b>	-21.82
<b>EV/EBITDA</b>	-23.15
<b>Gross Margin (%)</b>	83.07
<b>EBIT Margin (%)</b>	-30.69
<b>Net Margin (%)</b>	-31.02



Chart: Public comps (1/2)

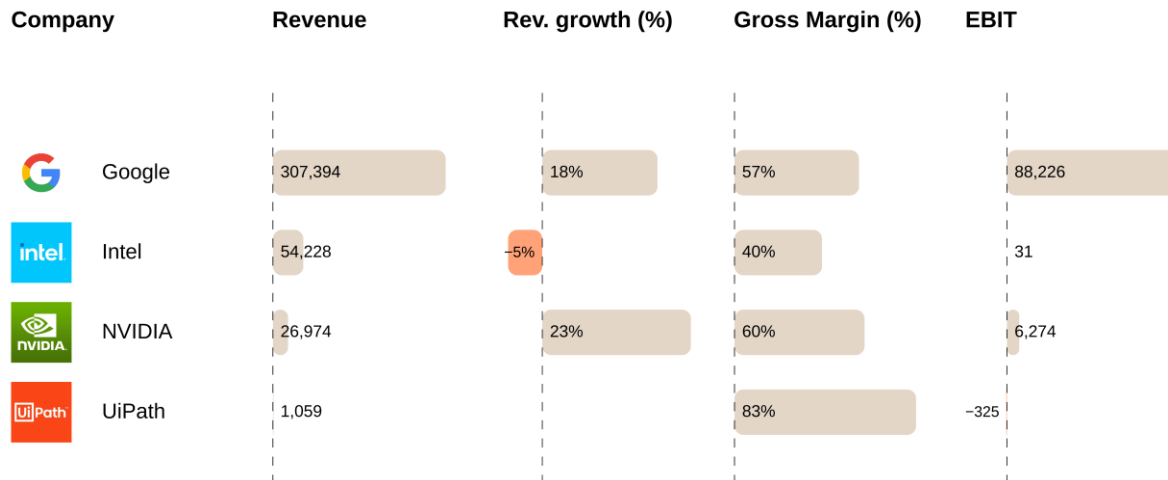
(Market cap and Revenue in \$M)



Source: Wokelo generated, Finnhub financials

## Chart: Public comps (2/2)

(All numbers in \$M)



Source: Wokelo generated, Finnhub financials

## Adjacent segments

Source: Wokelo synthesis

### Adjacent sectors

**AI Hardware:** AI hardware refers to the specialized hardware components and systems designed to support artificial intelligence and machine learning workloads. This includes processors, accelerators, memory systems, and interconnects optimized for AI computations.

**AI Software:** AI software encompasses the tools, frameworks, and libraries used to develop and deploy artificial intelligence and machine learning models. It includes programming languages, deep learning frameworks, data preprocessing tools, and model deployment platforms.

**AI Applications:** AI applications are software programs that utilize artificial intelligence and machine learning techniques to perform specific tasks or solve complex problems. These applications can be found in various industries such as healthcare, finance, retail, and transportation.

### Key segments and products

- AI Processors
- AI Accelerators
- Memory Systems
- Interconnects
- AI Storage

- Programming Languages
- Deep Learning Frameworks
- Data Preprocessing Tools
- Model Deployment Platforms
- AI Development Tools

- Healthcare AI
- Finance AI
- Retail AI
- Transportation AI
- Security AI

**AI Ethics:** AI ethics focuses on the ethical considerations and implications of artificial intelligence technologies. It involves ensuring fairness, transparency, accountability, and privacy in AI systems, as well as addressing potential biases and social impacts.









- Fairness and Bias
- Transparency and Explainability
- Accountability and Governance
- Privacy and Security
- Social Impacts








**AI Research:** AI research involves the study and development of new algorithms, models, and techniques to advance the field of artificial intelligence. It encompasses areas such as natural language processing, computer vision, reinforcement learning, and knowledge representation.

- Natural Language Processing
- Computer Vision
- Reinforcement Learning
- Knowledge Representation
- Machine Learning Algorithms

## Value chain

Note: This is representative based on analysis of companies in Wokelo database

Categories	Role	Key players examples	
<b>Research and Development</b>	This group focuses on advancing the theoretical foundations and developing new algorithms for AI. They are at the forefront of innovation, pushing the boundaries of what AI can do. Their work often leads to the creation of new products and services, and they collaborate closely with academic institutions. Business nuances include securing funding for research, publishing papers, and patenting new technologies.		DeepMind
			OpenAI
			IBM Research
<b>AI Chip Manufacturers</b>	These companies design and produce the specialized processors that power AI computations. Their products are critical for the performance of AI applications, from data centers to edge devices. Business nuances involve heavy investments in R&D, competition for market share, and the need to constantly innovate to keep up with the computational demands of AI.		NVIDIA
			Intel
			AMD
<b>AI Software Providers</b>	This group includes companies that develop AI software platforms, tools, and applications. They enable businesses and developers to build, train, and deploy AI models. Business nuances here include creating user-friendly platforms, ensuring data security and privacy, and providing		Google AI
			Microsoft Azure AI

	scalable solutions that can handle large volumes of data.		Amazon Web Services (AWS)
<b>Data Providers and Annotators</b>	Data is the lifeblood of AI, and this group supplies the high-quality, annotated data necessary for training AI models. They also provide services for data collection, processing, and labeling. Business nuances include maintaining a balance between data quality and cost, ensuring data privacy, and adapting to various data requirements of different AI applications.		Appen
			Lionbridge
			Figure Eight
<b>End-User Applications</b>	This group consists of companies that integrate AI into their products and services to enhance functionality and user experience. They span various industries, including healthcare, finance, automotive, and customer service. Business nuances involve understanding customer needs, ensuring seamless integration of AI into existing products, and staying ahead of the competition by leveraging AI for innovation and improved services.		Tesla (Autopilot)
			IBM Watson
			Salesforce Einstein





## Market insights

### Quantitative insights

**Source - Vantagemarketresearch: Artificial Intelligence Market Size USD 175.63 Billion by 2028** [\(link\)](#)

- The global Artificial Intelligence (AI) market size was \$65.32 billion in 2021.
- It is projected to grow to \$175.63 billion by 2028 with a CAGR of 41.23% from 2021 to 2028.

**Source - Nextmsc: Artificial Intelligence Market Size and Share | Analysis - 2030** [\(link\)](#)

- The market size of Artificial Intelligence (AI) was valued at \$95.60 billion in 2021.
- It is projected to grow to \$1,847.58 billion by 2030.
- The Compound Annual Growth Rate (CAGR) from 2022 to 2030 is expected to be 32.9%.

**Source - Marketdataforecast: Artificial Intelligence (AI) Market Size, Share, Growth Report** [\(link\)](#)

- The worldwide market for artificial intelligence (AI) is set to increase from USD 212.48 billion in 2023 to reach USD 1042.05 billion by 2028, growing at a CAGR of 37.44%.
- In 2021, the artificial intelligence market size was valued at USD 52.6 billion and is projected to reach USD 67.6 billion in the same year.
- Expectations suggest a significant uptick in the AI market, predicting a growth rate of 28.5% between 2022 and 2027, culminating in a market size of USD 236.8 billion by 2027.

**Source - Ibm: What is Computer Vision? | IBM** [\(link\)](#)

- The market size of computer vision within the context of Artificial Intelligence Computing is expected to reach USD 48.6 billion by the year 2022.

**Source - Nvidia: Computer vision – What Is It and Why Does It Matter? - NVIDIA** [\(link\)](#)

- The computer vision market is projected to grow at a rate of 47% annually.
- The market size of computer vision is expected to reach \$25 billion by the year 2023.

**Source - Polarismarketresearch: Artificial Intelligence (AI) in Computer Vision Market Share, Size ...** [\(link\)](#)

- The market for AI in computer vision was valued at USD 16.05 billion in 2020.
- The market is projected to grow at a CAGR of 26.2% during the forecast period.
- Revenue is forecasted to reach USD 81.10 billion by 2028.

**Source - Marketdataforecast: Artificial Intelligence (AI) Robots Market Size (2023-2028)** [\(link\)](#)

- The Artificial Intelligence (AI) Robots Market was valued at \$3.24 billion in 2022.

- It is forecasted to reach \$23.92 billion by 2028, growing at a CAGR of 28.41% from 2023 to 2028.

**Source - Technavio: Artificial Intelligence (AI) Robots Market Size, Share & Trends to 2027** [\(link\)](#)

- The market size of Robotics in the context of Artificial Intelligence Computing is estimated to grow by USD 19,157.78 million from 2022 to 2027. The growth rate is projected at a CAGR of 29.37%. APAC is expected to contribute 45% to the market by 2027.

**Source - Statzon: Global Artificial Intelligence Robots Market Grows at 25% to ... - Statzon** [\(link\)](#)

- The AI robots market was valued at USD 9.3 billion in 2022 and is projected to grow to USD 29.2 billion by 2027, at a CAGR of 25.7%.
- Markets and Markets estimated the AI robotics market at USD 9.6 billion in 2022, with a forecast for growth to USD 35.3 billion by 2026, suggesting a CAGR of 38.6%.
- Apollo Research Reports placed the AI robotics market value at USD 4.1 billion in 2022, anticipating growth to USD 52.6 billion by 2032, at a CAGR of 29.7%.

**Source - Omrglobal: Artificial Intelligence Market Size, Growth, Industry Forecast to 2028** [\(link\)](#)

- The global AI market is expected to grow at a CAGR of 38.7%.
- Total global investment in AI increased by 40% from 2019 to 2020.
- North America holds the largest share of the global AI market.

**Source - Analyticsinsight: Data Science Market: Size, Trends, and Forecast to 2024 - Analytics Insight** [\(link\)](#)

- The Data Science market size demonstrates significant growth with a Compound Annual Growth Rate (CAGR) of 30.0%, escalating from USD 37.9 billion in 2019 to a projected USD 140.9 billion by 2024.
- The volume of data has explosively grown in recent years, with approximately 90% of the world's data being generated in the last two years alone.

**Source - Bigdataanalyticsnews: 50+ Incredible Big Data Statistics for 2024: Facts, Market Size ...** [\(link\)](#)

- The market size of the Big Data industry, which encompasses data science as a component, is forecasted to grow from \$77 billion at the end of 2023 to \$103 billion by 2027.
- An estimated 147 zettabytes of data will be generated in 2024, with a projection of reaching 181 zettabytes by 2025.
- The industry expects a 9% growth rate in 2024, with a projected 7% growth rate annually from 2025 to 2027.

**Source - Straitsresearch: Natural Language Processing Market - Straits Research** [\(link\)](#)

- The natural language processing market was valued at USD 13.5 billion in 2021.

- It is projected to reach USD 91 billion by 2030, with a compound annual growth rate (CAGR) of 27%.

## Emerging trends

### AI Potential and Impact <sup>[1-11]</sup>

- Artificial Intelligence computing holds the potential to contribute \$15.7 trillion to the global economy by 2035, with China and the United States projected to account for nearly 70% of the global impact. Key players, such as Microsoft, Google Cloud, AWS, IBM, SAP, and Salesforce, are contributing to numerous advancements in AI, pushing it beyond its previous realm of researchers and technology industries into the public consciousness. AI in computing facilitates tasks requiring human intellect, uses accelerated hardware and software for machine learning, integrates with cloud computing technology, and is poised to surpass all forms of human intelligence. Applications span various sectors, like NLP, robotics, image analysis, financial services, healthcare, entertainment, among others, with AI finding its role in our everyday lives via tools such as ChatGPT, Google Maps, smart assistants like Alexa and Siri, among others. However, AI advancements also pose significant ethical considerations, ranging from bias in algorithms, job displacement, privacy concerns, to data ownership issues.

### AI vs Traditional Computing <sup>[2-8]</sup>

- AI's emergence has caused a significant shift from regular computing, relying on predefined instructions, to intelligent systems that learn and adapt without explicit programming. Rather than relying purely on deterministic decisions of the traditional computing, AI decision-making involves probabilistic reasoning, making AI excel in dynamic environments and diversely set problem arenas. Considering AI as an extension of human cognitive abilities, advancements in AI include natural language processing, image recognition, and the potential for AI to replicate and even augment human cognitive abilities, intertwining with regular computing ever more deeply. With this, hybrid systems combining the precision of traditional computing with the adaptability of AI are emerging.

### AI in Cloud Computing & Future of Work <sup>[3,7-9,11]</sup>

- AI brings transformative changes to both cloud computing and the future of work. AI's deep integration with cloud computing enhances revenue-generation opportunities, allows mass personalization of products/services, automates repetitive tasks, and optimizes cost and resource scaling, among others. AI on cloud reduces costs for organizations, enhances productivity, and simplifies routine processes in cloud infrastructure. In the context of employment, AI's growth is not expected to make human workers obsolete in the near future. Instead, it has the potential to fuel job creation, creating new sectors and transforming existing ones. Over 90% of leading businesses have invested in AI, reporting productivity increases. Through applications in sectors like healthcare, education, and productivity, AI is likely to democratize services that are currently expensive or inaccessible,

revolutionizing the software industry and marking the biggest shift in computing since the transition from command-line interfaces to graphical user interfaces.

Sources:

- [1] Simplilearn (30th Nov 2023) - What is Artificial Intelligence and Why It Matters in 2024? [\(link\)](#)
- [2] Rigb (12th Dec 2023) - What's the Difference Between AI and Regular Computing? [\(link\)](#)
- [3] Aithority (11th Nov 2023) - Top 20 Uses of Artificial Intelligence In Cloud Computing For 2024 [\(link\)](#)
- [4] Zdnet (21st Apr 2023) - What is AI? Everything to know about artificial intelligence [\(link\)](#)
- [5] Ibm (09th Feb 2024) - The most important AI trends in 2024 [\(link\)](#)
- [6] Rstreet (11th Jan 2024) - Accelerated Computing, Artificial Intelligence and the Computational Revolution [\(link\)](#)
- [7] Gatesnotes (09th Nov 2023) - AI is about to completely change how you use computers [\(link\)](#)
- [8] Itprotoday (03rd Aug 2023) - 5 Ways the AI Boom Could Reshape Cloud Computing [\(link\)](#)
- [9] Forbes (10th Mar 2021) - How AI Will Impact The Future Of Work And Life [\(link\)](#)
- [10] Builtin (27th Jul 2023) - Artificial Intelligence (AI): What Is AI and How Does It Work? [\(link\)](#)
- [11] Dataconomy (07th Nov 2022) - Big Data And Artificial Intelligence: What's The Future For Them? [\(link\)](#)

## Recent innovations and product launches

### Advancements in AI Chip Technology <sup>[1-2]</sup>

- Engineers from the University of Pennsylvania have developed a unique chip that harnesses light waves to perform AI computations, aiming to boost processing speeds while minimizing energy consumption. This silicon-photonics (SiPh) chip hints at the future of AI chip design, a merger of nanoscale material manipulation with the SiPh platform. The chip focuses on speeding up vector-matrix multiplication, a crucial operation for neural networks. Parallel to this, researchers have furthered the development of artificial neural networks using silicon microresonators. These silicon-based constructs trap and control light properties, leading to the sensitive handling of wavelength which serves as weight banks in photonic neural networks.

### Rise of Quantum Computing in AI <sup>[3-4]</sup>

- The realm of Quantum machine learning is gradually expanding as researchers combine the foundations of artificial intelligence with quantum computing. Heavyweight technology companies, along with numerous startups, are actively exploring the potential of quantum machine learning. Current investigations include assessing the benefits of quantum algorithms over classical machine learning processes. Proof-of-principle experiments, including those conducted on Google's Sycamore quantum computers, show the promise of quantum machine learning with a considerable speed advantage over classical methods. The real-world applications of quantum computing in AI are however still under investigation. Quantum computing is now being seen as an influential trend that could potentially outshine today's chip-based supercomputers.

### AI Boost in Cloud Computing and Infrastructure <sup>[5-8]</sup>

- Noteworthy strides in AI-infused cloud computing are taking place globally. The University of Texas at Austin is preparing to launch a Center for Generative AI, powered by a massive

cluster of GPUs, with a focus on advancing AI in the biosciences and healthcare sectors. Samsung is boosting its prospects in the AI market by establishing an AGI computing lab in Silicon Valley. Across the globe, in Beijing, a public AI computing platform named Shangzhuang project is set to launch, powered by Beijing Energy Holding. It aims to boost Beijing's AI ecosystem and the real economy. Meanwhile, the AI-as-a-service market is envisioned to be a booming \$55 billion market by 2028, foreseeing that AI will accelerate cloud computing spending, projected to reach \$600 billion in 2023.

### **Capacity Enhancement through AI Innovations** <sup>[9-13]</sup>

- Futuristic AI technologies designed to increase capacity and enhance productivity are surfacing across military and medical sectors. Drones furnished with computer vision models for reconnaissance and security are being introduced for future battlefield technologies. AI systems being designed intend to empower commanders to sense and act quicker than ever before. In contrast, researchers at the University of Florida and NVIDIA have designed an AI program capable of generating doctors' notes that are virtually indistinguishable from those written by actual doctors. One Stop Systems, focusing on industrial AI, has deployed its AI software to U.S. troops, enabling rapid decision-making on-site. Additionally, MIT Schwarzman College of Computing's funding of AI and HCI projects to boost workspaces, points at the future AI-powered utopia.

### **AI Research and Education Initiatives** <sup>[14-16]</sup>

- The effort to advance AI research and innovation is evident through numerous initiatives at renowned universities. An AI computing center named "Empire AI" is being spearheaded by New York governor Kathy Hochul. Cornell, Columbia, and several other universities have joined the consortium, aiming to position New York as a global leader in AI innovation. Columbia University, in particular, has expressed its key role in the initiative to drive responsible innovation to bolster New York's economy and U.S. national security. These initiatives signal a collective effort towards never-before-seen advancements in AI research and development.

Sources:

- [1] Upenn (21st Feb 2024) - New chip opens door to AI computing at light speed | Penn Today [\(link\)](#)
- [2] Miragenews (22nd Feb 2024) - Silicon Microresonators Boost Artificial Neural Networks [\(link\)](#)
- [3] Nature (02nd Jan 2024) - The AI-quantum computing mash-up: will it revolutionize science? [\(link\)](#)
- [4] Investors (16th Jun 2023) - After Artificial Intelligence, Quantum Computing Could Be The Next Big Thing [\(link\)](#)
- [5] Datacenterdynamics (30th Jan 2024) - University of Texas to host generative AI computing cluster [\(link\)](#)
- [6] Notebookcheck (21st Feb 2024) - Samsung establishes Artificial General Intelligence Computing Lab led by former Google TPU developer in Silicon Valley [\(link\)](#)
- [7] Datacenterdynamics (02nd Jan 2024) - New public AI computing platform launched in Beijing [\(link\)](#)
- [8] Globalxetfs (13th Oct 2023) - Generative AI Delivers a Boost to Cloud Computing [\(link\)](#)
- [9] Westpoint (02nd Nov 2023) - Fighting for Seconds: Warfare at the Speed of Artificial Intelligence - Modern War Institute [\(link\)](#)
- [10] Ufhealth (07th Dec 2023) - Medical AI tool from UF, NVIDIA gets human thumbs-up in first study [\(link\)](#)
- [11] Twst (02nd Oct 2023) - One Stop Systems (NASDAQ:OSS) Puts Artificial Intelligence Computing into the Hands of Battlefield Officers [\(link\)](#)
- [12] Ibm (10th Jan 2024) - Breaking down the advantages and disadvantages of artificial intelligence [\(link\)](#)
- [13] Scitechdaily (22nd Nov 2023) - The Future of AI: Self-Learning Machines Could Replace Current Artificial Neural

Networks [\(link\)](#)

[14] City-journal (12th Jan 2024) - Empire AI: Another New York Boondoggle? [\(link\)](#)

[15] Spectrumlocalnews (08th Jan 2024) - Hochul unveils plan for artificial intelligence consortium [\(link\)](#)

[16] Columbia (08th Jan 2024) - Columbia Joins New York Consortium on Artificial Intelligence [\(link\)](#)



## Annexure

### Key customer segments

- **Consumer Electronics Manufacturers<sup>[1-2]</sup>**: Companies that are integrating Groq's AI technology to enhance real-time language processing capabilities in consumer electronics, aiming to deliver immersive experiences for end-users.
- **Financial Technology Companies<sup>[1]</sup>**: FinTech enterprises leveraging Groq's low-latency AI inference engines to innovate in fields like high-frequency trading, risk assessment, and algorithmic decision-making.
- **High-Performance Computing Centers<sup>[1]</sup>**: Research and government institutions utilizing Groq hardware for AI Testbed systems to meet the computational requirements of large-scale, complex simulations and models.
- **Cybersecurity Solutions Providers<sup>[1]</sup>**: Organizations focused on national security and defense that implement Groq's AI hardware for fast and accurate cybersecurity anomaly detection.
- **AI and Machine Learning Platforms<sup>[1,3]</sup>**: Cloud service providers and platforms that offer machine learning and natural language processing services powered by Groq's technology to enhance the performance and scalability of AI applications.

[1] Groq - Groq TruePoint Technology [\(link\)](#)

[2] Groq (2021-12-16) - Talent, Locations, and Being Different. Groq's Growth Success - Groq [\(link\)](#)

[3] Groq - GROQ-ROCKS-NEURAL-NETWORKS.pdf [\(link\)](#)

### Other relevant news

Publisher	Date	Title
Prnewswire <a href="#">(link)</a>	Oct 27, 2021	Hiring - Groq Attracts Industry Best from Fortune 500 Companies and Beyond
Prnewswire <a href="#">(link)</a>	Oct 07, 2021	Hiring - Groq Fuels Talent Growth Beyond Expectations
Insidehpc <a href="#">(link)</a>	Nov 21, 2023	Customer - ALCF: Groq AI Online Workshop, Dec. 6-7 - High-Performance Computing News Analysis
Datacenterdynamics <a href="#">(link)</a>	Oct 19, 2023	Customer - Argonne Lab deploys Groq AI hardware - DCD
Prnewswire <a href="#">(link)</a>	Oct 25, 2022	Customer - US Army Analytics Group Confirms 1000x Performant Cybersecurity Technology by Entanglement AI™, Run on Groq ...





**Prnewswire** [\(link\)](#)

Nov 17, 2021

Customer - Groq Accelerates COVID Drug  
Discovery by 333x for Argonne National  
Laboratory





## References

#	Publisher	Title (link)
1	Company Website	Groq <a href="#">(link)</a>
2	Crunchbase Pro	Groq <a href="#">(link)</a>
3	PR Newswire	Groq Acquires Dataflow Systems Pioneer Maxeler Technologies <a href="#">(link)</a>
4	PR Newswire	Groq™ Partners With New Customer, OneNano™, Providing Ultra-low Latency for Next Generation Cryptocurrency Exchange (CEX) <a href="#">(link)</a>
5	Vantagemarketresearch	Artificial Intelligence Market Size USD 175.63 Billion by 2028 <a href="#">(link)</a>
6	Nextmsc	Artificial Intelligence Market Size and Share   Analysis - 2030 <a href="#">(link)</a>
7	Marketdataforecast	Artificial Intelligence (AI) Market Size, Share, Growth Report <a href="#">(link)</a>
8	Ibm	What is Computer Vision?   IBM <a href="#">(link)</a>
9	Nvidia	Computer vision – What Is It and Why Does It Matter? - NVIDIA <a href="#">(link)</a>
10	Polarismarketresearch	Artificial Intelligence (AI) in Computer Vision Market Share, Size ... <a href="#">(link)</a>
11	Marketdataforecast	Artificial Intelligence (AI) Robots Market Size (2023-2028) <a href="#">(link)</a>
12	Technavio	Artificial Intelligence (AI) Robots Market Size, Share & Trends to 2027 <a href="#">(link)</a>
13	Statzon	Global Artificial Intelligence Robots Market Grows at 25% to ... - Statzon <a href="#">(link)</a>
14	Omrglobal	Artificial Intelligence Market Size, Growth, Industry Forecast to 2028 <a href="#">(link)</a>
15	Analyticsinsight	Data Science Market: Size, Trends, and Forecast to 2024 - Analytics Insight <a href="#">(link)</a>
16	Bigdataanalyticsnews	50+ Incredible Big Data Statistics for 2024: Facts, Market Size ... <a href="#">(link)</a>
17	Straitsresearch	Natural Language Processing Market - Straits Research <a href="#">(link)</a>



## Important Notice:

Wokelo AI (Wokelo Inc., HQ in Redmond, WA) is a generative-AI powered research assistant platform and synthesis tool that uses LLMs (such as GPT-4) and other underlying models. Services provided by Wokelo AI are currently offered on a private beta basis and no warranties are provided for their accuracy, completeness, or appropriateness for the user's purposes. This report has been prepared by an AI agent (Wokelo.ai) on behalf of the end-user for their specific research topic and specific source and attribute selections. It includes sections interpreted, synthesized, and written by AI. Neither the report nor any part of its contents is intended or suited to constitute the basis of any investment decision regarding any company operating in the markets covered by this report or any similar markets (including, without limitation, the purchase of any securities of any listed company or in connection with the listing of any company). This report may contain certain forward-looking statements (i.e., estimates, forecasts, and projections about industries, trends, players, and consumers). The services may contain links to third party websites or services that are not owned or controlled by Wokelo. The provider has no control over and assumes no responsibility for the content, privacy policies, or practices of any third-party websites or services. There may be information on the application or report that contains typographical errors, inaccuracies, or omissions, including descriptions, pricing, availability, and various other information. Wokelo.ai does not assume any responsibility towards any persons for the correctness and completeness of the information contained in this report. Wokelo AI (Wokelo Inc.) makes every effort to keep our AI model and other content on our application as up-to-date and accurate as possible. However, we disclaim any warranty or representation, expressed or implied, about our service's accuracy, completeness or appropriateness for the user's purposes. You are responsible for Content that you post on or through Service, including its legality, reliability, and appropriateness. The application is our proprietary property and all source code, databases, functionality, software, and graphics on the application are owned or controlled by us or licensed to us and are protected by copyright and trademark laws and various other intellectual property rights and unfair competition laws of the United States, international copyright laws, and international conventions. Users must be at least 18 years of age to use the services. Customer will not, directly, or indirectly: reverse engineer, decompile, disassemble or otherwise attempt to discover the source code, object code or underlying structure, ideas, know-how, or algorithms relevant to the Services or any software, documentation, or data related to the Services or Software; modify, translate, or create derivative works based on the Services or any Software (except to the extent expressly permitted by Company or authorized within the Services)

Copyright © 2024 Wokelo.ai (Wokelo Inc.). All rights reserved





Wokelo.ai

Wokelo Inc. | Seattle, USA  
[support@wokelo.ai](mailto:support@wokelo.ai)

© 2024 Wokelo Inc. All rights reserved