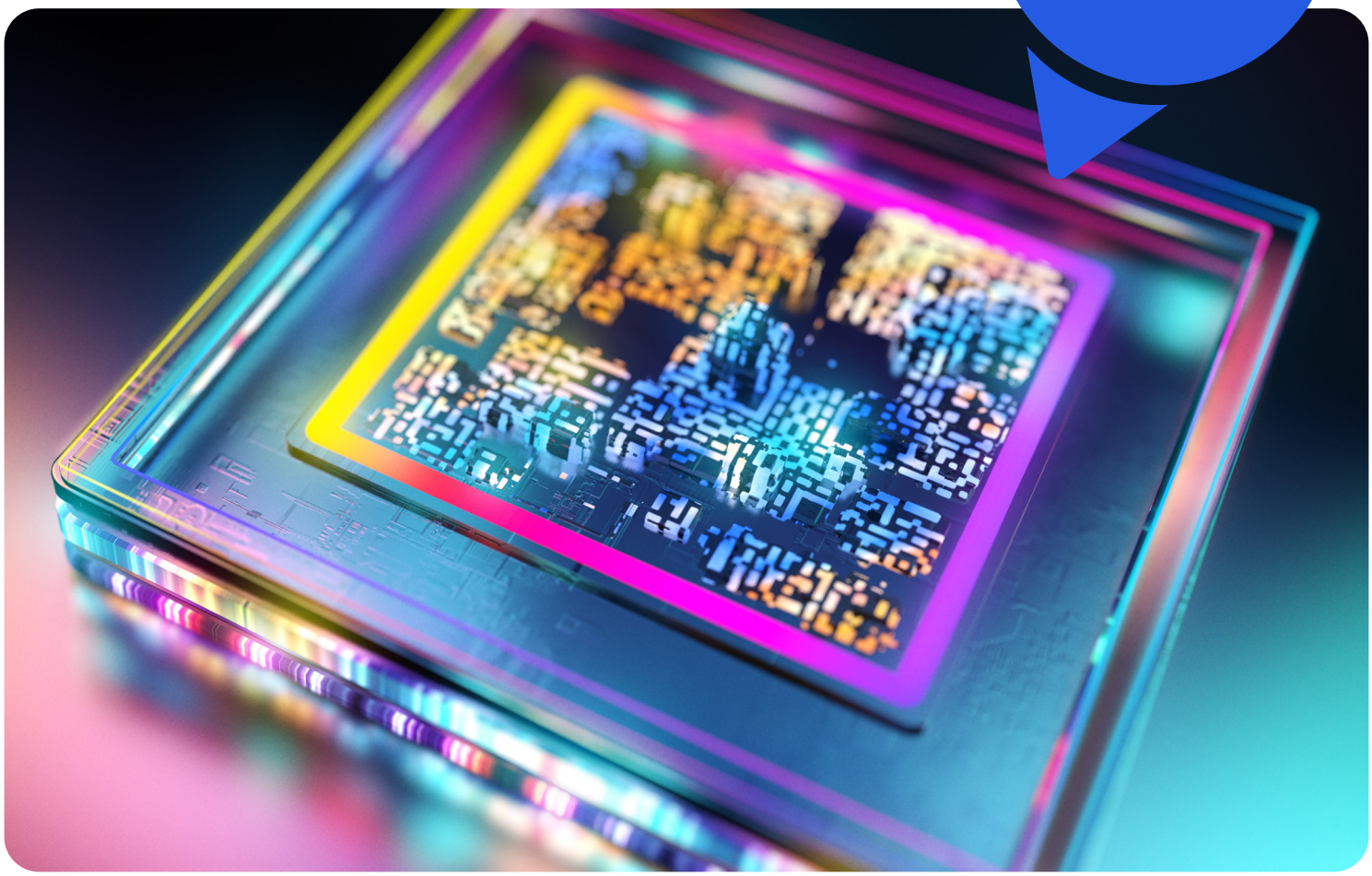


Disruptive Innovations X

Ten More Things to Stop and Think About

Citi GPS: Global Perspectives & Solutions
May 2024

Rob Garlick
Head of Innovation & Technology
Citi Global Insights
robert.j.garlick@citi.com

Ashiq Mubarack
U.S. SMID Biotech Analyst
Citi Research
ashiq.mubarack@citi.com

Yigal Nochomovitz, Ph.D.
U.S. SMID Biotech Analyst
Citi Research
yigal.nochomovitz@citi.com

Peter Verdult, CFA
Managing Director, Global Pharmaceuticals
Citi Research
peter.verdult@citi.com

Ronald Josey
U.S. Internet Analyst
Citi Research
ronald.josey@citi.com

Wenyan Fei
Analytical Insights Team
Citi Global Insights
wenyan.fei@citi.com

Adam Spielman
Head of Future of Health
Citi Global Insights
adam.spielman@citi.com

Amy Thompson
Social Economist
Citi Global Insights
amy1.thompson@citi.com

Canberk Benning
Future of Health Team
Citi Global Insights
canberk.benning@citi.com

Jason Gursky
U.S. Aerospace and Defense Analyst
Citi Research
jason.gursky@citi.com

Peter Lee
Asia Pac Semiconductors & Components Analyst
Citi Research
peter.sc.lee@citi.com

Avinash Mundhra
Business & Professional Services
Citi Research
avinash.mundhra@citi.com

Marc Van'T Sant
Commercial & Professional Services
Citi Research
marc.vantsant@citi.com

Arthur Truslove
Commercial & Professional Services
Citi Research
arthur.truslove@citi.com

Thomas A Singlehurst, CFA
Head of EMEA EMT Sector
Citi Research
thomas.singlehurst@citi.com

Ygal Arounian
U.S. Internet Analyst
Citi Global Insights
ygal.arounian@citi.com

Maxwell Moore
U.S Internet Research Team
Citi Research
maxwell.moore@citi.com

Tahmid Quddus Islam
Citi Global Insights
tahmid.quddus.islam@citi.com

Ronit Ghose, CFA
Head of Future of Finance
Citi Global Insights
ronit.ghose@citi.com

Sophia Bantanidis
Future of Finance Analyst
Citi Global Insights
sophia.bantanidis@citi.com

Kaiwan Master
Future of Finance Analyst
Citi Global Insights
kaiwan.hoshang.master@citi.com

Ronak S Shah
Future of Finance Team
Citi Global Insights
ronak.sharad.shah@citi.com

Maggie Xueting Lin
Commodities
Citi Research
maggie.x.lin@citi.com

DISRUPTIVE INNOVATIONS X

Ten More Things to Stop and Think About

What's the next big thing coming down the pike?

Citi's Disruptive Innovations series addresses this vital question head-on.

This is the 10th edition, each showcasing 'Ten Things to Stop and Think About', with a quick write-up on each. As such, with this year's picks we now have a catalog of 100 innovations, and we've added data sets and process around identifying innovation themes, too.

The report offers another diverse list of ideas from Citi experts as follows:

Antibody Drug Conjugates – ADCs combine the potency of chemotherapy with the precise targeting on monoclonal antibodies in the war against cancer. Approvals have been granted and revenues look set to rise.

Autonomous Agents – Digital assistants will increasingly be powered by Artificial Intelligence to carry out useful tasks. Large language models have captured the imagination over the last 18 months. Large action models could be next.

FemTech – Women are less well served than men in many health areas, including diagnosis, therapeutics, and pain management. FemTech can help close the gender health gap.

Joint All-Domain Command and Control – JADC2 uses sensors, communication, and AI capabilities to materially reduce critical decision-making timelines and deal with military threats across air, land, sea, space and cyberspace.

Neuromorphic Computing Architecture – Computing will take on more human brain-like architectures to reduce the bottlenecks between logic and memory as AI-related computing demand rises.

Piezoelectric Roads – Smart roads could harness electricity generated by vehicles travelling across these new roads.

Quantum Sensing – Sensors are the eyes and ears of our technological world. Quantum sensors can provide an unprecedented level of detail to generate new insights and aid decision making across multiple industries.

Retail Media – Digital adverts tied directly into a retailer's website, or on third-party websites, are growing fast in part due to stricter data privacy standards.

Smart Money – Generative AI has revolutionary potential in financial services, expanding the uses cases of ML and AI into unstructured data to change customer service, sales and marketing, content generation and coding.

White & Gold Hydrogen – Naturally occurring hydrogen sources are trapped below the earth's surface in underground deposits. Exploration techniques are being looked at by governments and private companies alike.

We hope you enjoy digging further into these in Citi's Disruptive Innovations X.

Disruptive Innovations X: Ten More Things to Stop and Think About

1

Antibody Drug Conjugates

Antibody drug conjugates are ***100-1,000x*** more potent than chemotherapy in killing cancer cells.

2

Autonomous Agents

AI assistants could reduce headcount in customer support functions ***20%-30% by 2026.***

3

JADC2

Joint All-Domain Command and Control (JADC2) could reduce the U.S. military's decision-making times ***from days or weeks to minutes.***

4

FemTech

Only 4% of healthcare research dollars go to women's health issues, underlining the scale of the unmet need.

5

Neuromorphic Computing Architecture

Neuromorphic computing could see ***adoption starting by end-2026***.

6

Piezoelectric Roads

Converting 5% of U.S. roads to piezoelectric could power ***5,000 households*** for a year.

7

Quantum Sensing

1 second per 300 billion years: Accuracy of optical atomic clocks using quantum sensing.

8

Retail Media

$110 billion Global retail media market (excluding China) forecast by 2027.

9

Smart Money

AI in Finance: ***2/3 of work*** in banking and insurance could be automated or augmented by AI.

10

White and Gold Hydrogen

Drilling for underground ***"white" hydrogen*** could be a clean source of hydrogen fuel.

Contents

Introduction

Rob Garlick
Head of Innovation &Technology
Citi Global Insights

In 10 editions of Citi's Disruptive Innovations, we've identified several innovations early in their lifecycles. And of course, we've had a number that have failed to meaningfully take off yet – like flying taxis.

But rather than use this intro to highlight our misses or do selective victory laps, here are a few data sets that may help you think about where we've been and where we're going.

Engineering objectivity

Citi GPS started in 2011 and included a diverse set of thematic reports. At the time, clients rightly asked whether there was a way of objectively looking at themes versus each other. It turns out there wasn't.

This led to the birth of a related product in 2013 from Citi Research called the Global Theme Machine. The Theme Machine weights exposure of over 5,000 publicly traded stocks to >90 themes, with significant analysis added using style factors, correlation, and other quantitative methodologies. In addition, the Theme Machine allows us to answer simple questions such as how different themes have performed over various time horizons. For example, in Figure 1, we show the best- and worst-performing themes from the original list over the past 10 years.

Figure 1. Top- and Bottom-Ranked Themes – 10Y Return

Top & Bottom Ranked Themes - Last 10Y Return			
TOP RANKED THEMES		BOTTOM RANKED THEMES	
Cloud Computing	301.2%	Value Healthcare Spend	94.0%
Defence	300.4%	Immunotherapy	86.6%
Data Storage (Big Data)	262.3%	EM Consumer	86.2%
Biotech	261.9%	Urbanisation	81.2%
Cyber Security	258.4%	Agriculture Demand	80.3%

Source: Citi Research, MSCI

New themes have been added over time. For example, Energy Storage was added in 2015, Virtual Reality in 2016, and Artificial Intelligence in 2017. When we look at the last 3-year performance in Figure 2, the picture is very different from the 10 year-view. While many themes seem to be enduring in theory, they often fall in and out of favor. The Theme Machine helps explain why. Timing also matters, and ideas can often be premature, taking a long time to go from germination to mass adoption.

Figure 2. Top- and Bottom-Ranked Themes – 3Y Return

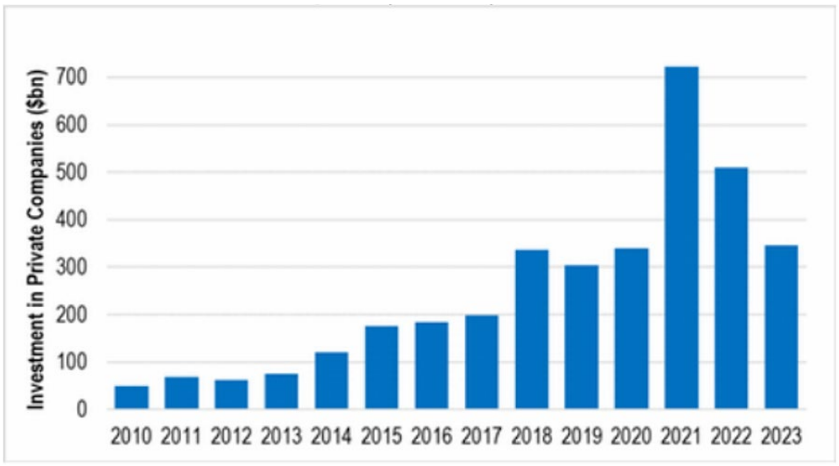
Top & Bottom Ranked Themes - Last 3Y Return			
TOP RANKED THEMES		BOTTOM RANKED THEMES	
Fossil Fuels	56.0%	Internet-driven business models	-25.0%
Nuclear Energy	45.8%	DNA/Genetics	-28.5%
Belt & Road	43.7%	Feminine Health & FemTech	-36.6%
Infrastructure	39.2%	Creator Economy	-36.8%
Biofuels	37.1%	Immunotherapy	-44.2%

Source: Citi Research, MSCI

Is the high innovation period behind us?

The Theme Machine only looks at listed companies. A large part of new innovation of course takes place via Venture Capital (VC) and we analyze which themes the VC's are backing in another Citi product called Mapping Innovation[1].' As seen in the chart below, the private market funding environment for innovation soared in the decade into 2021, rising 11-fold. Rising interest rates punctuated the VC investment wave with 2023 capital investment declining 52% to levels seen 6 years ago. Will a decline in innovation follow?

Figure 3. Investment in Private Companies (2010-2023)



Source: Pitchbook Data Inc

Citi Global Data Insights (CGDI) notes that examining the patent landscape globally, a step earlier in the innovation life cycle than VC funding, patents declined globally in both 2022 (-7%) and 2023 (-13%). This may have been a Covid-19 anomaly, but combined with declines in VC capital being deployed, it is worth noting.

The hope of course is that Generative AI is about to unleash a significant new wave of innovation, capital, and productivity.

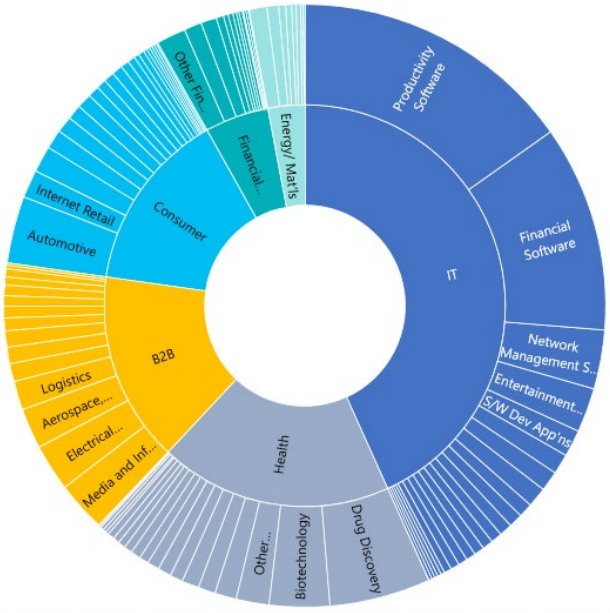
A generation of new Einsteins would be great, but the capabilities of bigger computers have become increasingly significant. AI became the largest area of VC funding in 2023. It may be surprising how concentrated the VC market is with funding dominated by two sectors: technology (38%) and healthcare (19%).

Our Disruptive Innovation series has been more diverse, with 18% of our ideas in the Healthcare sector and 14% in Technology. Green Energy ideas occupied the top slot (at 20%), followed by Finance (13%), Consumer (12%), Industrials (11%, including Transportation), and Communications (8%).

Exciting and substantial change areas exist in Technology and Healthcare, but several other huge changes are taking place including the retooling of the world's energy and power sectors, a reimaging of the world's transport sector (from autonomous vehicles to air-taxis), or the growth of Blockchain and Web3.0 in Finance. So, despite the pullbacks in VC funding and patents, there's no shortage of interesting innovation.

[1] Mapping the Innovation Landscape: Comparing 100 Emerging & Disruptive Technologies, Business Models, and Ideas

Figure 4. Investment in Private Companies by Overall Sector and Specific Industry 2023

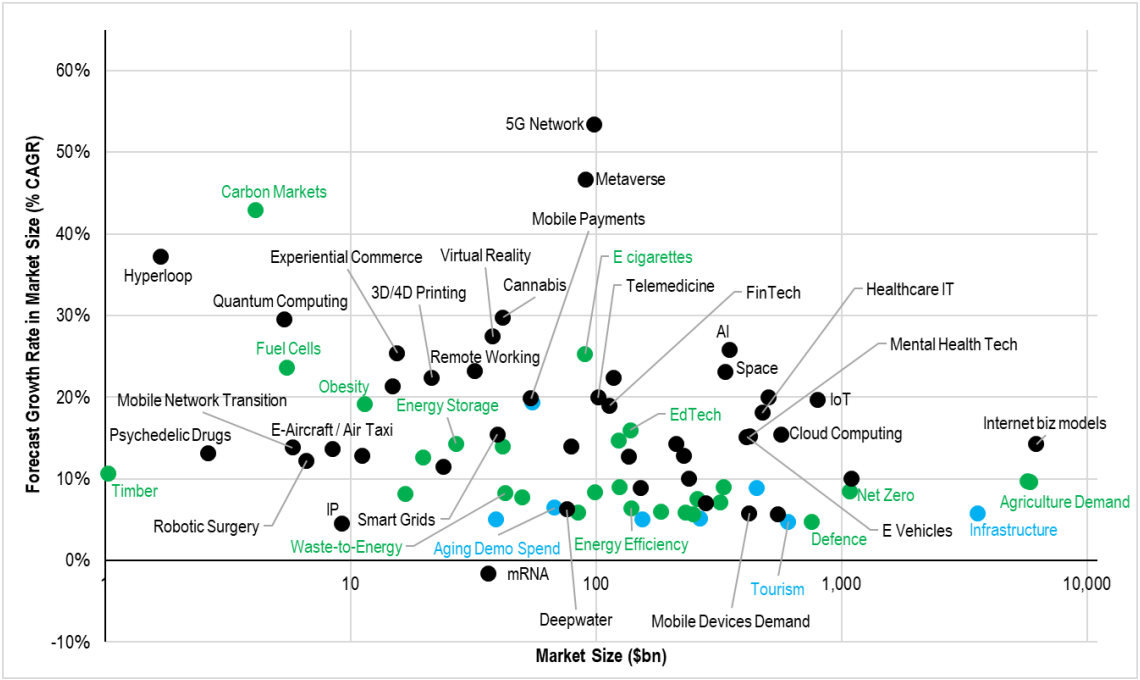


Source: PitchBook Data Inc.

How big?

A big factor behind any new area of innovation is how big the market can become, a measure known as TAM (Total Addressable Market). Citi's Mapping Innovation report offers a unique data set on 100 theme TAMs out to 2030. Many of these cover ideas from our Disruptive Innovations series.

Figure 5. Forecast Growth Rate (2023-2030) vs. Current Market Size (2023) for Areas of Innovation



Black = Technology; Green = Sustainability; Blue = Prosperity
Source: Various 3rd-Party Sources

When looking at 2023 TAMs versus CAGRs to 2030, we note: as expected, AI has a combination of a high current market size and high growth, but space and IOT are also in this top right quadrant; metaverse and 5G networks do not get as much airplay currently, but still have very high growth rates and large TAMs; bottom-right, agriculture demand has a large TAM and robust growth CAGR; in the top-left quadrant, Carbon Markets forecast growth stands out.

What's coming next?

Since our first Disruptive Innovation report in 2013, Citi has added data sets to allow analysis of a growing list of innovation ideas. A question that often gets asked when we present a long list of ideas and data is: what's next? This report aims to provide the answers.

We ask Citi experts for new areas of innovation they find interesting and select a varied list that we hope will make readers stop and think. The summary of the ten ideas from this 10th report is found on page 3, with more detail in the pages that follow.

How our next 10 innovation ideas captured in this report plug into the innovation data sets described above is of course yet to be determined.

Some areas are already substantial in revenue generation – such as Antibody Drug Conjugates or Retail Media. In many other cases, the topics could become big one day but are still nascent – such as Neuromorphic Computing, Piezoelectric Roads, or White Hydrogen.

In other cases, it appears that new markets could grow very quickly – such as Autonomous Agents, FemTech, Joint All-Domain Command & Control, GenAI in Finance, or Quantum Sensors.

For some of these topics, we already have in-depth reports that can provide more detail. Others are coming, and we welcome feedback on which areas would benefit from further exploration, or indeed which topics we may have missed.

In the meantime, we hope you enjoy dipping into these fresh innovation ideas.

Antibody Drug Conjugates

Ashiq Mubarack
U.S. SMID Biotech Analyst
Citi Research

Yigal D Nochomovitz, Ph.D.
U.S. SMID Biotech Analyst
Citi Research

Peter Verdult
Managing Director, Global Pharmaceuticals
Citi Research

Antibody drug conjugates (ADCs) have become a critical weapon in the war against cancer and have already disrupted the treatment of many cancer types.

Conceptually, ADCs combine the potency of chemotherapy with the precise cancer-targeting capabilities of monoclonal antibodies into a safer and more effective package. Sales of the eleven currently FDA-approved ADCs are expected to more than double by 2026 to around $10.3 billion in the U.S. alone according to a consensus analyst estimate from VisibleAlpha.

Despite this growth, we believe the ADC storyline remains in its infancy. Technological advancements in individual ADC components are now overcoming key historical limitations, thus allowing for further innovation in oncology. Additional variations of ADCs such as radiopharmaceuticals and immune-stimulating ADCs are being rapidly commercialized and clinically developed. Researchers are also exploring the potential of ADCs in areas outside of oncology such as autoimmune disease and infectious disease. Large pharma is investing heavily in the promise of ADCs through significant deal-making, most prominently including Pfizer's acquisition of Seagen, announced in 2023.

Overall, we believe technological improvements and significant investment by large pharma companies mean ADCs will increasingly disrupt and improve the treatment of cancer. This chapter gives a conceptual overview of ADCs and discusses their commercial potential, improvements in ADC technology, large pharma deal-making around ADCs, and other future areas of potential for ADCs.

Setting the Stage

Historically, cancer treatments can be divided into five main pillars: (1) surgery, (2) chemotherapy, (3) targeted therapy, (4) radiation, and (5) immunotherapy. In this chapter, we focus on chemotherapy, radiation, and immunotherapy.

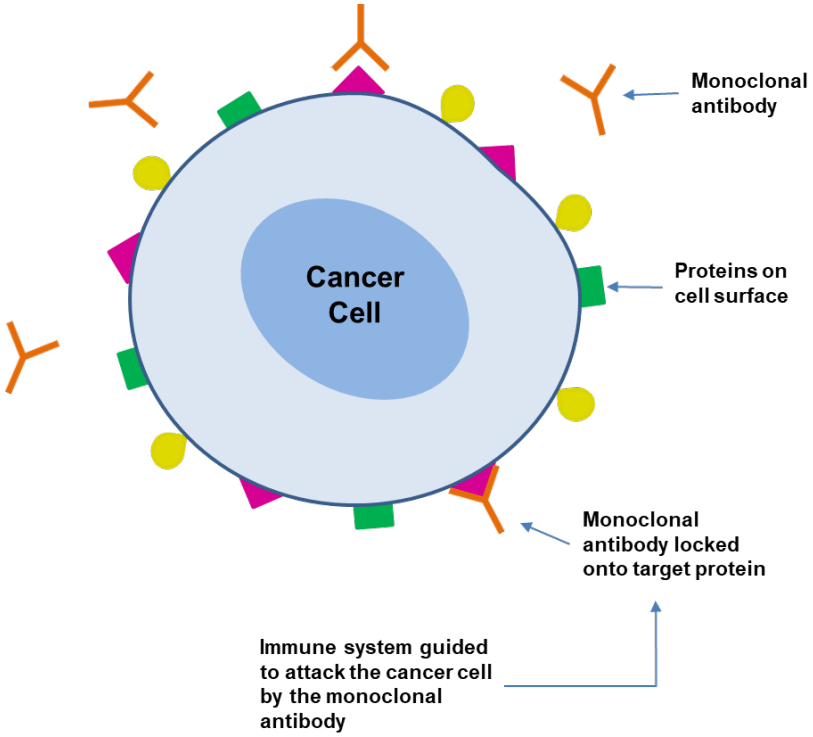
- **Chemotherapy** was introduced in the 1940s and offered a potent way to kill rapidly dividing (cancerous) cells. **The problem with chemo is that it is a very imprecise and toxic way to attack cancer, causing significant damage systemically to non-cancerous areas in the body, resulting in a very poor side-effect profile.**

- **Radiation** dates back even earlier, with Chicago physician Emile Grubbe claiming to have treated a breast cancer patient with x-rays in 1896. Over the next 100 or so years, the field of radiology grew more sophisticated, with the addition of radium and other isotopes to the radiation treatment arsenal and the introduction of more sophisticated delivery technologies such as external-beam therapy or brachytherapy. **The problem with radiation is that it, too, is an imprecise way to kill cancers, resulting in irradiation of surrounding non-cancerous tissues and frequently causing secondary cancers to arise.**

- **Immunotherapy**, the idea of hijacking the body's immune system to attack cancer cells, has seen the most significant cancer treatment innovation in recent years. One of the foundational advances in immunotherapy was the introduction of monoclonal antibodies to the cancer treatment paradigm. Monoclonal antibodies, which are Y-shaped proteins, were introduced starting in the 1980s and offer a much more precise way to target cancers, binding to cell surface markers that are primarily found only on cancer cells and not healthy tissue (Figure 6). Once bound, the monoclonal antibody can activate and guide the immune system straight to the cancer. **The problem with monoclonal**

antibodies is that they do not offer direct intrinsic cancer-killing capabilities. Instead, they require the antibody to hijack the immune system to attack the cancer or have anti-cancer effects via other secondary mechanisms.

Figure 6. Monoclonal Antibodies Can Precisely Target Cancer Cell Surface Proteins



Source: Citi GPS

Enter the Antibody Drug Conjugates

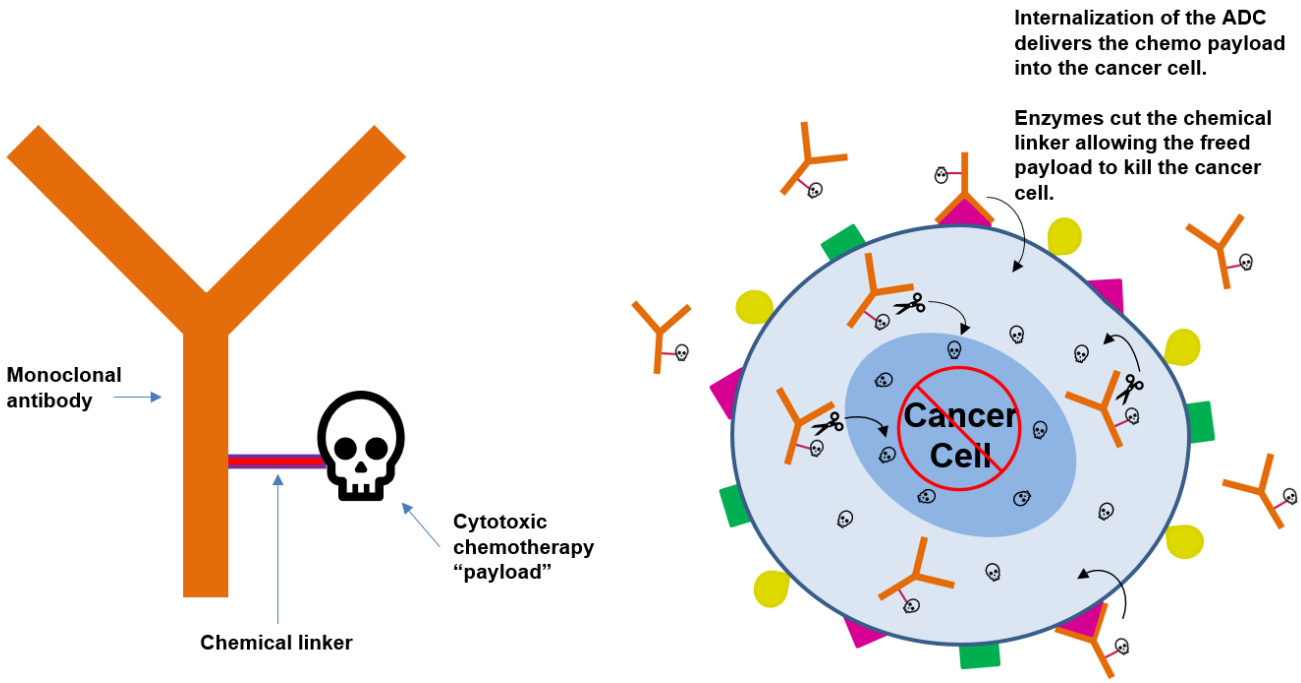
An ADC is a monoclonal antibody with a very small amount of highly potent chemotherapy chemically attached to it via a "linker." The benefit of this design is that an ADC not only has the potency and cancer-killing capabilities of chemo, but also the precise targeting capabilities of a monoclonal antibody. Theoretically, an ADC approach could be much safer than traditional chemo given the targeted destruction of cancer cells.

ADCs are designed deliberately and are generally composed of **three parts** (see Figure 7):

- **A monoclonal antibody** backbone targeting a chosen specific cancer protein marker. Once locked onto the target protein, the ADC can enter the cancer cell via a process called "internalization."
- **A cytotoxic payload** (aka chemo) carried by the monoclonal antibody into the cancer cell. These payloads, while delivered in small amounts, are 100-1,000x more potent than standard chemo options.
- **A chemical linker** attaching the cytotoxic payload to the monoclonal antibody. Once inside the cancer cell, the chemical linker is clipped by enzymes, allowing the freed payload to kill the cancer cell. The process through which the linker is

attached to the monoclonal antibody is called “conjugation”; hence, the name antibody drug conjugate.

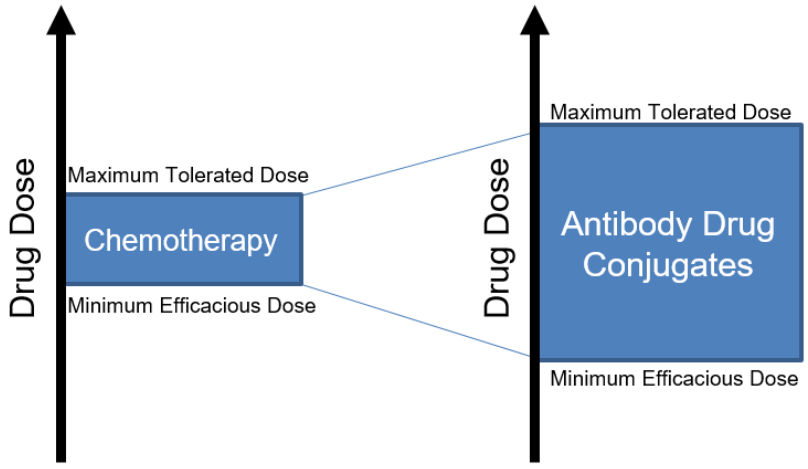
Figure 7. ADCs Carry Chemo Payloads into Cancer Cells, Resulting in Precise Tumor-Killing



Source: Citi GPS

In theory, because they specifically target cancer cells and not healthy tissues, ADCs can be far safer than chemotherapy and offer a wider therapeutic window (ie, range of efficacious doses) (Figure 8).

Figure 8. ADCs Can Increase the Therapeutic Window of Chemotherapy



Source: Citi GPS

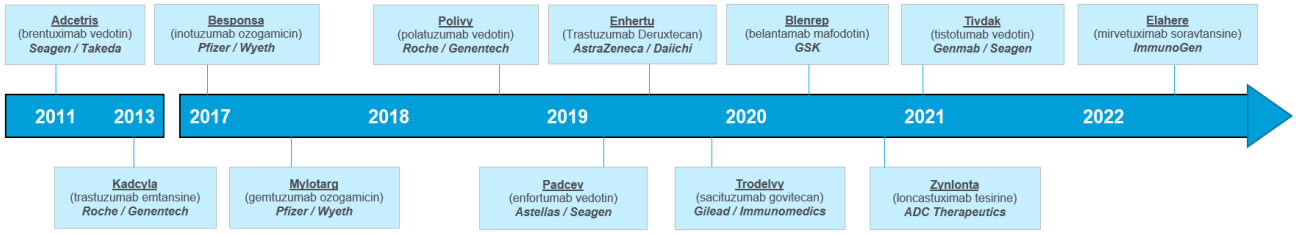
Technology Improvements Are Now Allowing ADCs to Reach Their Full Potential

Since 2011, 11 ADCs have been approved and are currently available on the market for use in a wide variety of solid and hematologic cancers (excluding Blenrep, which was withdrawn in November 2022).

Many approved and previously discontinued ADCs faced significant safety challenges in patients, mainly because of limitations in protein, linker, conjugation, and chemo payload engineering. For example, a common problem and past reason for poor ADC efficacy and safety was an unstable linker. One can imagine that if the linker disintegrates too soon in the body, the payload would be unintentionally released prior to the ADC reaching the cancer, thus causing chemo-like systemic toxicities and negating the targeted benefits of the monoclonal antibody.

That said, protein, linker, conjugation, and chemo payload technologies have all improved dramatically in recent years, allowing for: (1) a more stable ADC product; and (2) increased efficacy and safety. **These improvements have advanced the promise of ADCs and have accelerated research and investment interest from both biotech and pharma.** Note that eight of the 11 available ADCs were approved in the past five years alone, highlighting this acceleration (Figure 9).

Figure 9. Timeline of ADC FDA Approvals for Cancers

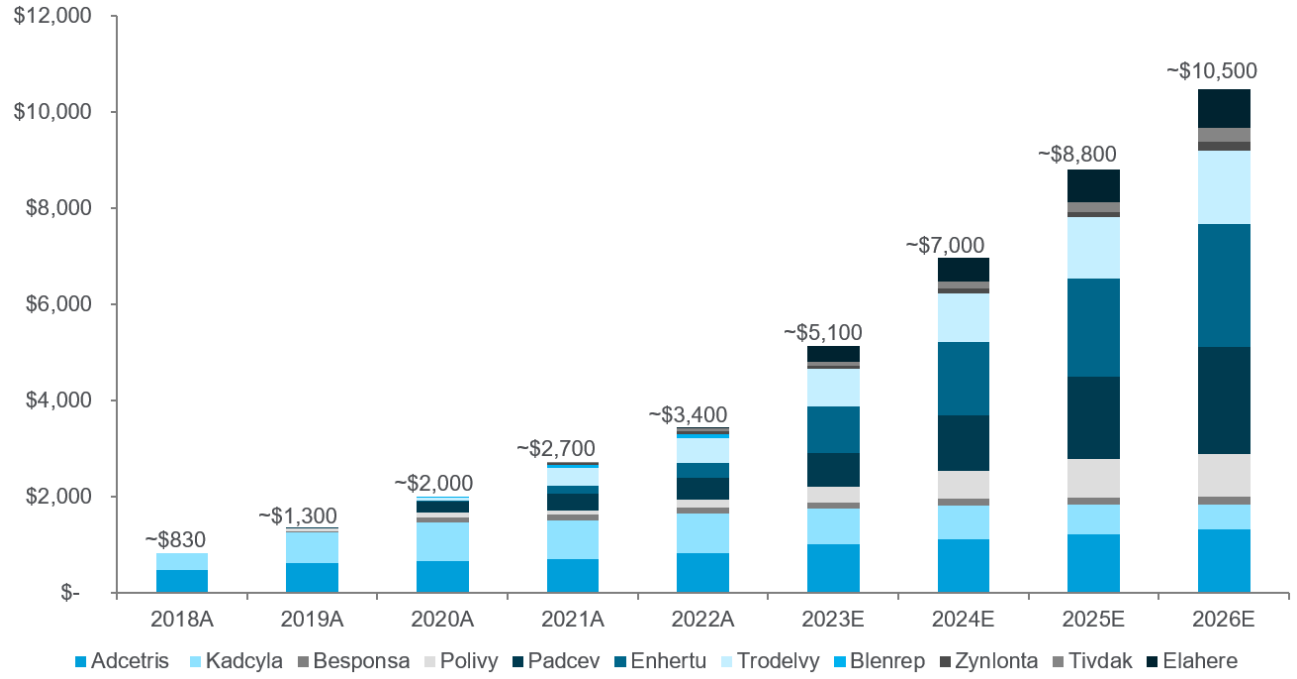


Source: Citi GPS

The 11 approved ADCs have been proven to be commercially viable, with sales expected to grow to over $10 billion by 2026 in the U.S. alone, doubling from 2023 projections (Figure 10). **One of the key commercial advantages of ADCs is their design complexity. The inclusion of multiple components (each with its own intellectual property) means they may be difficult for biosimilar (generic) competitors to replicate.** Because of this, ADCs could have long, uninterrupted commercial lives compared with other oncology agents, which are easier for generic competitors to replicate. Pfizer reported this dynamic [2]as a key reason for its acquisition of leading ADC company Seagen.

[2] Pfizer website, "Pfizer Invests $43 Billion to Battle Cancer | Pfizer", March 13,2023

Figure 10. Currently-Approved ADCs Are Expected to Generate Over $10 Billion in Sales in the U.S. ($mn, 2018-26E)



Source: Visible Alpha, Citi GPS

The ADC Field Continues to Innovate…and Big Pharma Is Taking Notice

Given the improving technology and increasing commercial success, the field of ADCs continues to innovate and broaden its reach. There are currently around 300 ADCs in preclinical and clinical development covering a wide range of cancer types. Substantial M&A activity around ADCs (and similar payload-bearing therapeutics) highlights interest from both big biotech and pharma.

Outside of M&A, dealmaking around ADCs has also been done through large collaboration and partnership agreements. There are too many examples of ADC partnerships and collaborations to list here, though the largest historically was signed in October 2023.

We believe the significant investment being made by large pharma into ADCs and similar modalities highlights the promise of the class and is driving further R&D spending. **Other variations of ADCs, which are also being rapidly commercialized and clinically developed, include:**

- **Radiopharmaceuticals:** Radiopharmaceuticals are conceptually very similar to ADCs. The key difference is the use of a radioactive isotope in place of a cytotoxic chemotherapy payload. Radiopharmaceuticals can be thought of as a more precise way to deliver radiation to cancers (see the “Radiation” discussion on the first page of this chapter), analogous to ADCs being a more precise way to deliver chemotherapy. Radioisotope payloads are orders of magnitude more potent than chemo payloads and importantly do not require “internalization” to kill the cancer cell, given that radiation from the isotope is lethal even if externally located to the cancer cell.

  While they are more potent, the radiation emission makes storage and use of radiopharmaceuticals logistically more challenging than ADCs.

Radiopharmaceuticals require the use of specialized radiation oncology and nuclear medicine facilities and thus are not easily utilized by community oncologists. Unlike cytotoxic chemo payloads, which are stable in storage, radioisotopes face the challenge of very short half-lives, limiting the shelf life of radiopharmaceuticals to as little as a few days post-manufacturing.

Nonetheless, radiopharmaceuticals have been proven commercial successes in the treatment of cancer (eg, Xofigo, Lutathera, and Pluvicto). Dealmaking in the radiopharmaceutical field is also very active.

- **Immune-stimulating ADCs (ISACs):** ISACs are a variation of ADCs that utilize an immune-stimulating payload instead of a cytotoxic chemo payload. Cancers generally utilize multiple mechanisms to escape recognition by the body's immune system. The idea with ISACs is to stimulate immune signaling in cancer cells, thereby potentially allowing the immune system to recognize and destroy the cancer cell. ISACs have yet to show meaningful clinical efficacy in late-stage studies, though the industry continues to invest heavily in R&D.

- **ADCs for use outside of cancer:** Most ADC development is focused on oncology. However, there is growing interest in utilizing ADCs for the targeted killing of pathogenic cells in other disease areas. For example, in autoimmune diseases, ADCs could be used to destroy out-of-control immune cells driving the disease. Similarly, ADCs could be used to destroy infectious diseases. They could therefore contribute to solving the problem of antibiotic resistance, which is drawing rising concern. These areas are currently under investigation by researchers.

Autonomous Agents: AI Assistants on the Rise

Ron Josey
U.S. Internet Analyst
Citi Research

Wenyan Fei
Analytical Insights Team
Citi Global Insights

Digital assistants are already among us. Most of us think nothing of asking them to play some music, help us avoid traffic, or point us in the direction of the best meal in town. As the full impact of AI is realized, how much more help might we need – or be offered?

We think the answer is a lot, and we believe digital assistants and agents will play an increasingly central role in our lives across a range of use cases.

We are already beginning to see the adoption of AI agents and the benefits they offer to engagement, commerce, and customer service. OpenAI has seen adoption across consumers (c.100 million weekly active users), developers (over 2 million), and businesses (OpenAI has partnered with 92%+ of Fortune 500 companies).[3] Snap's My AI counted over 200 million users in October 2023, and these users have generated over 20 billion cumulative messages since the tool's launch in February 2023 (or around 2.5 billion messages per month), suggesting around 100 messages per user.[4]

Usage spanned across shopping (over 16 million conversations on clothing and apparel), travel (over 5 million conversations on top tourist destinations), and food and dining (over 5 million conversations on restaurant recommendations).[5] In many ways, this engagement is a form of search. Meta's LLaMA (large language model Meta AI)-based code generation models have been downloaded over 30 million times through the Hugging Face platform.[6]

Digital Assistants Upgraded to "AI Assistants" Through Large Language Models

We believe the next frontier area of AI innovation is autonomous agents.

Almost all U.S.-based developers (92%) are already using AI coding tools both in and outside of work, according to a June 2023 GitHub Survey.[7] Building upon their ability to understand instructions and turn them into lines of code that can be executed by computers, large language models (LLMs) can be fine-tuned to go beyond simply interacting with humans or producing code.

Instead of prompting the LLMs with a single line of instruction, more complicated goals can be inputted that may take multiple steps to achieve (eg, call a taxi, book a flight, reserve a table). We can "teach" LLMs how to complete such tasks with memories and loops – prompting them to break the task down into sub-tasks, come up with an execution plan, and follow that plan step by step.

---

[3] Jon Porter, "ChatGPT Continues to Be One of the Fastest-Growing Services Ever," The Verge, November 6. 2023.
[4] Andrew Hutchinson, "Snapchat Crosses 400M Users, Sees Revenue growth in Q3," Social Media Today, October 24, 2023.
[5] Riyah Shah, "New AI Features for Snapchat, Includes AI-Generated Snaps & News Tools," Social Nation, December 31, 2023.
[6] Jonathan Vanian, "Meta's Unique Approach to Delivering AI Puzzles Wall Street, But Techies Love It," CNBC, October 16, 2023.
[7] Inbal Shani and Github Staff, "Survey Reveals AI's Impact on the Developer Experience," Github Blog, June 13, 2023.

In this way, LLMs can be tuned to function as agents or digital assistants to humans for tasks.

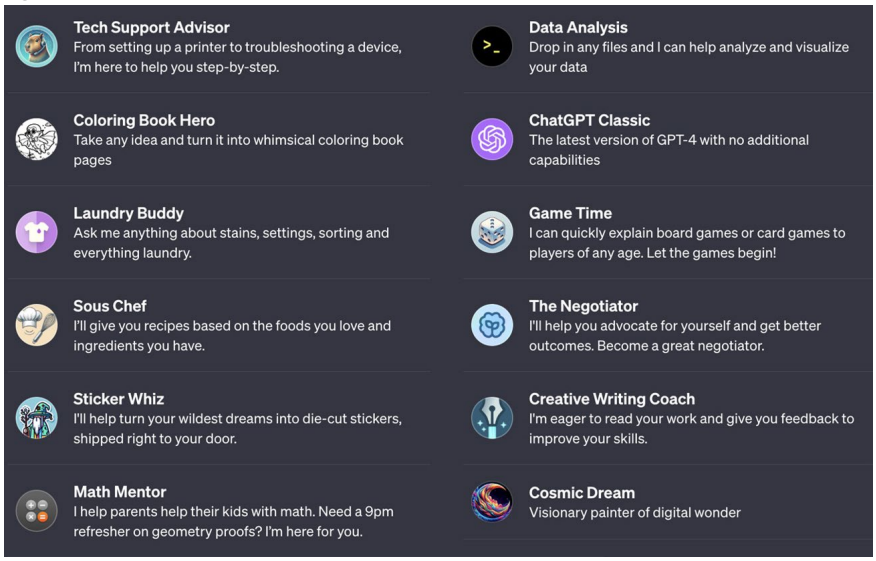
Toran Bruce Richards, the founder of the video game company Significant Gravitas, made this a reality first in late March 2023 when he launched a prototype application called Auto-GPT on GitHub. It quickly took GitHub by storm and was ranked 24th among over 370 million project repositories on the site as of January 9, 2024.[8]

Mustafa Suleyman, CEO of Inflection AI and co-founder of DeepMind, mentioned in a September 2023 interview with the *Financial Times* that he expects to see AI systems to add applications including "a very coherent expert role that can coordinate and decide and plan and reason and use its judgment" in the near future.[9] "In future, AI is going to be participating in the economy in a material way, unlike the way that Excel participates in the economy. It's going to be orchestrating actions using APIs [application programming interfaces]," he said. "It's going to be booking and buying and planning and organizing."

AI Assistants for Consumers

GPT-4 has already demonstrated its ability to design and execute chemistry experiments, browse the web, and use software tools including other AI models. In November 2023, OpenAI offered more flexibility to users by launching new GPTs at its first ever Developer Day. GPTs are customizable AI assistants that users can create on their own with specific instructions, domain knowledge, and any combination of skills for specific purposes. Users can now fully leverage the power of the latest LLMs and move beyond having conversations to executing calls to many external tools, including various apps and web browsers, to complete tasks. Within 48 hours of launch, over 10,000 GPTs were created by users, with some examples noted in Figure 11.

Figure 11. Examples of GPTs



Source: Citi GPS, OpenAI

[8] GitHub, "GitHub Ranking: Top 100 Stars," accessed January 9, 2024.
[9] Richard Waters, "U.S. Should Use Chip Leadership to Enforce AI Standards, Says Mustafa Suleyman," *Financial Times*, September 1, 2023.

Meta launched its own AI assistant, Meta AI, and 28 verticalized AI personalities (based on celebrities) that are intended to drive incremental engagement across Meta's family of apps (including WhatsApp, Messenger, Instagram), while its Meta-powered Ray-Ban smart glasses and Quest 3 mixed reality headset have incorporated Meta AI into the devices.

The AI assistant also leverages Meta's search partnership with Bing that enables the LLM to get real-time info updates and includes an image generation tool. Potential use cases include cooking, travel, and sports.

Google launched its AI assistant, Bard, in February 2023 and has evolved its search function with the introduction of Search Generative Experience (SGE) to create a more interactive, conversational internet experience. These offerings represent the early evolution of what could be the future state for Google Search, with Bard in particular representing an incremental surface for user engagement and unlocking new query types (ie, follow-up conversations).

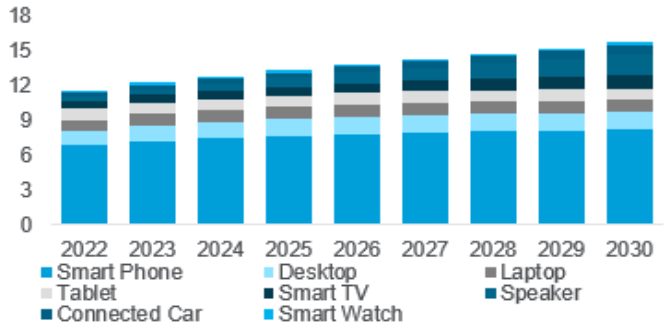
Future use cases could be even broader than anticipated – intelligent digital assistants may be able to perform more complex physical tasks by linking to intelligent robots, or embodied AI assistants. We also expect AI to bring a massive upgrade to current digital assistant models like Apple's Siri, Amazon's Alexa, and Google Assistant. This will in turn lead to faster adoption of digital assistants.

Total Addressable Market (TAM) Analysis for Consumer Market

To assess the opportunity in this area, we break down the consumer digital assistant device market into eight different segments: (1) smartphones, (2) laptops, (3) desktops, (4) smart speakers (eg, Alexa), (5) tablets, (6) smartwatches, (7) connected cars, and (8) smart TVs. We then model two different scenarios: (1) the base case and (2) the upside case.[10]

In our base case scenario, we assume the total number of devices with digital assistant capabilities will grow from 12 billion to 16 billion units by 2030. We discuss revenue models and see a wide range of potential market sizes by 2030, from $57 billion to $339 billion, as shown in Figure 13. The key takeaway is the digital assistant market is likely to become a big new market. In turn, this will likely attract more capital, drive more innovation, add more value, increase penetration, and create a battle to become your preferred intelligent digital assistant.

Figure 12. Total Number of Digital Assistant Devices (by device, in bns)



Source: Citi Global Insights

Figure 13. 2030 Digital Assistant TAM Sensitivity Table (in $bns)

% Devices Subscribed to Premium Model	Premium Model Monthly Subscription Fee ($)						
	3.0	3.5	4.0	4.5	5.0	5.5	6.0
10	56	66	75	85	94	104	113
15	85	99	113	127	141	155	169
20	113	132	151	169	188	207	226
25	141	165	188	212	235	259	282
30	169	198	226	254	283	311	339

Source: Citi Global Insights

[10] For more information, see Citi Global Insights, *AI Time: AI Assistants Are Coming for You*, August 1, 2023.

AI Assistants for Enterprises

Workplace Productivity

AI-powered digital assistants can make a big difference by helping with simple tasks such as emails – in 2021, the average U.S. employee spent around 2.5 hours per day checking emails. And that does not include the nonstop pings from Slack, Teams, Skype, or Google Chat.

Microsoft launched Copilot, its GPT-powered assistant, for Microsoft 365 applications and services in 2023. The Copilot assistant is in essence a form of Auto-GPT that is fine-tuned to complete designated tasks for Microsoft Office. The Outlook Copilot, for example, can help users triage their inbox, prioritize important and urgent emails, summarize long email threads, and draft replies.

An internal study of Copilot at Microsoft in November 2023 found 70% of users felt more productive, and overall, users were 29% faster in a series of tasks – 64% of users spent less time processing email, 85% of users said they could write first drafts faster, and 75% of users said Copilot saved time by finding content across files.[11] In a survey that includes over 2,000 developers on the GitHub Copilot system, the average time for a software developer to complete the same coding task is 71 minutes with Copilot versus 161 minutes without Copilot – a 55% boost in productivity. Further, 88% said they were more productive, 74% said they could focus on more satisfying work, and 77% noted it helps them spend less time searching for information or examples.[12]

Freeing users from repetitive tasks allows them to focus on more creative opportunities. As one senior software engineer said, "(With Copilot) I have to think less, and when I have to think it's the fun stuff. It sets off a little spark that makes coding more fun and more efficient."[13]

Customer Service and Engagement

AI-powered agents can improve customer engagement. For example, imagine a Taylor Swift-trained bot that directly interacts with her fans. Or, picture a brand creating an LLM-based customer service agent that could improve conversion rates and sales. As the use cases of agents evolve, we believe search and social usage, ad budgets, and customer service are likely to evolve as well.

Through the help of an AI assistant and more natural language queries, users can find new products that suit their preferences, purchase a product, and receive ongoing product support, all without the involvement of human representative.

As an example, a user could click on a future Facebook ad for a shirt or sweater; ask follow-up questions about the material, price, or delivery; and then purchase the product in the same conversation. This more seamless purchase experience (potentially from an ad in a user's feed) could improve conversion rates, which could lead to greater ad revenue and overall revenue across search, social, and the broader Internet.

[11] Microsoft, "What Can Copilot's Earliest Users Teach Us About Generative AI at Work?", November 15, 2023.
[12] Eirini Kalliamvakou, "Research: Quantifying GitHub Copilot's Impact on Developer Productivity and Happiness," GitHub Blog, September 7, 2022.
[13] Ibid.

AI assistants could also generate cost savings and efficiency across customer service operations by improving agent productivity and automating a greater volume of customer interactions.

Toward productivity, a McKinsey analysis suggests that generative AI applied to customer care functions could improve productivity by 30%-45%, which could prove conservative as the AI assistants evolve.[14] And with improving AI assistant capabilities, Gartner projects customer service and support agent organizations can operate at an approximately 20%-30% lower headcount by 2026.[15]

In the business-to-business (B2B) domain, OpenAI supports business applications by offering plugins (eg, ChatGPT can be leveraged for Expedia, Instacart, and Shopify) and is beta-testing an Assistants API to support custom AI assistant development.

Meta is also launching AI Studio, which will enable businesses and creators to build custom bots in a low-or no-code environment using Meta's Llama 2 LLM. We believe this initiative could drive agent adoption, increase customer engagement, and improve monetization for advertisers. CEO Mark Zuckerberg suggested it could be broadly available in 1H 2024, and the service could be a natural complement to Meta's Click-to-Message ad product. We can envision a world whereby celebrities and small and midsized businesses can train agents to chat with fans or customers.

Innovation is also taking place at Google, with Duet AI (a productivity AI assistant) and a Vertex AI platform built on Google's core generative AI capabilities based on PaLM 2 (and soon to be Gemini) LLM, allowing businesses to build (or leverage pre-built) AI models, assistants, APIs, and applications within Google Cloud. A generative AI upgrade of the Google Assistant is currently underway.

While enterprise adoption can take longer than consumer adoption, the enterprise market for AI assistants is likely to be significant. Based on the productivity opportunities it can deliver, Microsoft is charging $30 per month for its new Copilot services. This contrasts the $3-$6 per month average revenue per user (ARPU) we modeled in our consumer market scenarios. The conclusion, however, is similar – AI assistants are coming for you.

[14] Michael Chui et al., "The Economic Potential of Generaitve AI," McKinsey & Co, June 2023.
[15] Gartner, "Customer Service and Support Leaders Should Assess Generative AI Technology Options to Enhance Their Organization's Function," August 2023.

FemTech

Adam Spielman
Citi Global Insights

Amy Thompson
Social Economist
Citi Global Insights

Canberk Benning
Citi Global Insights

What Is FemTech?

The word "FemTech" was coined in 2016 by Ida Tin, one of the cofounders of the period-tracking company **Clue**. She wanted to help "legitimize the female health tech market, thus driving forward innovation, attracting investment, and helping to normalize conversations about female health."[16]

The sector ranges from products that might be thought of as "consumer" – for example, simple period-tracking apps – to ones that use deep science, for example, prescription drug development. However, the space is a continuum:

- **Period-tracking shades into contraception and fertility – and these areas can be highly medicalized.** It is possible to argue that the period-tracking app on the Apple Watch is purely a consumer product, but often period tracking has a quasi-medical function. Between a third and a half of users use the technology purposefully either to try to become pregnant, or to avoid doing so.
- **Female-friendly medical devices shade into consumer products.** One example comes from the Matrix Female Health & Care. This is a pre-seed start-up that is looking to redesign the speculum – a device that's been around since Roman times – and replace it with something that is much more pleasant for women patients and, as a result, ends up looking as much like a consumer product as a medical one (Figure 14).
- **BioPharma, MedTech, and health service companies focusing on women's health.** One leading biopharma company focuses on therapeutics – mainly for women – in areas including reproductive health, heart disease, oncology, immunology, dermatology, allergy, and asthma. **CooperSurgical** is a division of a company that develops medical devices and services mainly used for fertility and OB/GYN surgical procedures, including reproductive genetics and ART — or assisted reproductive technology.

Figure 15 focuses just on startups, showing the split by vertical. The dark blue bars represent issues having to do with maternity and what is sometimes called "bikini medicine"; the royal blue shows more general issues related to women's health.[17] Not only does the chart show that the biggest focus of investment is in maternal and menstrual health and fertility, but it also shows that that general women's medicine remains much less of a priority than bikini medicine.

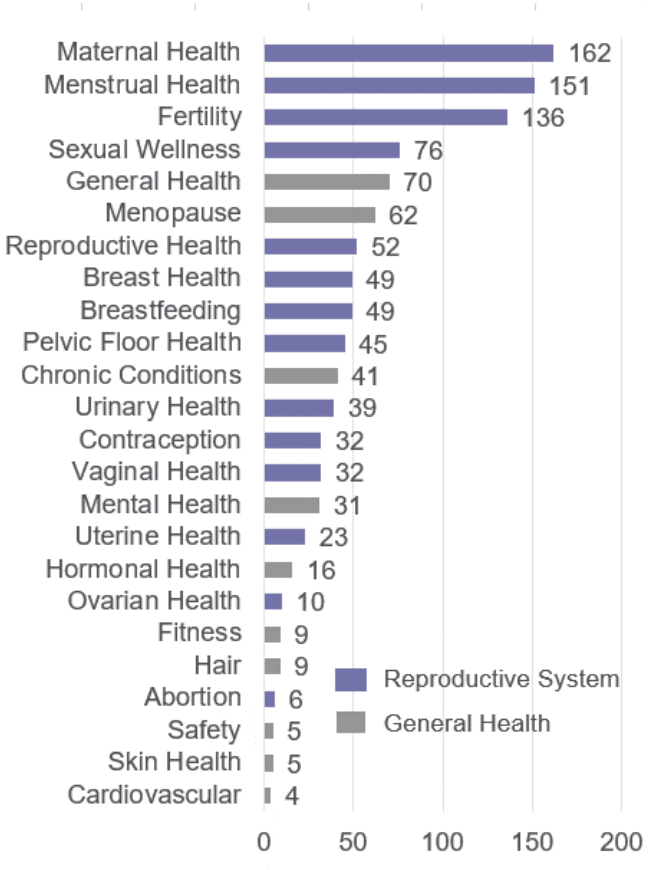
[16] Femtech.Live, "FemTech Founder: An Interview with Clue CEO, Ida Tin," February 11, 2021.
[17] The term "bikini medicine" refers to the mistaken belief that women's health only differs from men's in the parts of the body that a bikini would cover. See: George Institute for Global Health, "Time to Shift Research Focus From 'Bikini Medicine' to What Is Really Ailing Women," March 30, 2022.

Figure 14. Proposed Replacement for Speculums

- Speculums are devices to open up orifices -- for example vaginas -- for inspection. Archeologists have found Roman which are similar to modern versions.
- This replacement can be used either by the women concerned, potentially under instruction by a health care professional in another room, or directly by a health care professional.
- The tip includes a light, camera and device that can retrieve a sample.
- The purple sheath is designed to be sterilized; the blue part is wiped clean.

Source: Matrix Health and Care

Figure 15. FemTech Startups by Subsector (Number of Companies)

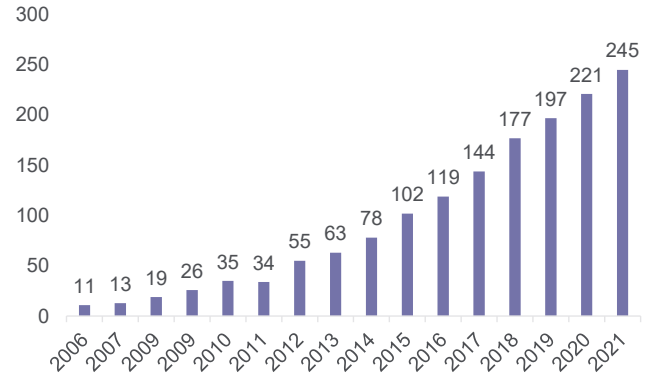


Source: Citi GPS, FemHealth Insights

Growing, But Not Huge

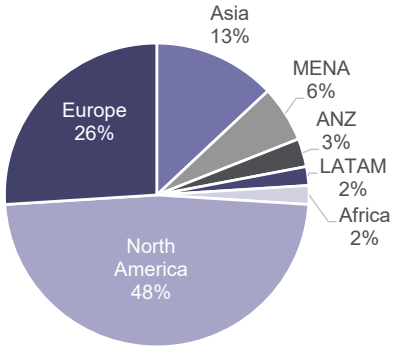
Figure 16 shows that although the number of FemTech startups has risen significantly over time, the total number is relatively modest. So far, about half are in North America and a quarter in Europe.

Figure 16. Number of FemTech Companies Founded



Source: FemTech Analytics, Citi Global Insights

Figure 17. FemTech Companies by Region



Source: Pitchbook, Citi Global Insights

Why Is FemTech Needed?

There is plenty of evidence that women are less well-served in many domains by the healthcare ecosystem – from individual healthcare providers through to ongoing research and development – than men. We see three distinct issues: worse diagnosis, worse therapeutics, and worse pain management.

- **Worse diagnosis:** Women are more likely to be misdiagnosed or diagnosed late. For example, women are more likely to receive an incorrect diagnosis for serious cardiovascular events like heart attacks and strokes. [18]
- **Worse therapeutics:** There are insufficient safe and effective therapeutics for women. Worldwide, women account for around 60% of adverse drug reactions, with modest variation globally.[19] There is a particular problem for pregnant women because very few drugs are tested for safety and efficacy on pregnant women.[20] In fact, only 4% of healthcare research dollars go to women's health issues – which explains why relatively common issues like endometriosis are under-researched.
- **Worse pain management:** First, it appears that women suffer more pain. Between 45% and 95% of women who menstruate experiencing dysmenorrhea,

[18] British Heart Foundation, *Bias and Biology: How the Gender Gap in Heart Disease Is Costing Women's Lives*, September 2019; David E. Newman-Toker et al., "Missed Diagnosis of Stroke in the Emergency Department: A Cross-Sectional Analysis of a Large Population-Based Sample," *Diagnosis (Berlin, Germany)*, Vol. 1, No. 2, April 3, 2014.

[19] We must note that adverse drug reactions in men were more likely to be serious or fatal, and there is a gap in understanding the difference between men and women in that direction as much as the one being highlighted in this report.

[20] Martina Ayad, MD and Maged M. Constantine, MD, "Epidemiology of Medications Use in Pregnancy," *Seminars in Perinatology*, Vol 39, No. 7, September 8, 2015.

which is painful menstruation that cannot be explained by any named disease.[21] In addition, half of chronic pain conditions are more common in women.[22] Furthermore, women are less likely to receive pain relief and, when they do, they have to wait longer for it. One meta-analysis of numerous studies on gender differences in the treatment of pain concluded that "women receive less treatment for pain than men."[23]

FemTech will not totally close the gender health gap, but it can help in two ways:

- **FemTech can support women to take full advantage of existing healthcare systems** by making them more empowered consumers of healthcare; and
- The rise of **FemTech incrementally shifts healthcare innovation toward a more equal gender balance** by focusing on women's health.

Clue, the period-tracking app, is a good example of how FemTech uses digital technology to help women understand their menstrual cycles and – if needed – communicate about them in detail.

- Women input data on their menstrual cycles, and the app generates insights, including predictions on periods and ovulation.
- It provides insights on cycle lengths to observe patterns and discover how cycles can impact other things, like mood, skin, and energy levels.
- Clue is free for users, but Clue Plus, which costs €1/month, provides deeper analysis and more advanced features, including fertility information, pregnancy tracking, and indications of perimenopause.
- The app has 10 million active monthly users. It has downloads in 190 countries and supports over 20 languages.

Unlocking the Value of FemTech Requires the Health Ecosystem to Evolve

Certain structural factors cause companies in many of the subsectors within FemTech to generate low (or even negative) margins. The emerging vertical could have a greater impact if the wider healthcare ecosystem evolved to better support women's health.

First, **health systems do not provide reimbursement for many women's health issues.** Many female-affecting health conditions apply to either effectively all women (like menopause) or a very large percentage (like painful periods). Throughout history, women have simply had to deal with these conditions as best they could. Perhaps as a result, modern healthcare systems generally do not pay for women to access therapeutics for these issues, and products that help address them are rarely reimbursable.

Period-tracking apps are an example of the profitability problem. Many of the largest wearables – including the Apple Watch and Fitbit – offer period tracking

[21] Stella Iacovides, Ingrid Avidon, and Fiona C. Baker, "What We Know About Primary Dysmenorrhea Today: A Critical Review," *Human Reproduction Update*, Vol. 21, No. 6, September 7, 2015.
[22] Natalie R. Osborne and Karen D. Davis, "Sex and Gender Differences in Pain," *International Review of Neurobiology*, Vol. 164, No. 277-307, June 13, 2022.
[23] Diane E. Hoffmann and Anita J. Tarzian, "The Girl who Cried Pain: A Bias Against Women in the Treatment of Pain," *The Journal of Law, Medicine & Ethics*, Vol. 29, No. 1, March 1, 2001.

effectively for free. Of course, companies like Clue offer tools that the wearable companies do not, but it is hard to compete profitably. In theory, period trackers could generate revenue by selling data. But in practice, all the main companies have given absolute commitments that they will never do this. It is hard to imagine many women uploading information about their menstruation if they thought this would be sold on.

Women-friendly medical devices also need to demonstrate value. The medical device segment, often called MedTech, is certainly a profitable industry, with many of the largest companies generating adjusted earnings before interest, taxes, depreciation, and amortization (EBITDA) margins in the high-20s. Typically, the margins for any product depend on how important they are in a clinical sense and the competition in their subsector. A device that enables a unique life-saving procedure is likely to generate much higher margins than (say) a (medical) thermometer that's designed for use at home. The question for companies that make female-friendly MedTech products is: to what extent are their customers – who are usually healthcare providers, not patients – willing to pay a premium.

Joint All-Domain Command and Control

Jason Gursky
U.S. Aerospace and Defense Analyst
Citi Research

U.S. defense strategy is focused on deterring conflict with enemies, particularly near peers. To accomplish this, the country invests over $300 billion annually in the modernization of its military, including its nuclear arsenal and its more tactical and conventional systems.

Joint All-Domain Command and Control (JADC2) is a military concept created by the Department of Defense (DoD) that aims to improve decision-making timelines through the use of enhanced communications, cloud-based interactions, and artificial intelligence (AI) and utilizing military assets across all domains – including space, air, land, sea, and cyber. The goal is to give the U.S. an advantage on the battlefield, and perhaps to prevent conflict from starting in the first place, by demonstrating that the country can both detect and eliminate enemy targets before even being seen by adversaries.

Imagine the following: The President of the United States is awakened early one morning to address an emerging situation. Just minutes prior, a newly formed enemy group assembled a small force that is a potential threat to U.S. special forces units operating overseas. With his top advisors and generals at his side, he looks at a bright digital screen displaying the exact location of the enemy on a satellite map produced by the Air Force, with live video footage from an Army drone depicting the menacing assets. The screen displays an array of available Army, Navy, and Air Force weapons and solutions in the area to efficiently deal with the enemy and protect the soldiers. Within a matter of minutes, the President chooses a plan and executes it, and the threat is promptly neutralized before the enemy forces are even aware of the presence of U.S. special operations units in the vicinity.

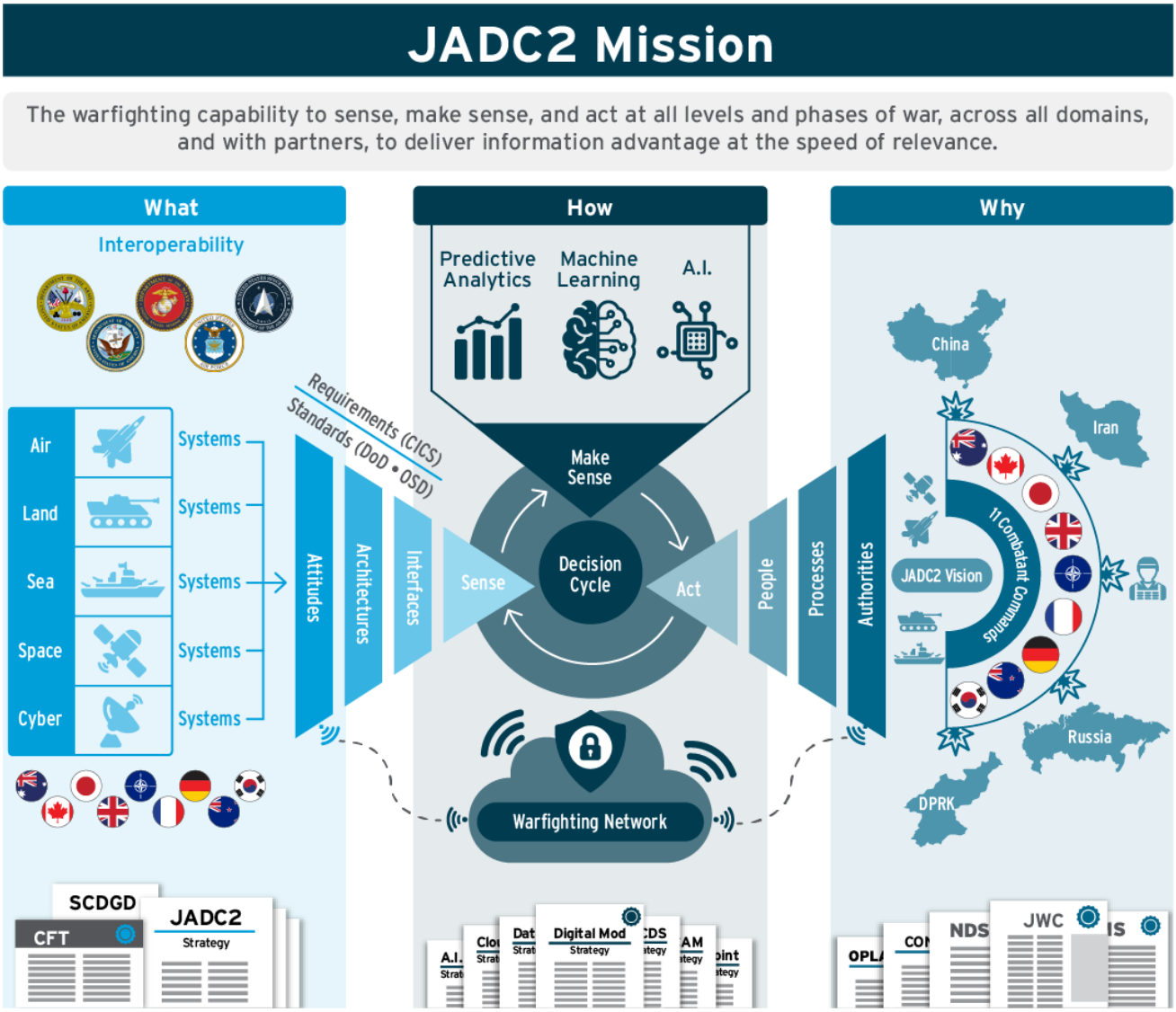
While we frequently see such scenarios in movies, they are largely a fantasy today as timelines using existing technology would take hours, if not days, to provide the President and his military leaders with the intelligence and situational awareness needed to execute such a mission. JADC2 is trying to solve for this using a collection of new and existing sensor, communications, and AI capabilities to materially reduce critical decision-making timelines to deal with threats. Its three phases (outlined below) seek to provide a more stable and efficient set of capabilities and procedures across the military to address existing and emerging threats:

- **The "Sense" phase** involves threat detection using space, air, land, cyber, and sea-based sensors.
- **The "Make Sense" phase** involves analysis of the threat using AI.
- **The "Act" phase** involves the neutralization of targets.

Investments in these areas are likely to be disruptive to the defense industrial base, as well as to military leaders as they attempt to adapt to the emerging technologies and operational constructs that will be needed to successfully deploy a military of the future – one that relies on cutting-edge sensing and AI capabilities to drive ever-faster decision-making timelines. While legacy defense contractors are benefiting from these investments given their exposure to existing systems, a cadre of new companies focus on providing ever-more capable solutions at a faster pace, in some cases using non-traditional modes of delivery (ie, delivering a service versus hardware that is then operated by a soldier). We expect many of these newer companies to be successful in carving out niches in critical areas of spend over time

given their technical know-how, lower costs (ie, pension and improved manufacturing techniques), and corporate structures that afford quicker development timelines.

Figure 18. JADC2: The Big Picture



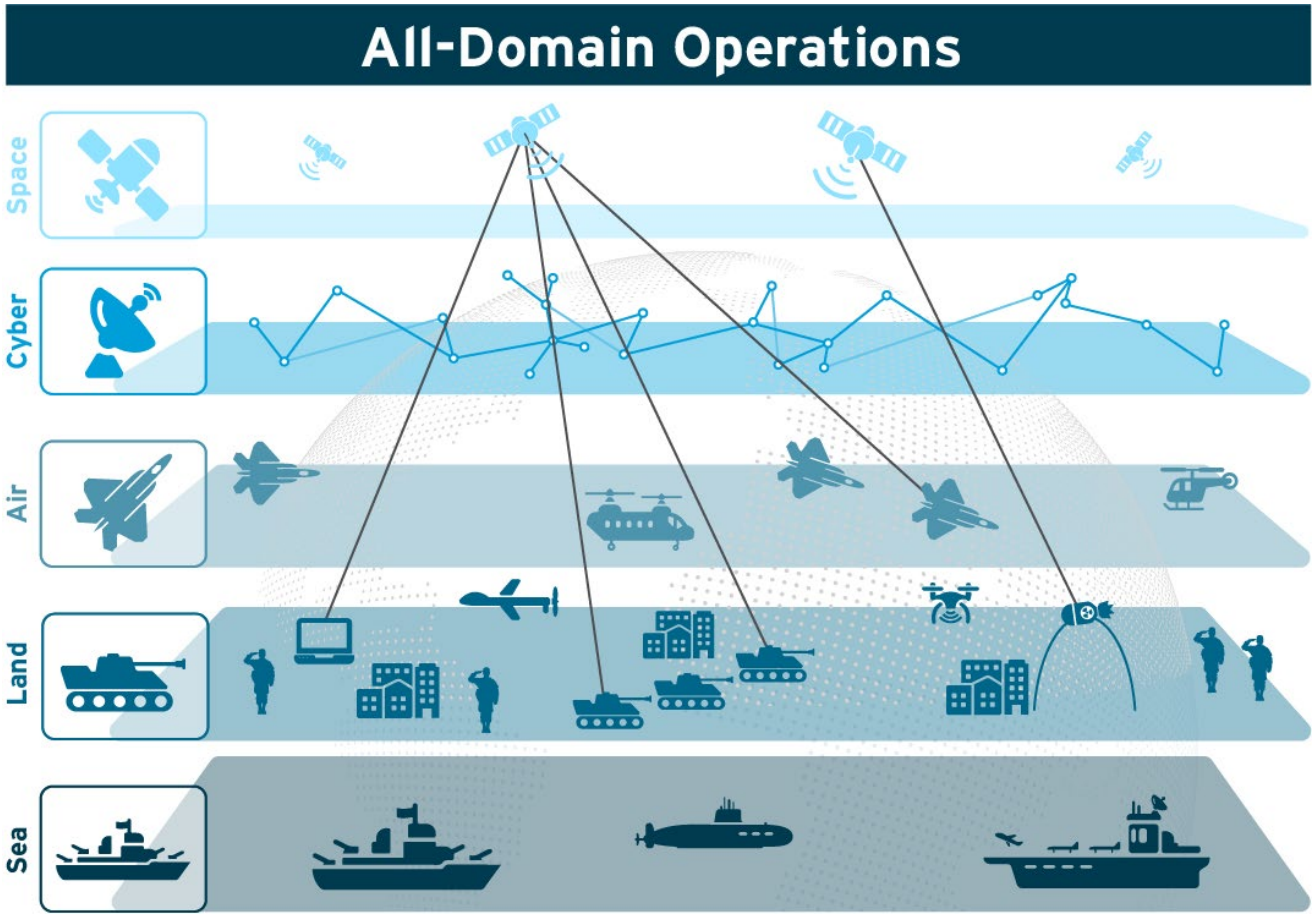
Source: Citi Research

Sense

The initial "Sense" stage involves collecting data from a variety of sources and receptors. In order to gather as much data as possible, the JADC2 initiative aims to install these sensors into a variety of devices across all domains, such as the Air Force's F-35 aircraft, the Navy's upcoming Large Unmanned Surface and Undersea Vehicles, and even satellites deployed by the Space Force. The goal is essentially to create a fabric where the combined sensors form an information web, becoming able to transfer relevant and important information to each other with ease. One major hurdle that this stage seeks to address is improving intra-military communications, as each military branch possesses a unique communications system that cannot easily interact with the others. Dubbed Project Convergence, a goal has been set to unify these systems between branches. As of January 2023,

the project has included the Army, Navy, Marine Corps, and Air Force. In fact, the Air Force has already made great strides in developing their own communications gateway, called the Advanced Battle Management System (ABMS), which will be integrated into JADC2.

Figure 19. JADC2: Sensing

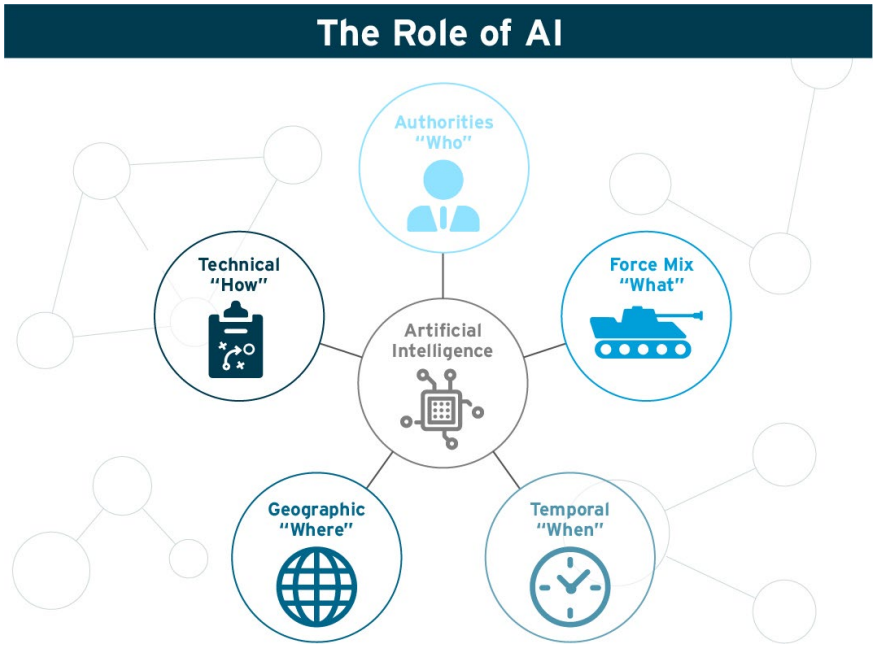


Source: Citi Research

Make Sense

The second stage, "Make Sense," focuses on the collection and interpretation of the multi-sourced data, primarily using AI. While the specific measurements and analysis conducted by the AI unit will vary, it is not difficult to understand how AI entities operate given our already normalized predisposition toward them, such as Apple's Siri, Amazon's Alexa, and now OpenAI's ChatGPT. The goal of using AI is to accelerate a commander's decision-making time. This can be done by assessing the threat levels of unannounced attacks, determining the location and defenses of enemy leaders and/or hostages, and even providing potential solutions and accounting for their potentially positive or negative ramifications. Furthermore, this analysis will not be verbally communicated, so there is less risk of data interception. Overall, this process, which would ordinarily take days or weeks, can now in theory be reduced to a matter of minutes, allowing the military to act and neutralize threats before they grow out of control. Moreover, global conflicts can potentially be avoided, as the U.S. government has expressed interest in promoting JADC2 to deter enemies from endangering others given the now much-quicker potential U.S. response time.

Figure 20. JADC2: Making Sense

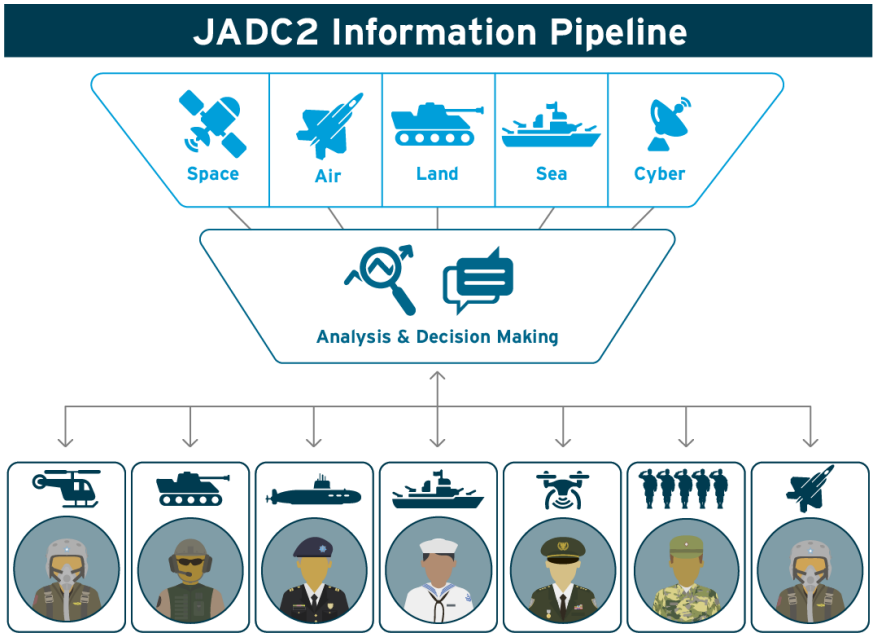


Source: Citi Research

Act

The final stage is "Act," which involves carrying out the course of action the commander chooses with the best available assets, no matter the branch of the military. The necessary orders will be sent out via the now-unified communication gateways to the responsible branches and carried out as planned. The length of time between receiving these instructions and executing them fully will depend on the mission, but transmission of the orders will be faster and more secure than by conventional means. Another goal is preserving the already existing protocols for carrying out orders. The current "Mission Command approach" allows subordinates to operate efficiently by being aware of their supervisor's intent for each course of action, even when lines of communication are broken. This stage completes the cycle and brings any confrontation that arises around the globe to a quick conclusion.

Figure 21. JADC2: Acting



Source: Citi GPS, Citi Research

Barriers to Adoption

In our view, there are three barriers to the immediate and sustainable adoption of a JADC2 approach to warfighting: inertial, technical, and ethical. Budget constraints are real in the U.S., particularly with a divided government. There is therefore risk that a "good enough" approach could be taken to the implementation of JADC2. For example, legacy sensors and communications networks might work against non-nation state actors and could save money, but they might prove vulnerable to the cyber capabilities of near peers. It will be up to Congress to assess the risks associated with that choice. There are also technical hurdles to be resolved, particularly in the areas of secure and resilient communications networks and the design and implementation of accurate AI models. Further, since AI is poised to play a significant role not only in gathering and inferencing data, but also in providing potential solutions, ethical concerns are likely to dominate the headlines in the years ahead. Are the data supplied to the AI machines accurate? Are the AI algorithms accurate? Should a machine be given the authority to provide advice/solutions in situations where human life is at stake?

Market Dynamics

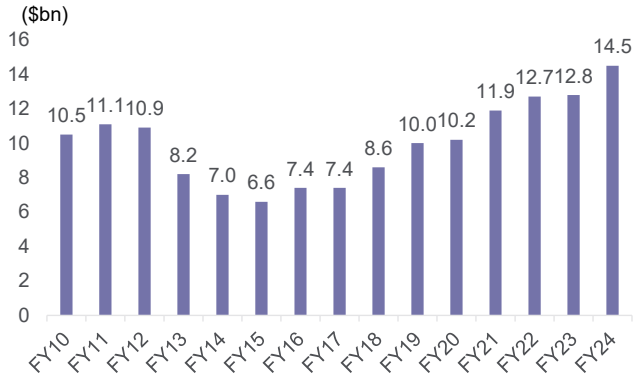
We expect investments in military modernization by the U.S. and its allies to grow around 5% annually through 2030 to support the strategic and defense doctrines of these countries – which emphasize deterring near-peer conflict. Spending is likely to focus on recapitalizing nuclear assets to assure future deterrence and on bolstering the capabilities needed to support the JADC2 construct. For example, budgets for space-based assets and operations have more than tripled in the U.S. over the past seven years, in large part driven by the need for a resilient network of sensors and communications capabilities to implement JADC2 effectively. Further, spending on intelligence-gathering and analytical systems (C4I, or command, control, communications, computers, and intelligence) has more than doubled over that same time. Importantly, it is not just traditional defense contractors that have benefited from market expansion. Hundreds of companies have been started over the past decade that focus on providing sensing, communications, and analytical

capabilities across all domains to support the warfighter and the JADC2 initiative, as barriers to entry are lower in these areas than in more traditional military systems. Many of these firms have reached critical scale, while others have become acquisition targets. Going forward, we expect the U.S. to continue to focus its spending growth in the areas in and around JADC2, given the benefits that a connected battle space with superior situational awareness affords the country's troops. This points to some level of market disruption as non-traditional defense companies take a greater share of the overall spending pie.

Figure 22. DoD Space-Related Budgets

Source: Citi Research, Department of Defense, OMB

Figure 23. DoD C4ISR-Related Budgets



C4ISR = Command, Control, Communications, Computers and Intelligence Surveillance and Reconnaissance

Source: Citi Research, Department of Defense

Neuromorphic Computing Architecture

Peter Lee
Asia Pac Semiconductors & Components Analyst
Citi Research

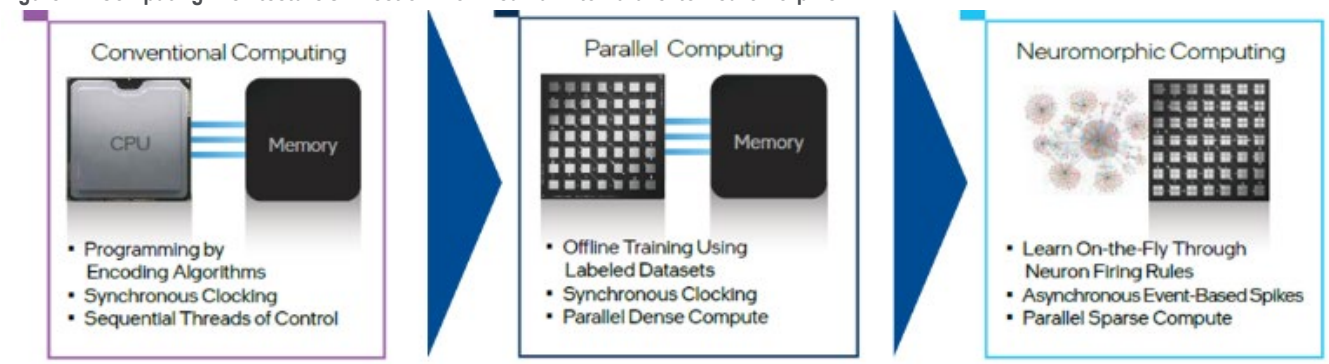
Until now, computing architecture has generally adopted one of two types: von Neumann or Harvard structure. The Von Neumann structure is the first computer method with a built-in program, and it can perform different calculations by merely replacing the software without having to rearrange wires in hardware.

The Harvard architecture divides the memory into a place where instructions are stored and a place where data is stored, which is a slight modification to the Von Neumann architecture to solve the bottleneck problem. However, given the continued increase in data bottlenecks between logic and memory, we expect the current architecture is becoming unsuitable to meet computing demand growth, especially in light of rising artificial intelligence (AI)-related demand structure.

Von Neumann Architecture → AI Computing Architecture

As such, we expect future computing architecture and hardware to initially develop in the direction of relieving the bottleneck issue with the current von Neumann structure. Eventually, we think it will take on a more human brain-like architecture that is optimized to meet rising AI-related computing demand.

Figure 24. Computing Architecture's Direction: Von Neumann to Parallel to Neuromorphic



Source: Citi Research

We expect the market change will occur in the order of: (1) integration of logic and memory; (2) neural-based parallel computing NPU (Neural Processing Unit); and (3) neuromorphic computing architecture.

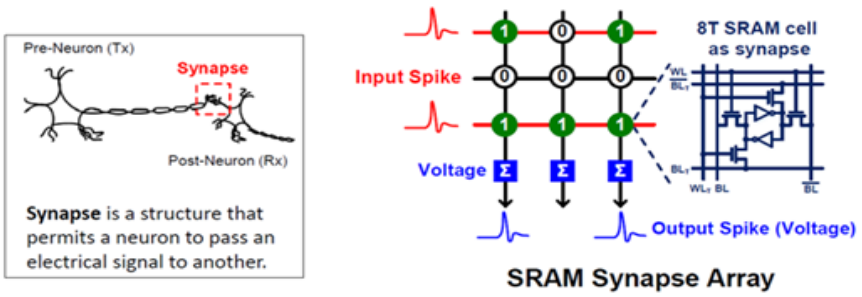
Neuromorphic computing architecture mimics the human brain

Neuromorphic computing, which mimics the human brain structure, follows a completely different paradigm from today's traditional von Neumann architecture task processing. While current computing systems with the von Neumann architecture transmit and modify precise numerical data, the human brain transmits and processes data through a series of stimulations called spikes.

Neurons receive these spikes through synapses and make small changes in the cell membrane. Neurons integrate these potential changes over time. When many spikes arrive in a short amount of time, the neuron then outputs a spike.

Neurons, which perform calculations in the human brain, correspond to logic chips in a computer, while synapses (the connections between neurons) would correspond to buffer memories such as static random-access memory (SRAM) and dynamic random-access memory (DRAM) in a computer. As such, if several billions of logic chips and several tens of trillions of SRAM/DRAM chips are connected in parallel, the theory is that the human brain structure can be artificially reproduced.

Figure 25. Mimicking Synapses with SRAM



Source: Google Search, Citi Research

Neuromorphic computing is expected to be adopted from 2026

We believe computing architecture will eventually resemble the human brain structure and expect the first step to be through increased demand for advanced packaging. The backend technology's importance should increase to speed up data transfer between heterogeneous chips.

In the memory space, we expect AI computing to generate increased demand for HBM3 (the latest High Bandwidth Memory) and DDR5 (the latest DRAM) memory. In terms of timeline, logic and memory integration and neural computing adoption have already begun, whereas neuromorphic computing is expected to be adopted starting in 2027E.

Neuromorphic computing is not well known to the market

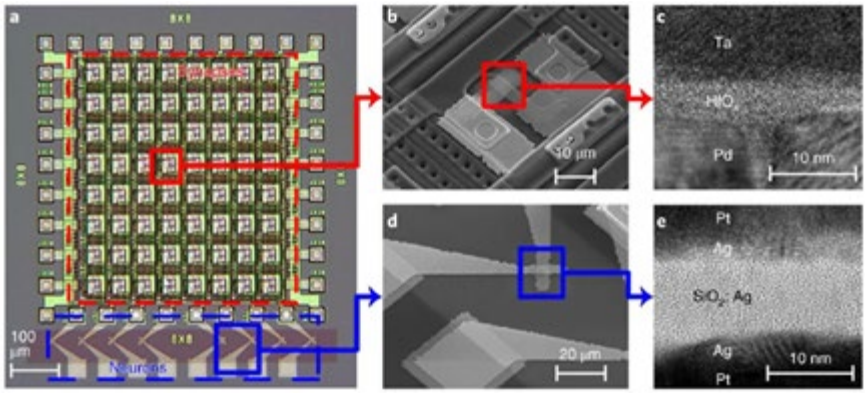
Neuromorphic computing is not well-known to the market due to technology difficulty. Most of the neuromorphic semiconductors developed to date have been developed with the conventional silicon-based complimentary metal-oxide semiconductor (CMOS) transistor technology.

If we can divide neuromorphic computing into "generations" depending on how it operates, the operation of first-generation neuromorphic semiconductors is based on reading and writing of weighted synapses between the CPU and memory chips. However, the second-generation neuromorphic semiconductors integrate logic and memory chips through a "memristor," which has characteristics of both memory and peripheral resistors at the same time.

"Memristor" is a compound word of memory and resistor, and in a memristor, having a gradual switching resistance is crucial to handling subdivided weights. Second-generation neuromorphic semiconductors vary depending on the material and implementation method of the device, but memristor is currently the most studied path.

One of the latest technologies currently being pursued is an AI accelerator, which is a hardware or microchip specially designed for fast processing of AI algorithms. Like other accelerators, AI accelerators are designed to efficiently perform specific AI-related tasks that are inefficient with a typical x86-based CPU. Purpose-built accelerators provide higher performance, more features, and higher power efficiency when completing a specific task. A technology key to this trend would be memory-based Processing in Memory (PIM), which corresponds to the concept of Near Memory.

Figure 26. Memristor with ReRAM



Source: Google Search, Citi Research

How big could the opportunity be?

As mentioned, we expect the first step towards this new architecture will be through increased demand for advanced packaging for the integration of logic chips. We foresee backend technology's importance will increase in the future to increase data transfer speed between heterogeneous chips.

In the memory space, we project Near Memory will emerge as an important direction and expect AI computing will generate increased demand for HBM3 and DDR5. In addition, we think integration of memory and logic chips through PIM will increase going forward. Theoretically, 1,000 to 10,000 synapse (DRAM) is needed per neuron (CPU/GPU), which indicates that the DRAM demand could increase radically with the adoption of neuromorphic computing.

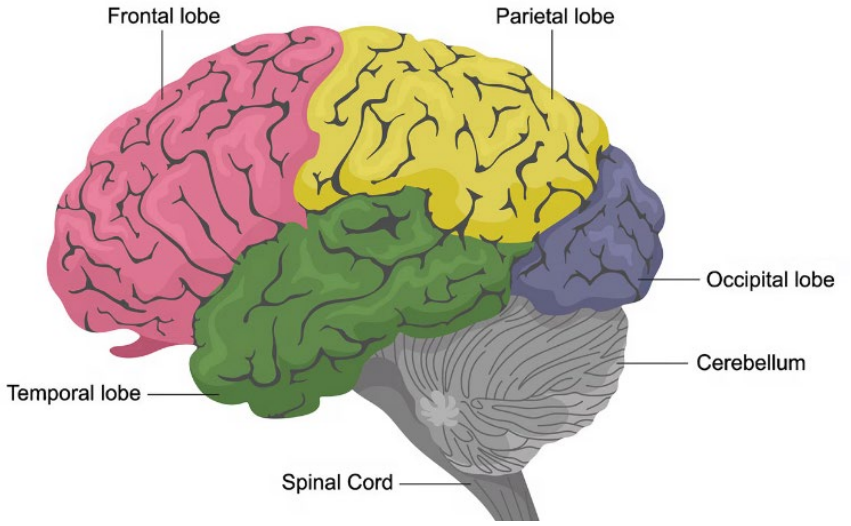
The barriers to adoption: difficult implementation due to complexity

It is difficult for neuromorphic computing to fully implement the human brain structure as it is currently difficult to implement artificial connections of billions of neurons and synapse like human brain. Therefore, while we expect neuromorphic computing architecture to follow the human brain structure, and to be implemented in a simplified brain structure in the early stage.

The human brain consists of 85 billion neurons and 85 trillion to 850 trillion synapses. In other words, the number of synapses per neuron is about 1,000 to 10,000 times the number of neurons. Human neurons correspond to the logic part of a computer, such as CPUs, and human synapses correspond to the memory part, such as DRAMs. Meanwhile, NAND in a semiconductor plays the same role as the hippocampus in a human brain.

In the human brain, a large portion of computational tasks are handled by the frontal lobe, where neurons and synapses are organically connected to process parallel computation, thereby enabling quick computation and transfer of data. However, in the current von Neumann computing structure, logic and DRAM are separated, resulting in a data bottleneck and thereby lowering performance compared to the human brain.

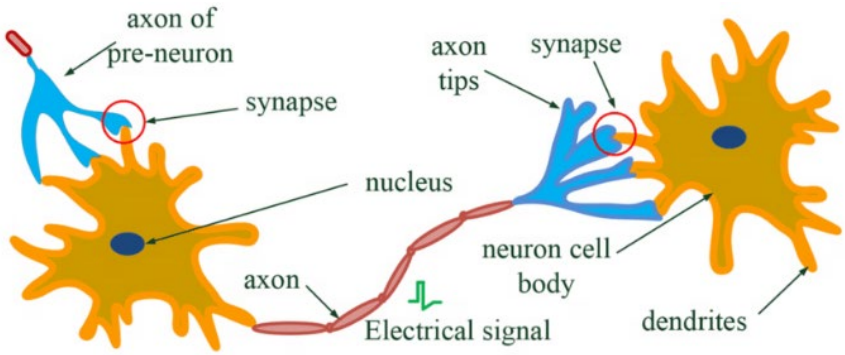
Figure 27. Human Brain



Source: Google Search, Citi Research

Neurons and synapses handle computations in the human brain. Calculations that are processed by neurons are temporarily stored in synapses, similar to how computation tasks processed by the CPU in a computer are temporarily stored in a buffer memory DRAM. However, as human brains can simultaneously handle computation and data transfer (unlike computers that are structured in the von Neumann architecture), human brains can complete learning and inference tasks more efficiently.

Figure 28. Brain: Neuron and Synapse



Source: Google Search, Citi Research

Notable features of the human brain structure include: (1) parallel computation to minimize data bottlenecks; and (2) real-time communication through neurons (which play the role of logic in a computer) and synapses (which play the role of DRAM in a computer). We predict computing architecture will develop in a way that increasingly resembles the human brain more to alleviate data bottleneck burdens.

Winners and losers

With the wider commercialization of artificial intelligence, we believe semiconductor chips will evolve towards the direction of merging logic and memory functions, similar to the way human brain functions.

Piezoelectric Roads: The Next Smart Road

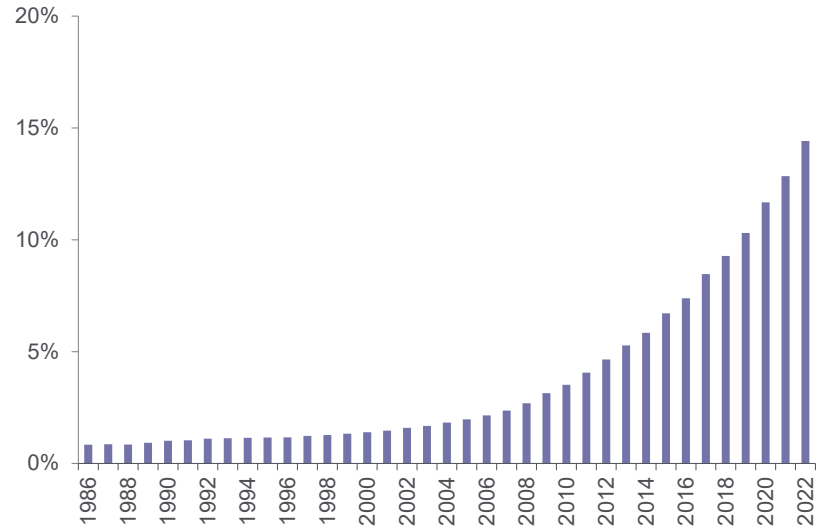
Avinash Mundhra
Business & Professional Services
Citi Research

Marc Van't Sant
Commercial & Professional Services
Citi Research

Arthur Truslove
Commercial & Professional Services
Citi Research

Rising energy demand is fast outpacing the rate at which we can renew conventional forms of energy, both for practical and emissions-related reasons. As alternative forms of energy generation become increasingly vital, piezoelectric roads could be a next big innovation.

Figure 29. World Renewable Electricity Generation as a % of Total



Source: Citi Research, energyinst.org

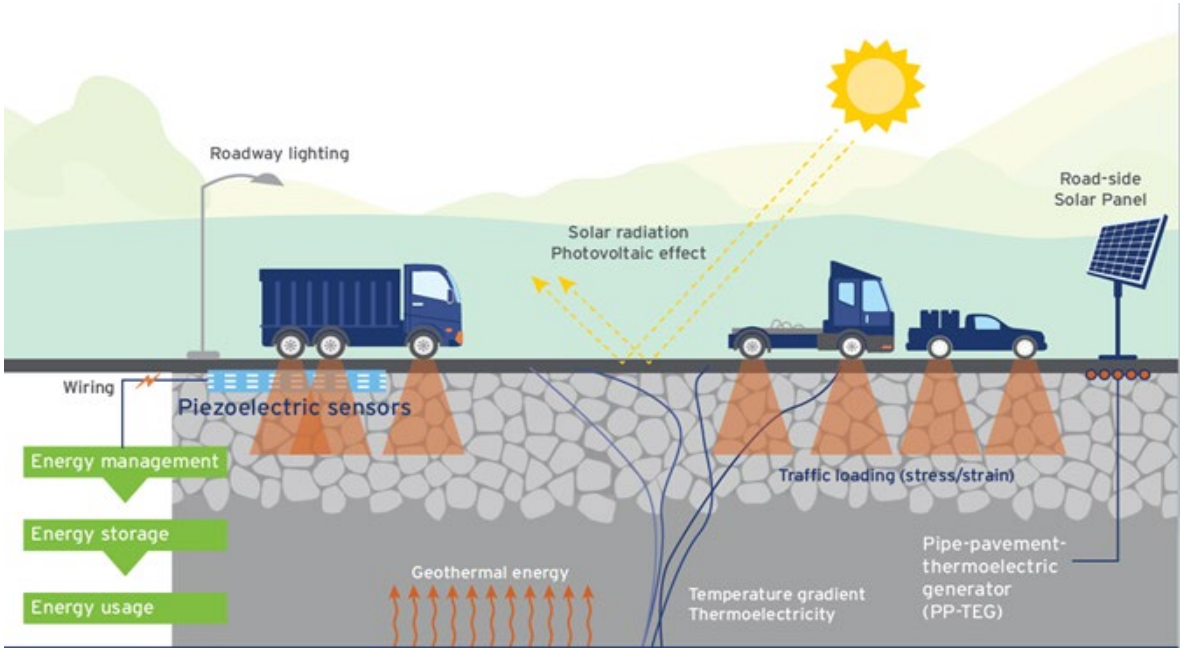
Piezoelectricity refers to the generation of electricity by subjecting a material to mechanical stress. The pilot versions of the technology included application in various forms including barbeque lighters and dance floors.

In 2008, the first large-scale piezoelectric implementation was made by the East Japan Railway Company in the Tokyo station; this involved the generation of electricity from passengers' footsteps to power ticket machines. Currently, around the globe, research and development (R&D) is ongoing to commercialize piezoelectric technology for wider application on roads, especially in cities with heavy traffic.

Piezoelectricity generation is similar in concept to wind energy, where electricity is generated by converting mechanical power from wind turbines. The concept of piezoelectric roads arises from harnessing the potential generation of electricity from vehicular movements.

This involves the conversion of mechanical energy from vehicular movement into electricity. With the amount of available car park and miles driven data constantly increasing, there is ample scope for tapping this source of energy. With traffic increasing around the globe, the concept of harvesting electricity through the movement of vehicles is certainly worth considering.

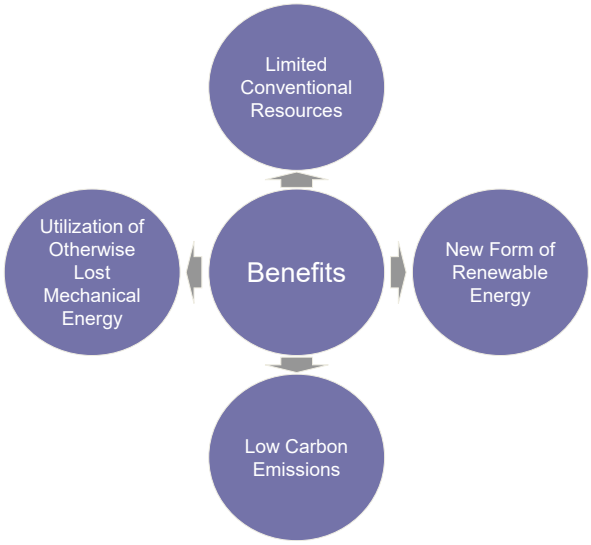
Figure 30. Renewable Energy: Various Forms



Source: Citi Research

Past Projects and Studies

Figure 31. Benefits of Piezoelectric Roads



Source: Citi Research

R&D for this technology is in progress in a few countries, including various pilot projects and experimental installations. New materials and designs are being explored to improve the performance and cost-efficiency.

In 2009, Israel's Innowattech launched an experiment in which it established the world's first piezoelectric road: a 1-kilometer (km) stretch of piezo road on a highway in northern Israel. The company claimed that for a single-lane, 1 km-long piezoelectric highway (PEH), it was possible to generate as much as 100 kilowatts

(kW) of electricity per hour. However, neither detailed information regarding project specifications nor experimental raw data were made available by the company.

A recent study[24] on piezoelectric roads, prepared for the California Energy Commission by authors from the University of California, Merced Department of Mechanical Engineering, concluded that one lane of a one-mile-long roadway could potentially generate 72,800 kilowatt-hour (kWh) of energy every year based on mid-sized vehicles. This could go up to 907,873 kWh for heavy trucks, which the report claimed implied a reduction of 300 metric tons of CO_2. The project is part of California's goal of generating 100% of its electricity from renewable and zero-carbon energy sources by 2045.

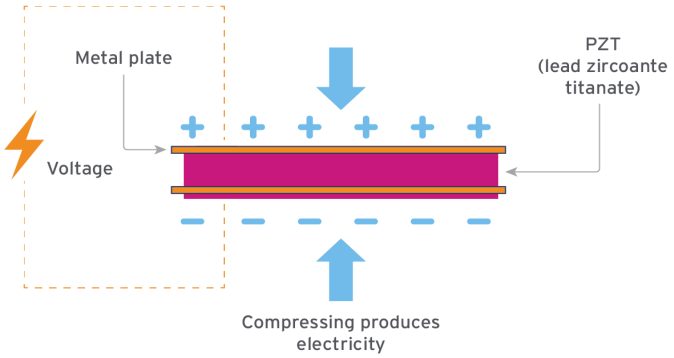
In 2021, in another study published by International Journal of Creative Research Thoughts (IJCRT), it was claimed that one lane of one-mile-long roadway could potentially generate around 70,800 kWh of energy every year. [25] However, no details were given regarding the size of vehicles assumed.

Process

To create a piezoelectric road, strips or tiles of piezoelectric materials, mostly Lead Zirconate Titanate placed between metal plates, are embedded within a road's surface.

These materials can come in various forms (eg, pads or sensors) and are strategically positioned to capture the mechanical energy produced by vehicles passing over them. The moving vehicles put vertical force and apply pressure on the piezoelectric materials, generating electricity.

Figure 32. Piezoelectric Mechanism



Source: Citi Research

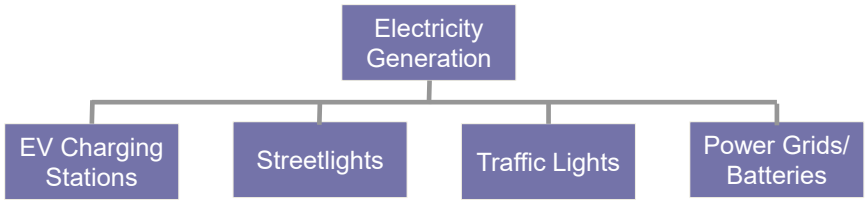
Consumption Possibilities

The energy generated from piezoelectric roads can be consumed or stored in various forms.

[24] Jian-Qiao Sun, Tian-Bing Xu, and Atousa Yazdani, *Ultra-High Power Density Roadway Piezoelectric Energy Harvesting System*, California Energy Commission, June 2023.

[25] Pravin Wale et al., "Generation of Electricity from Roads by Using Piezoelectric Sensors," *International Journal of Creative Research Thoughts*, Vol. 9, No. 1, June 2021.

Figure 33. Consumption and Storage Possibilities



Source: Citi Research

- **EV charging machines:** As EV penetration increases, more EV charging stations will likely be set up around highways; these stations could source their energy from piezoelectric power generation.
- **Streetlights:** Piezoelectric roads could be linked up directly to power streetlights. This could be useful on highways where the power grid is set up just to light roadside lampposts.
- **Traffic lights:** Piezoelectric roads could produce energy to power traffic lights in urban areas.
- **Storage:** Unconsumed piezoelectricity could be stored in the grid for household and commercial consumption, as well as in battery form.

Comparison to Solar and Wind Energy

Two comparable low-carbon energy sources to piezoelectricity are wind energy and solar energy. We compare the three energy sources in the table below.

Figure 34. Comparison vs. Solar and Wind Energy

	Piezoelectric Roads	Solar Energy	Wind Energy
Source of Electricity	Vehicle Movements	Sun	Wind
Weather Dependency	Not Dependent	High	High
Carbon Emissions	Low	Low	Low
Relative Space Requirement	Low (set up on already-built roads)	Medium (mostly built on roofs)	High
Urban Deployment Opportunity	Medium-to-High	High	Low
Adoption	Low	High	High

Source: Citi GPS, Citi Research

Bottlenecks

Limited research is available on the feasibility and sustainability of piezoelectric roads, as they are still in an experimental stage and not yet implemented on a large scale. Initial investment costs remain an area of concern around the deployment of piezoelectric materials.

Unanswered questions remain regarding the longevity of the materials, as the technology is still evolving. To measure the potential lifetime of a system, more extensive fatigue tests are needed.

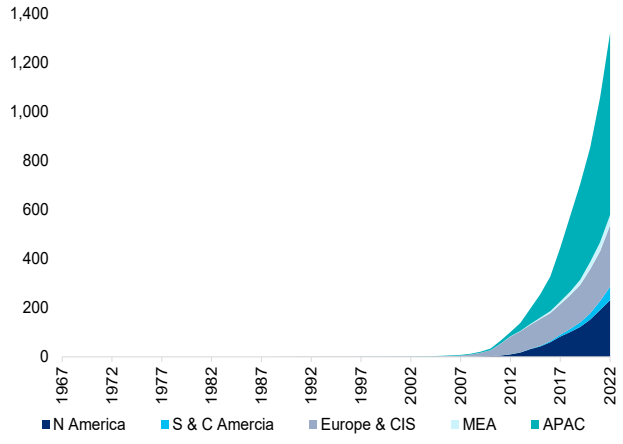
Future Potential

Piezoelectric road technology is still at an early stage, but with focus rising on low-emission renewable energy sources, piezoelectric penetration could increase significantly, as we have seen for solar and wind energy (Figure 35 and Figure 36) over the past decade.

With R&D, the technology is being developed on a bigger scale for roads, especially within the cities with heavy vehicular movement. Piezoelectric roads could be a radical development, as they use mechanical energy that is otherwise wasted, as well as existing infrastructure.

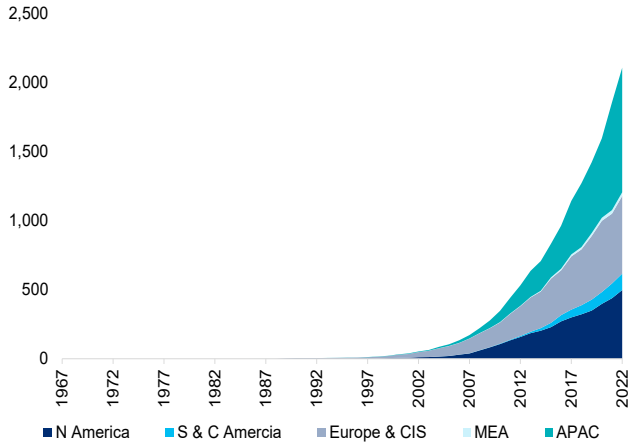
Ongoing research on feasibility, efficiency, sustainability, and cost-effectiveness will be pivotal for achieving widespread adoption.

Figure 35. Solar Capacity 1967-2022 (TWh)



Source: Citi Research

Figure 36. Wind Capacity 1967-2022 (TWh)



Source: Citi Research

Implementation Potential and Cost Analysis

Based on the California Energy Commission study discussed earlier, if the U.S. decides to convert only 5% of the current road infrastructure (4.19 million miles) into piezoelectric roads, this could generate around 50 billion watt-hours of renewable energy, which is equivalent to 5.6% of 2022 renewable electricity generation in the U.S. and around 1.2% of total U.S. electricity generation. This would be enough to power 5,000 households for one year, based on average 2021 consumption of around 10,600 kWh/household/year (source: EIA).

Figure 37. Scenario Analysis Based on Various Levels of U.S. Implementation of Current Infrastructure

U.S. Implementation	5%	10%	15%	20%
KWh per year / Mile (80% Mid-sized; 20% Heavy vehicles)	239,815	239,815	239,815	239,815
Total U.S. lane miles	4,190,000	4,190,000	4,190,000	4,190,000
Total miles based on 20% implementation	209,500	419,000	628,500	838,000
Estimated electricity generated/year (BWh)	**50**	**100**	**151**	**201**
Total Electricity generated, 2022 in BWh	4,231	4,231	4,231	4,231
Piezo as % of Total Electricity	**1.20%**	**2.40%**	**3.60%**	**4.70%**
Renewable electricity generated, 2022 in BWh	901	901	901	901
Piezo as % of Total Electricity	**5.60%**	**11.20%**	**16.70%**	**22.30%**
U.S. energy-related CO^2 emissions, 2022 (million metric tons)	4,964	4,964	4,964	4,964
Estimated CO^2 reduction (million metric tons)	63	126	189	251
Piezo-led CO^2 Reduction	**1.30%**	**2.50%**	**3.80%**	**5.10%**

Source: Energy Information Administration (EIA), American Road & Transportation Builders Association (ARTBA), Citi Research

Cost Analysis

Initial investment costs remain an area of concern around the deployment of piezoelectric materials. Although the exact cost dynamics are still unknown, a 2013 study prepared for the California Energy Commission suggested a levelized cost of energy (LCOE) in the range of $0.08-$0.20/kWh.[26]

Figure 38. LCOE of Renewable Energy Sources in 2022 vs. 2010

	Total Installed Costs (2022 $/kW)			Capacity Factor (%)			Levelized Cost of Electricity (2022 $/kW)		
	2010	2022	% change	2010	2022	% change	2010	2022	% change
Solar PV	5,124	876	-83%	14	17	25%	0.445	0.049	-89%
Onshore Wind	2,179	1,274	-42%	27	37	37%	0.107	0.033	-69%
Offshore Wind	5,217	3,461	-34%	38	42	11%	0.197	0.081	-59%

Source: Citi Research, IRENA.org

The dollar cost per Watt for solar energy was down around 90% in 2022 versus 2010. We do not expect a similar reduction in cost for piezoelectric technology, but there will be cost improvements helped by economies of scale if this technology is adopted at a large scale.

Parallel Technology

Piezoelectric road technology could be seen as a potential upgrade to the solar road technology that has already been adopted in various parts of the world, including the U.S., France, the Netherlands, Japan, and China. These solar roads are embedded with energy-harvesting modules that can power streetlights and melt snow. However, they are heavily weather-dependent, and the road infrastructure does not make for a smooth driving experience, as solar roads lack strong tire grips and demand high maintenance costs.

Conclusion

In summary, piezoelectric road technology is still in the pilot stage. However, it does provide an attractive potential option for sustainable energy generation. Piezoelectric roads could be an important addition to existing forms of renewable energy, like wind and solar, especially since wind and solar energy require wind and sun, respectively, and are thus largely weather-dependent. Piezoelectric roads could help harness the energy already applied to roads on a large scale.

[26] Davion Hill, Arun Agarwal, and Nellie Tong, *Assessment of Piezoelectric Materials for Roadway Energy Harvesting: Cost of Energy and Demonstration Roadmap*, DNV KEMA Energy & Sustainability, January 2014.

Quantum Sensing

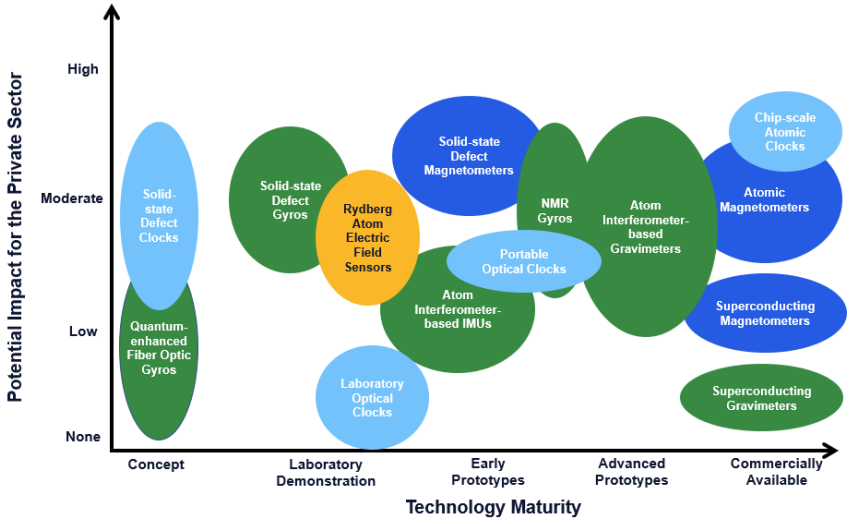
Tahmid Quddus Islam
Citi Global Insights

Data is increasingly integral to the global economy. Sensors – the eyes and ears of our technological world – collect the data that enable some of society's most critical systems. As we push the boundaries of what is possible in fields such as precision manufacturing, conventional sensors are being challenged by the limits of classical physics. In recent years, a new type of advanced sensor technology has been emerging – next-generation quantum sensors.

Quantum sensors use the principles of quantum physics (ie, the physics of low energy or small-scale systems, where energy exists in discrete packets called "quanta") to make measurements at the level of atoms or photons. They have the potential to be extremely sensitive to their environment, detecting the miniscule changes in physical quantities that would be imperceptible to conventional sensors. The ability to probe the physical world with an unprecedented level of detail, and the new insights this is expected to bring, will generate a new pool of data to help power decision-making. As such, quantum sensors have the potential to bring benefits to society from areas as broad as navigation to healthcare, defense, and even agriculture (see our Citi CGI report Quantum Sensing: Tech's New Eyes and Ears).

The challenge with understanding quantum sensors is that there are many different types, all at different stages of technological maturity and commercialization, as seen in Figure 40. Some quantum sensors measure time, acceleration, or rotation, whereas others measure the likes of electric fields, magnetism, or gravity. And not all quantum sensors work in the same way.

Figure 39. Potential Impact of Quantum Sensors vs Technological Maturity



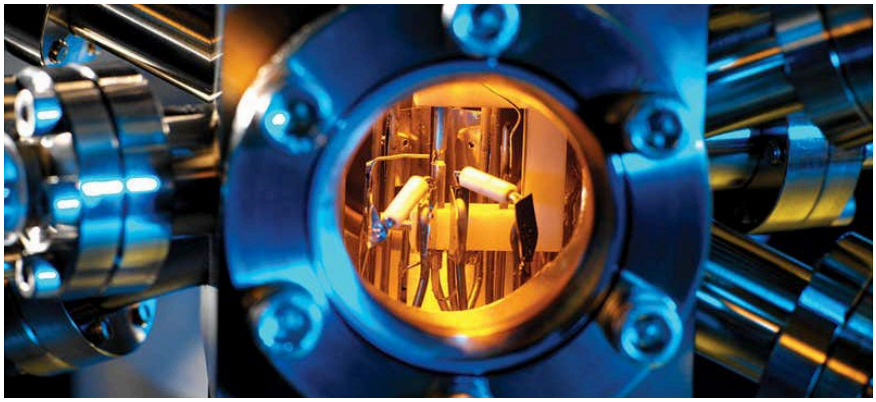
Source: Quantum Economic Development Consortium (QED-C), "Quantum Sensing Use Cases: Prospects and Priorities for Emerging Quantum Sensors," September 2022.

Quantum sensing, therefore, is really an umbrella term for all the different types of quantum sensors. It is also one of the three main key types of quantum technologies – alongside quantum computing (see our Citi GPS report Quantum Computing: Moving Quickly From Theory to Reality) and quantum communications. While there are numerous types of quantum sensors, we focus on four examples that we believe have the most near-term potential for disruption.

Atomic Clocks

Atomic clocks measure time using the resonance frequencies of atoms and have been used commercially for decades. They underpin much of our infrastructure, including telecommunications, and have an estimated market size of over $500 million.[27] Today, the best conventional cesium-based atomic clocks are accurate to around 0.03 nanoseconds per day, equating to an error of approximately 1 second per 100 million years.[28] The next generation of atomic clocks, optical atomic clocks, have been reported to improve on this 100-fold in laboratory experiments, being accurate to around 1 second per 30 billion years – twice the age of the universe.

Figure 40. An Optical Atomic Clock



Source: National Physical Laboratory (NPL)

Conventional atomic clocks are already used in numerous applications, such as those listed below. The increased precision that optical atomic clocks offer is expected to further help test theories of fundamental physics and could lead to the redefining of the definition of a second.[29] Some practical areas they are expected to help in include:

- **Synchronizing High-Speed Networks:** Any data being sent over a network is entirely reliant on the precise timing of data packets, without which errors can occur or communication lines can become congested. This is why the International Telecommunication Union (ITU) places limits on the difference in speed between two networks to around 1 second per 3,000 years – an accuracy and stability only achievable using atomic clocks.[30] It is thought that the next generation of ultra-stable, highly accurate optical atomic clocks could help improve the synchronization of high-speed networks.[31]

- **High-Frequency Trading:** The use of atomic clocks to provide extremely accurate timekeeping to ensure high-frequency trading (HFT) gaps are

[27] Future Market Insights, "Atomic Clock Market: Report Preview," August 2023.
[28] NIST, "Timekeeping and Clocks FAQs," accessed December 13, 2023.
[29] Nathaniel Sherrill et al., "Analysis of Atomic-Clock Data to Constrain Variations of Fundamental Constants," *New Journal of Physics*, September 6, 2023; Jeremy Hsu, "Optical Atomic Clocks Are Ready to Redefine Time," IEEE Spectrum, May 27, 2020.
[30] Thales Group, "Atomic Clocks and the Importance of Being on Time," November 18, 2014.
[31] UK National Physical Laboratory (NPL), "Optical Atomic Clocks," accessed December 13, 2023.

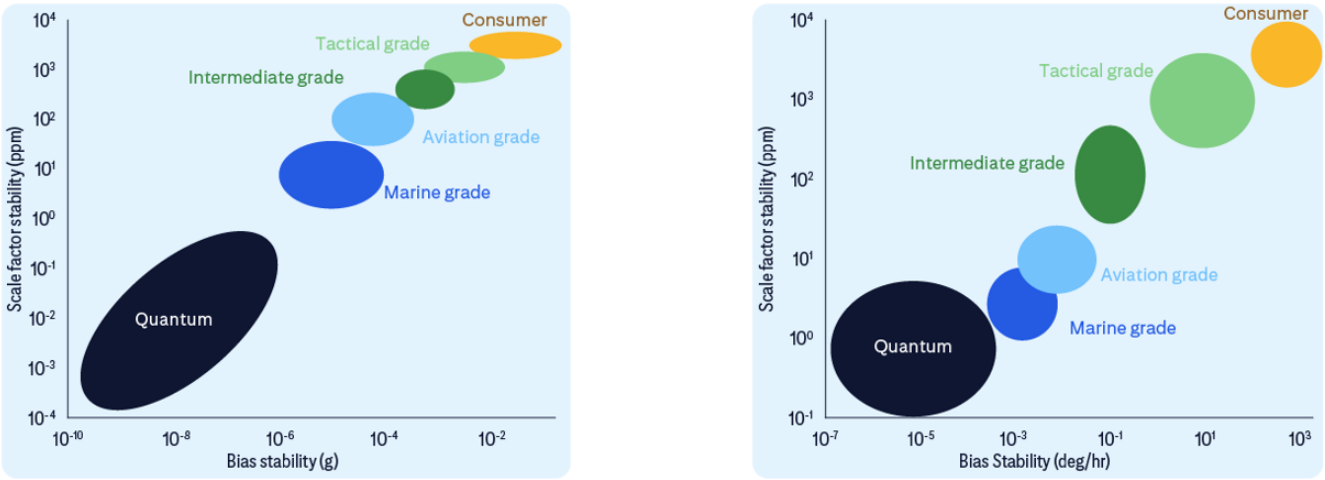
addressed has been discussed for some time now.[32] The steady growth of HFT since the 2008 Global Financial Crisis has meant the accurate timestamping of trades has become of increasing importance. In 2008, the Financial Industry Regulatory Authority (FINRA) required stock market trading accuracy to 1 second, but by 2016, the U.S. Security and Exchange Commission (SEC) approved a tighter standard of 50 milliseconds for certain activities. Hedge funds have also been reported to be looking at using atomic clocks to identify novel ways for "executing synchronized trades in multiple exchanges."[33]

- **Global Positioning System (GPS) Accuracy:** GPS operates on signals received from satellites at the speed of light (~300,000,000 meters per second), meaning that even a small error in timing of just 1 microsecond (a millionth of a second) can result in positioning errors of around 300 meters. The ability of optical atomic clocks to be accurate to picoseconds (trillionths of a second) could enable significant improvement in the location accuracy and autonomy of GPS.

Quantum Inertial Sensors

Global navigation satellite systems (GNSSs), such as GPS, underpin many of our daily activities, from determining the best route to work to ensuring that goods are delivered all around the country. The GPS market size is estimated to have surpassed $100 billion in 2023, but systems are very vulnerable to jamming, blocking, or spoofing.[34] One report put the potential loss to the UK economy alone from a GNSS outage at over £1 billion (≈$1.2 billion) per day, with two-thirds of the impact affecting critical infrastructure such as emergency services and roads – meaning that such an outage is also likely to cost lives.

Figure 41. Stability of Quantum vs Classical Accelerometers and Gyroscopes



Source: Glasgow University

As with any type of GNSS, GPS is not able to be used in locations that do not have access to signals from GPS satellites, such as below ground or under water. In GNSS-denied environments, inertial sensors, such as accelerometers and

[32] Jacob Aron, "Atomic Time Lord to Battle Sneaky High-Speed Trades," *New Scientist*, April 16, 2014.
[33] Miles Weiss and Zachary Mider, "Legendary Hedge Fund Wants to Use Atomic Clocks to Beat High-Speed Traders," *Bloomberg*, July 7, 2016.
[34] Grand View Research, "Global Positioning Systems Market Trends: Report Summary," November 2023.

gyroscopes, can be used in navigation, measuring acceleration and rotation that is then converted into a location. The challenge with using any type of inertial sensor for navigation is that their accuracy eventually degrades, a phenomenon known as "bias drift." This is where errors (or "biases") accumulate over time, ultimately making the sensors less stable and unreliable for long-distance travel. Quantum inertial sensors use atoms to measure acceleration and rotation, and in the case of accelerometers, have been shown to provide a 50-fold improvement in stability over their classical equivalents.[35]

Unlike atomic clocks, most quantum inertial sensors for navigation are still in the early field-testing stage. However, a number of practical use cases are being investigated:

- **Below Ground:** A grant has been awarded in the UK to investigate the use of quantum sensing on the London Underground train network, 45% of which resides below ground. Quantum inertial sensors may be able to help trains register their position on the network more accurately, thus offering greater reliability and a safer experience for passengers.[36]

- **Maritime:** The UK's Royal Navy has been testing quantum inertial sensors for navigation on one of its research vessels.[37] Maritime travel is prone to having areas of poor GNSS connection. A significant proportion (around 20%) of the estimated £1 billion/day loss to the UK economy from GNSS failure is expected to come from the maritime sector – something that quantum sensors could help mitigate.[38]

- **Air:** Following the U.S. Executive Order "Strengthening National Resilience Through Responsible Use of Positioning, Navigation, and Timing [PNT] Services," the U.S. Department of Transportation (U.S. DOT) is investigating alternative navigation methods.[39] In late 2023, the U.S. DOT initiated a Request for Information (RFI) for "technology capable of providing critical infrastructure users and operators positioning and/or timing information that is derived independently from Global Navigation Satellite Systems (GNSS)."[40] Quantum inertial sensors are thought by many to be a possible solution.

Quantum Magnetometers

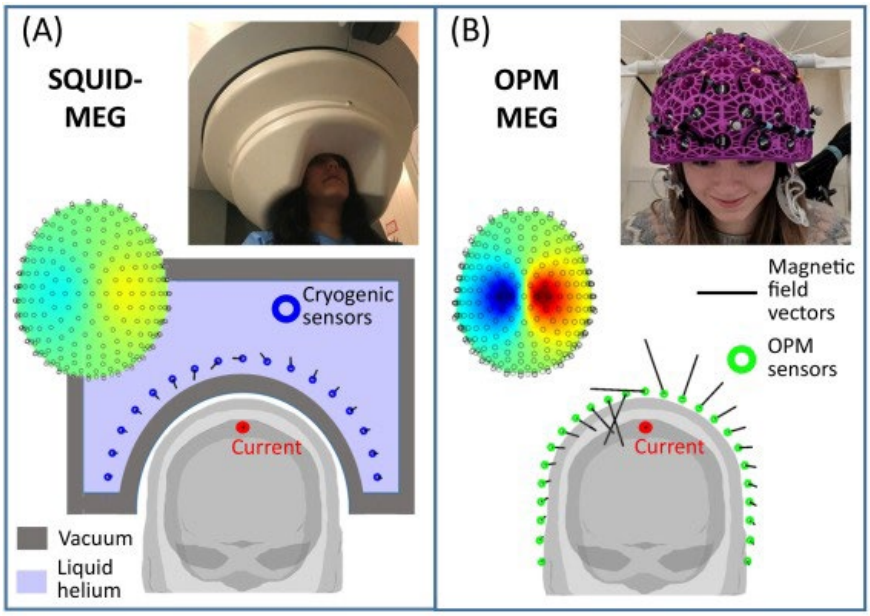
Sensors that detect magnetic fields (or "magnetometers") are a cornerstone of diagnosing health conditions. Furthermore, quantum sensors that detect magnetism, such as Magnetic Resonance Imaging (MRI) technology, have been heralded as a revolution ever since the first patient scan in the 1970s.[41] Similarly,

---

[35] Simon Templier et al., "Tracking the Vector Acceleration with a Hybrid Quantum Accelerometer Triad," *Science Advances*, Vol. 8, No. 45, November 9, 2022.
[36] UK Engineering and Physical Sciences Research Council, "Quantum Sensing on the London Underground," April 27, 2023.
[37] UK Royal Institution of Naval Architects, "Research Ship Enables Royal Navy to Test Cutting-Edge Navigation System and New Radars," July 7, 2023.
[38] London Economics, *The Economic Impact on the UK of a Disruption to GNSS*, August 2023.
[39] SAM.gov, "RFI -- Complementary Positioning, Navigation, and Timing (PNT)," accessed December 13, 2023.
[40] Billy Mitchell, "Transportation Department Seeks GPS Alternatives to Fortify US's Position, Navigation, and Timing System," *Fedscoop*, September 25, 2023.
[41] American Physical Society, "This Month in Physics History: July 1977: MRI Uses Fundamental Physics for Clinical Diagnosis," July 2006.

brain scans that use a process known as magnetoencephalography (MEG) are able to measure the magnetic fields generated by electrical signals in the brain.[42] The MEG market size is estimated to be over $270 million and currently uses a type of quantum magnetometer known as a Superconducting Quantum Interference Device (SQUID).[43]

Figure 42. SQUID-MEG vs. OPM-MEG



Source: "Magnetoencephalography with optically pumped magnetometers (OPM-MEG)"

SQUIDs need to be cooled to near "absolute zero" (around −273.15°C), which requires cryogenic cooling. This results in brain scanners being bulky and expensive to maintain. However, a new type of quantum magnetometer, known as an Optically Pumped Magnetometer (OPM), has recently seen considerable progress in its development. OPMs get their names from the fact that they use lasers to "pump" atoms into a specific state. These do not need to be supercooled and consequently can be made much more compact, resulting in wearable brain scanners. Whilst it's important to note that there are numerous potential applications of this technology in various industries, in the examples below, we focus on healthcare applications:

- **Epilepsy Monitoring:** OPMs could potentially be used to enable the high-fidelity monitoring of epilepsy without surgery (doctors currently make assessments by removing part of the skull to attach electrodes directly to the brain).[44] Furthermore, in instances where surgery is required, there is notable investment being made into the use of OPM-MEG for pre-surgical planning.[45]

[42] Athinoula A. Martinos Center for Biomedical Imaging, "A History of fMRI," accessed December 14, 2023.
[43] Grand View Research, "Magnetoencephalography Market Size: Report Summary," accessed December 14, 2023.
[44] UK National Quantum Technologies Programme, "Quantum Brain Scanner," PDF, October 2021.
[45] Wellcome Centre for Human Neuroimaging, "MEG Team Awarded Funding to Enhance Technology Aiming to Improve Accuracy of Pre-Surgical Planning for Epilepsy," accessed December 14, 2023.

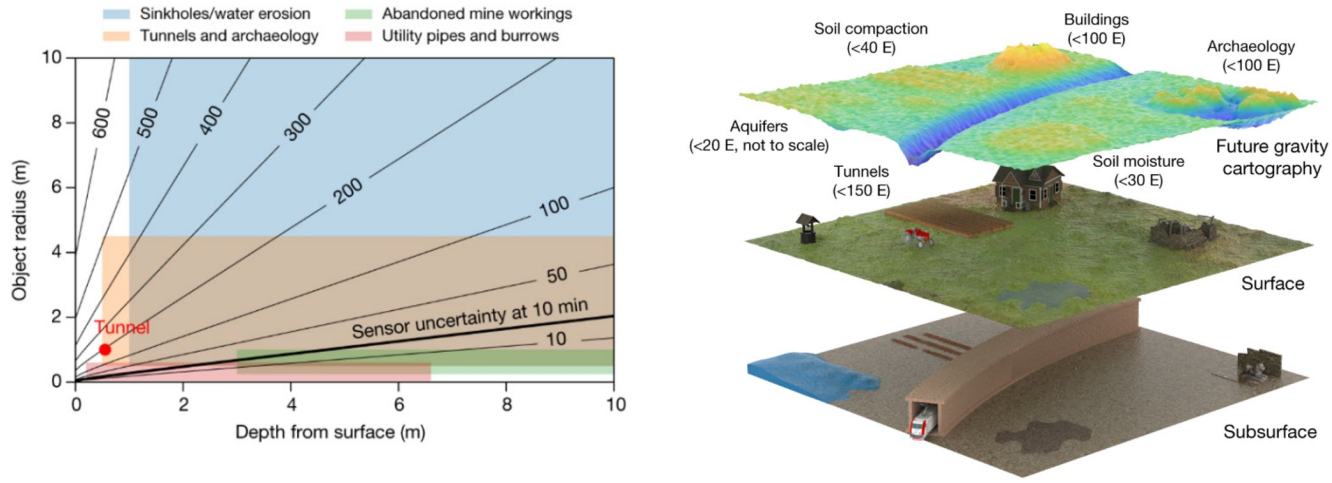
- **Alzheimer's Diagnosis:** The latest OPM-MEG sensors can be put in close proximity to the head, providing the additional advantage of enabling more accurate 3D scans. Such scans could provide a way of identifying neurological conditions. Many neurological conditions are characterized by the slowing down of messages in the brain. By monitoring real-time changes, OPM-MEGs have the potential to assess the pace at which neural signals travel across the brain. A decline in signal speed could indicate the presence of conditions such as Alzheimer's or other brain-related diseases.[46]

- **Autism Research:** A hospital in Toronto is investigating the use of OPM-MEG technology to scan children identified as having a higher likelihood of developing autism.[47] The trial involves the regular scanning of the children's brains starting at 12 months of age to track their brain function development and compare it to that of children who develop autism. The compact helmet design being used by researchers allow scans of children that would otherwise not be possible, as the inability of children to sit still makes using traditional SQUIDs impractical.

Quantum Gravimeters

Sensors that measure gravity (or "gravimeters") typically use classical methods such as measuring a spring with a mass attached to it or the free-fall of a mass in a vacuum. Quantum gravimeters, on the other hand, use a process known as atom interferometry, which exploits the wave-like nature of atoms and their interference patterns – effectively measuring the pull of gravity on individual atoms.

Figure 43. Resolution of Quantum Gravitational Sensors



Source: "Quantum sensing for gravity cartography" Paper

Quantum gravimeters offer numerous benefits, such as the potential to be more stable over time, as well as the opportunity to remove the noise arising from ground vibrations. For instance, by using two quantum gravimeters set-up as a gradiometer (which measures the gradient of gravitational changes), vibrations can be identified and removed – this promises to reduce measurement times from several minutes to just a few seconds. As it currently stands, quantum gravimeters can achieve a level of precision not too dissimilar to the classical gravimeters used in industry. However,

[46] Medical Design Briefs, "Quantum Brain Sensors Could Spot Dementia," January 1, 2022.
[47] UK Quantum Technology Hub: Sensors and Timing, "Wearable Brain Imaging System Installed in Toronto for Autism Research," August 24, 2021.

the ability for quantum gravimeters to mitigate the noise from ground vibrations means that the future improvements in the sensitivity of these devices could translate directly into either higher spatial resolution or faster reading times. Consequently, it has been suggested that there could yet be a 10x to 100x improvement in instrument sensitivity that could be deployed in the field.[48]

Historically, the accuracy and physical size of quantum gravimeters have restricted them to laboratory settings. However, recent improvements have resulted in them being used in the field for the first time, and it is thought that quantum gravimeters may be especially valuable in applications where detecting minute details can make a significant difference, such as:

- **Civil Engineering:** Described as potentially "Google Maps for the Underground", the use of quantum gravimeters could help save time and money by more accurately scanning underground infrastructure or detecting hazards such as sinkholes. For example, by one estimate, 16% of damage to the UK's utility infrastructure caused by contractors is thought to be due to a lack of awareness of the existing pipes underground, with the cost of this damage being estimated to be up to £5.5 billion (≈$6.8 billion). A collaborative project between industry players and academics, called "Quantum Technology – Potential for Rail Infrastructure" (QT – PRI), is exploring the use of quantum sensors to identify underground assets. The collaborators believe there are around 190,000 different assets underground that quantum sensors could help identify.

- **Energy & Mineral Exploration:** Gravity surveys (that measure local gravitational fields at locations of interest) are regularly performed in the energy sector to map structures in the search for new energy reserves. Quantum gravimeters have recently been tested on a ship at sea as a new way of mapping the ocean floor.[49] This technique could be used to help navigation, and companies like BP have already started hosting workshops on the use of quantum sensors and to train their employees on the technology.[50]

- **Environmental Monitoring:** Another potential use case for quantum gravimeters is improving earthquake detection. Once an earthquake occurs, it is possible to use quantum gravimeters to detect immediate gravitational changes resulting from mass shifts in the earth, even before seismic waves are felt. A gravity-based alert might only offer a few extra seconds, but that time can be crucial in saving lives.

Conclusion

In conclusion, the potential use cases for quantum sensors are vast, and the derivative economic impact could be hugely transformative for countless industries.

The proprietary analysis in our *Quantum Sensing: Tech's New Eyes and Ears* report suggests most underestimate the potential of the quantum sensing market. For instance, given navigation is critical for commercial and military aircraft, there could potentially be up to a market of $3.25-5bn for just one type of quantum sensor, quantum inertial sensors, in just the aviation industry alone.

[48] Nature, "Quantum sensing for gravity cartography," February 23, 2022 https://www.nature.com/articles/s41586-021-04315-3

[49] UK Quantum Technology Hub: Sensors and Timing, "Quantum Sensor for Gravity Gradiometry Validated at Sea," September 18, 2023.

[50] UK Quantum Technology Hub: Sensors and Timing, "Exploring Quantum for the Oil and Gas Industry," May 22, 2018.

Retail Media

A Fast-Growing Digital Ad Format Benefitting Both Advertisers and Retailers/Marketplaces

Thomas A Singlehurst, CFA
Head of EMEA EMT Sector
Citi Research

Ygal Arounian
U.S. Internet Analyst
Citi Research

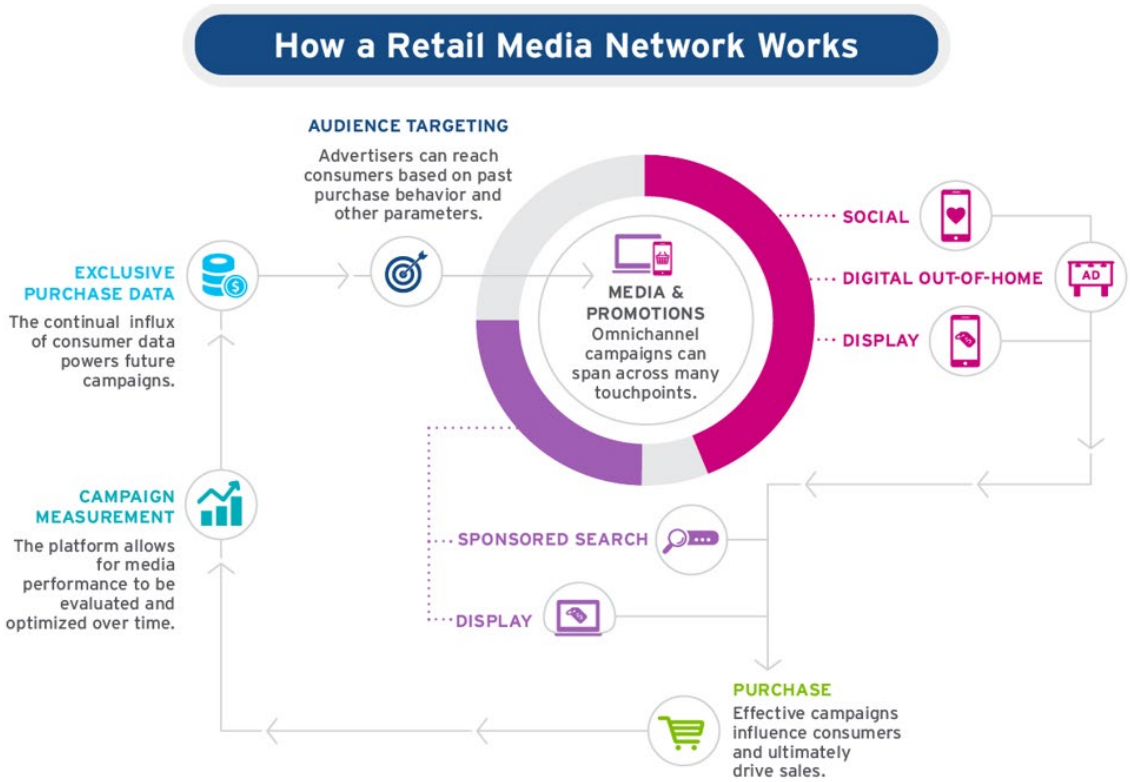
Maxwell Moore
U.S Internet Research Team
Citi Research

Retail media is a digital advertising format tied either to a retailer's website or to its products on a third-party website. The advertising format leverages a retailer or brand's first-party data to target consumers within their shopping journey and utilizes closed-loop measurement for attribution of that sale (directly tying a sale to the ad). This advertising model is leading to both revenue and profitability growth across multiple companies and sectors.

Retail media has propelled by a confluence of factors – the growth of e-commerce; the shifting privacy landscape; and better data collection, targeting, and measurement of ads – to become one of the fastest-growing verticals within digital advertising. Retailers are looking to improve monetization, and advertisers are drawn to a model that targets shoppers on their purchase journey.

As stricter data privacy standards have made third-party (3P) data harder to use for targeted advertising, retailers and advertisers, are increasingly focusing on this growing but still nascent ad format.[51]

Figure 44. Retail Media Networks' Closed-Loop Measurement Capabilities Are Critical Factors of Adoption for Advertisers



Source: Citi Research, Quotient

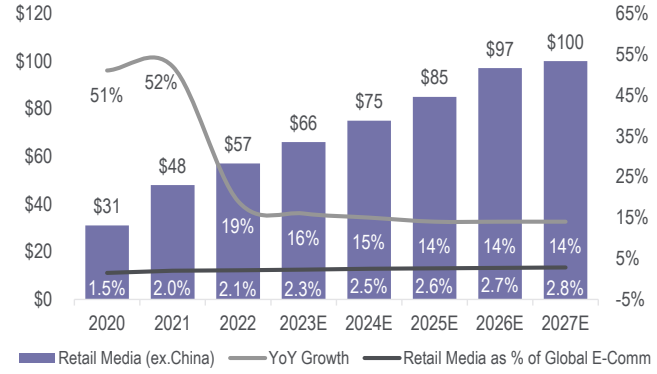
[51] One relevant data privacy initiative is Apple enabling users to block the identifier for advertisers (IDFA) – which facilitates third-party tracking – as part of its App Tracking Transparency initiative. Cookies are also being deprecated on many web browsers, including on Google Chrome beginning in 2024.

A Large, Fast-Growing Market

We estimate that retail media is currently a $65 billion market globally (excluding China) and expect it to grow to $110 billion by 2027. That represents around 12% of total digital advertising dollars today (excluding China) and an estimated 14% in 2026. Those dollars are coming from various sources, and much is brand new spend, but a large portion is coming from offline sources as well.

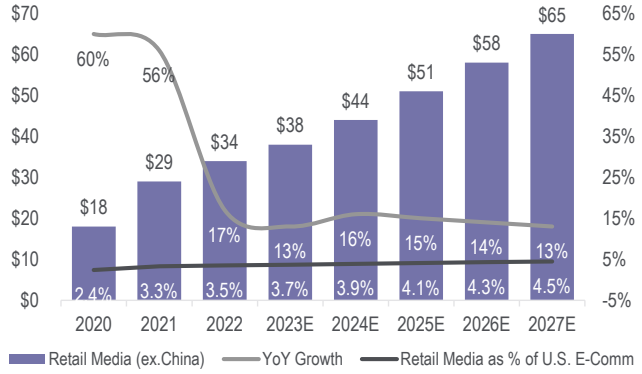
Much of the projected spend will be driven by consumer packaged goods companies (CPGs), who historically are the largest spenders of shopper marketing budgets (which we expect will continue to shift online as e-commerce grows). However, spending should rise across more verticals over time, too.

Figure 45. Global Retail Media Forecast (ex. China) 2020-27E ($ bns)



Source: Citi Research Forecasts, GroupM

Figure 46. Retail Media Forecast 2020-27E ($ bns)



Source: Citi Research Forecasts, GroupM

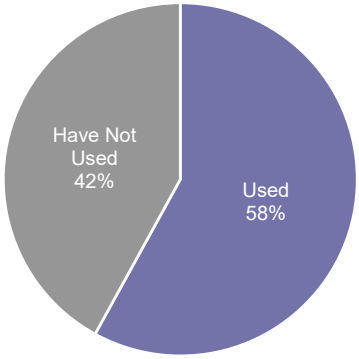
Taking Share of Digital Ad Spend and Offline Budgets

As shopping moves online, how trade and consumer promotion budgets are being spent has substantially changed. Less important are aisle-end displays and in-store promotions, as these budgets are now being repurposed to buy prominence in search results on e-commerce platforms, optimize listings via better content and tagging, and manage influencer communities. For advertisers, this represents a significant structural overhaul of how they do business, forcing them to integrate sales with marketing as trade promotion budgets (eg, advertising spent in store) shift online to retail media. For the agencies, it is a significantly novel discipline that potentially brings with it substantial incremental advertising-related budgets. For AdTech firms, it is an opportunity to be an intermediary for a new vertical, drive a supply and demand marketplace, and supply technology to retailer platforms. And for e-commerce platforms, it represents a massive incremental source of revenue and profit.

We see retail media's share of total digital ad spend increasing in the U.S. There is still a lot of room for adoption by marketers to rise, with just over half having used retail media networks (RMNs) in 1H22. We estimate that in 2022, retail media in the U.S. represented 14.7% of total digital ad spend, up from 3.4% in 2017, and we expect it to account for 16.3% in 2025. This compares to other leading ad channels' 2022 share in the U.S., for example, social at 27.1%, video at 30.1%, and search at 41.6% (note there is overlap between the categories). We also look to China given its more mature retail media market; the country sees about a 40% share of digital ad spend in retail media or e-commerce-related advertising. While there are structural differences (particularly within search) that

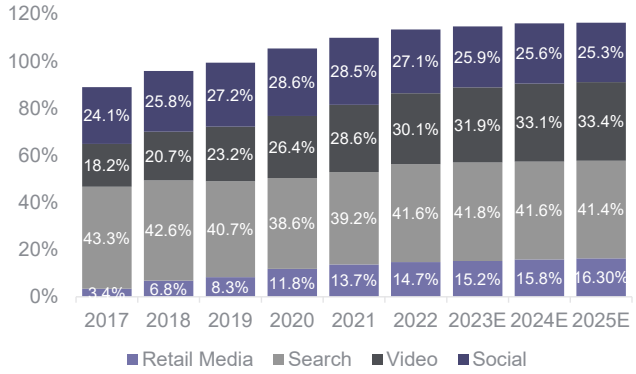
make such a high-water mark unlikely to be achieved in the U.S., China's case remains a useful illustration of the large potential opportunity.

Figure 47. Marketers Who Used Retail Media Networks in the Past Year (Survey Conducted in Summer 2022)



Source: Citi Research, Association of National Advertisers: Retail Media Networks: A Forced Marriage or Perfect Partnership?

Figure 48. U.S. Digital Ad Formats as a % of Total U.S. Digital Advertising Spend



Note: Retail Media is based on Citi's own forecasts for retail media and U.S. digital ad spend. Search, Video, and Social are based on eMarketer projections from March 2023. Also note that there are overlaps in ad spending between some formats, so they should not be thought of as adding to 100%.

Source: Citi Research, eMarketer

Two Primary Retail Media Formats: Onsite and Offsite

Onsite advertising is any retail media ad that is bought and/or displayed on the retailer's website or platform. The most common ad formats are sponsored (promoted) listings, which are when a seller or brand pays a retailer website to place its product higher up on the search or browsing results page to increase the product's visibility. The other primary format is a display ad on a retailer website. Both ad formats allow the retailer to monetize its digital storefront while keeping the consumer on its owned properties throughout the buying process.

Figure 49. The Retail Media Path to Purchase Spans Across Retailer Sites, the Open Web, and Physical Retail

Source: Citi Research, CitrusAd

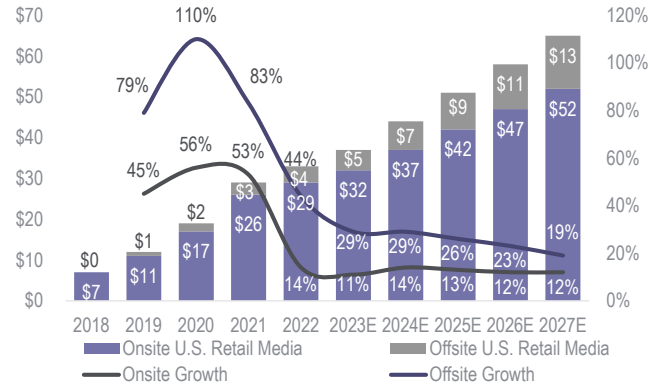
Offsite retail media is when a brand, agency, or retailer creates a personalized ad for a product that is served across the open web on a property that is not the same as where the product transaction will occur. This allows both the retailer and brand to increase their reach beyond the content walls of its retailer partner, which can help improve conversion rates. Offsite ads can be in any format – eg, display, connected TV (CTV), or video – as long as the ad is a joint one between the brand and the retailer and leads back to the retailer's website for the transaction.

An offsite ad could be a co-branded display ad from a brand and retailer – for example, an ad for a Sony TV being sold on Best Buy's website placed on *The Wall Street Journal*'s homepage. When a consumer clicks through the display ad, they are redirected to the brand's product on the retailer site where they can seamlessly complete the purchase.

Onsite represents the majority of retail media spend today, but offsite has been growing at around double the pace, which we expect to continue – we forecast offsite ad dollars to triple in the U.S. by the end of 2027 to $13 billion and account for 20% of total retail media spend, up from 13% in 2022. We see offsite's growth continuing and accounting for a larger portion of the overall retail media mix as adoption expands to new channels like CTV. Growth in offsite should support publishers, social, search, and AdTech providers.

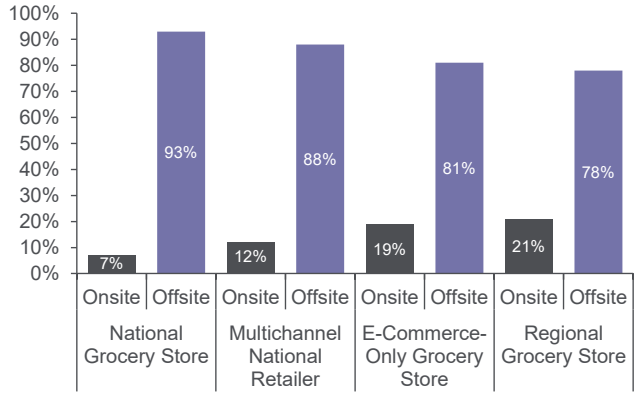
However, offsite should support retailers as well, and, importantly, the broader digital advertising ecosystem. Offsite retail media is so critical because for many retailers, most customers (and particularly, potential customers) are reachable outside their owned properties. Data from Epsilon clients (an agency of Publicis, eMarketer 9/6/22) shows that most customers for key CPG retailers are reachable across the web rather than on its own website.[52]

Figure 50. U.S. Onsite and Offsite Retail Media Spend ($ bns)



Source: Citi Research, GroupM, eMarketer

Figure 51. Brands' Reach to U.S. Customers via Onsite vs. Offsite Ads by Media Type



Note: Onsite represents the percentage of a retailer's customers that visited the retailer's website, and offsite represents the percentage of retailers who didn't but are reachable on the open web.

Source: Citi Research, Epsilon, August 3, 2022, eMarketer

[52] Andrew Lipsman, *The Digital Shelf for CPG Brands 2022*, Insider Intelligence | eMarketer, September 6, 2022.

First-Party Data & Closed-Loop Measurement Advantage

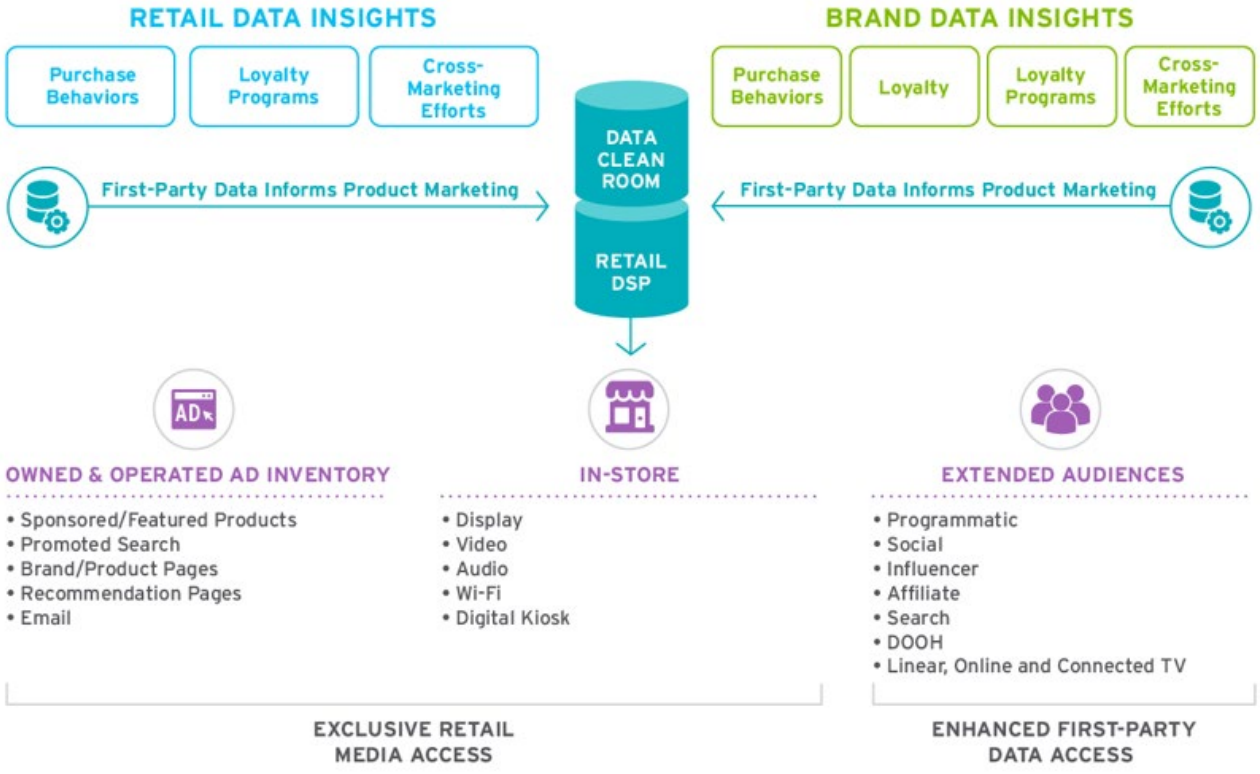
Other critical components of retail media's rising prominence are first-party (1P) data and closed-loop measurement. Specifically, the ability to target and retarget shoppers on social platforms has degraded since Apple's App Tracking Transparency (ATT) initiative made obtaining 3P data more difficult.

Since this initiative, ad innovations have improved certain targeting capabilities and re-invigorated return on ad spend (ROAS) on platforms like Meta. However, the focus on 1P data is continuing to rise as advertisers seek more measurable ad formats, aiming to further leverage their own data (and retailers' data) to connect with audiences. Also critical is the ability to gauge performance and attribution more accurately through closed-loop measurement of sales on a retailer's site or in store.

This ecosystem is allowing brands, advertisers, and retailers the ability to leverage one another's data to better understand their shoppers, which is particularly the case for verticals like CPG that were never able to have their own closed-loop measurement and attribution data because their sales always came through a retailer's site or store. They can effectively now shift more dollars online.

The Trade Desk recently introduced an enhanced data product that takes retail data to build offsite retail media audience targeting.[53] Criteo's new Commerce Max DSP was designed to capitalize on its access to data and provide closed-loop measurement, as it provides advertisers with access to both onsite inventory and can leverage audiences built on real shopper behaviors for offsite targeting ads.

Figure 52. First-Party Data from Retailers and Brands Can Be Incorporated Into Retail Media Campaigns



Source: A Guide to Retail Media, IAB Canada, March 2023

[53] Citi Research, *Kokai Product Launch Displays Best-in-Class Ad Buying Platform*, June 12, 2023.

Smart Money: AI in Finance

Ronit Ghose, CFA
Head of Future of Finance
Citi Global Insights

Sophia Bantanidis
Future of Finance Analyst
Citi Global Insights

Kaiwan Master
Future of Finance Analyst
Citi Global Insights

Ronak S Shah
Future of Finance Team
Citi Global Insights

All general-purpose technologies – from the introduction of steam power and electricity in earlier eras, to AI in our era – have a time lag between their invention and their broader integration into the economy and society.

Business models need to adapt, regulations need to be introduced and/or updated, and human skills need to evolve. The time lag between the invention and industrialization of a technology can be significant. And the finance industry presents the additional challenges of legacy technologies and regulatory hurdles.

Despite the above caveats, we believe generative AI (Gen AI) has revolutionary potential in financial services. In many respects, finance is the perfect sector for the application of Gen AI. The sector is data-rich with words and numbers its raw materials.

For many years, AI – or, strictly speaking, machine learning (ML) – has been used in finance on structured data and for quantitative tasks. Gen AI will expand these use cases to unstructured data. Most enterprise data (about 80-90%) is unstructured – this means the data are locked away in emails, transcripts, documents, and reports.[54]

> Based on conversations with hundreds of clients, the banking and financial services industry is absolutely sprinting right now and is ahead in terms of ideating and understanding use-cases of artificial intelligence in finance and how it can bring value to the firm.
>
> – SANDEEP ALUR, MICROSOFT TECHNOLOGY CENTER, MICROSOFT

Today, AI is used in finance primarily for risk management and pricing, but Gen AI will expand the use cases to information management, search and retrieval, and coding. Software development and implementation will be another important use case. Below are some of the near- to medium-term use cases in the financial services industry in which we expect to see increased adoption of Gen AI:

- **Code Generation & Software:** AI will help enable automation of routine tasks, codification of existing policies/controls, faster generation of code, and improved vulnerability testing.
- **Searching, Summarizing, and Insight Generation:** AI-powered web scraping tools will be increasingly capable of generating research insights, concise summaries, and quarterly reports, for example.
- **Sales and Marketing:** AI will make sales more hyper-personalized through improved customer insights, lead generation from call notes, and campaign optimization.

[54] Tam Harbert, "Tapping the Power of Unstructured Data," MIT Sloan School of Management, February 1, 2021.

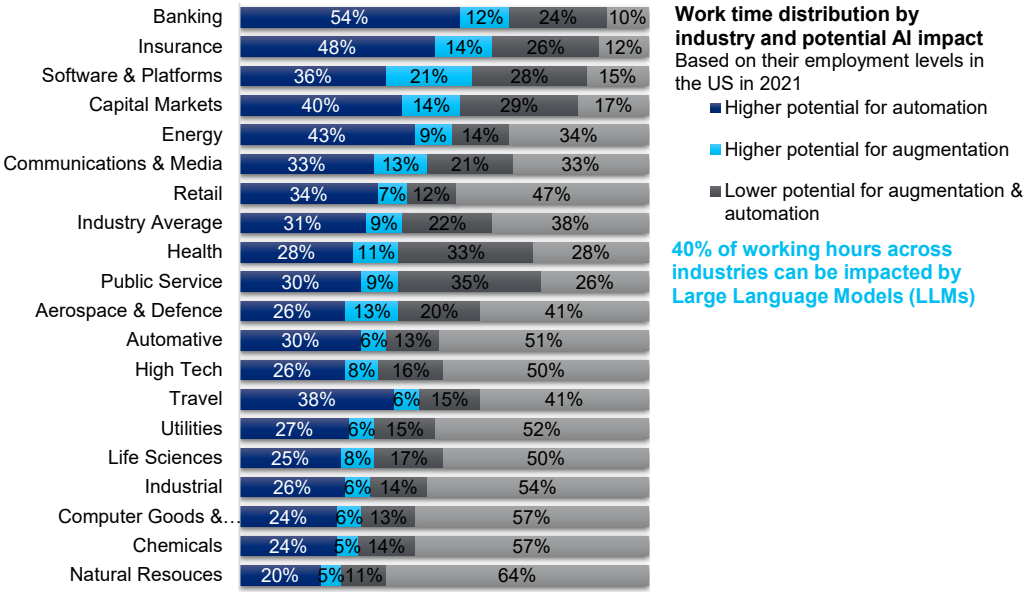
- **Customer Engagement:** AI is poised to bring massive improvements to voice assistants, sentiment and intent analysis, and real-time language translation.
- **Legal Tech:** Tasks in which AI can be used include drafting contracts, conducting risk assessment on documents, detecting internal risks of collusion and insider trading, and performing due diligence.

We believe software development and coding will be one of the key areas of productivity improvement. This holds true for all industries but is particularly relevant for finance given the number of developers and coders who work for large banks and the information-intensive nature of the sector. Based on company disclosures, between 15-25% of large U.S. bank staff are employed in technology roles.

How AI Changes Finance Jobs and Tasks

Jobs and tasks in finance have already seen significant change in recent decades, and they are likely more susceptible to AI-driven change than those in other sectors. According to World Economic Forum analysis, nearly a quarter of all jobs globally will change in the next five years.[55]

Figure 53. Potential for AI-Led Job Displacement



Source: World Economic Forum

Historically, adoption of technological innovations in finance has not led to reductions in the workforce, but rather to the nature of the work changing.

AI will most likely bring productivity gains to both the finance sector and the broader economy by automating and augmenting current tasks and roles. Technology and coding roles in finance will be at the forefront of such productivity gains, as Gen AI will likely lead to 30-50% productivity gains in more basic coding tasks[56]. Banks today rely heavily on tech-driven solutions, boasting a significant number of employees in technology roles (around 15-25% of the staff of big U.S. banks are employed in direct tech roles). The headcount in tech roles has increased in the past few decades as focus increases on digitization and digital transformation.

[55] World Economic Forum, Future of Jobs Report, May 2023
[56] McKinsey, Unleashing developer productivity with generative AI, 27 June 2023

While Gen AI will very likely change individual finance jobs and tasks, as well as how they are staffed and managed, it is notable that the adoption of waves of new technology has not historically led to rapid changes in the finance-sector workforce. Despite the 2010s tech revolution, U.S. financial jobs as a percentage of total employment remained stable at around 5.5%, while U.S. commercial bank employees as a percentage of total finance jobs declined marginally to 16% (2022) from 18% (2011).[57]

Similarly, the introduction of the ATM from the late 1960s onwards did not lead to a drop in the number of human tellers employed by banks. In fact, during the 1970s onwards until the mid-2000s, the number of human tellers employed in physical bank branches exploded as the U.S. economy and financial sector both grew rapidly in size. Over the past 15 years, as the mobile internet revolution went mainstream and banks also faced cost and capital return pressures, manual bank teller roles finally declined.

In the case of Gen AI, both supply- and demand-side factors are worth considering. The amount of human labor required for current roles and tasks will decline due to the AI automation and augmentation discussed above. So, from a supply-side perspective, less human labor will be required. But new roles will emerge, for example in AI governance, model training, and output monitoring. Or, existing human staff, in coding or in other roles, will have more time to focus on "value-added" tasks.

In addition, the amount of output required from Gen AI to maintain the status quo will likely increase. For example, financial service clients, both business and consumer, will likely be AI-enabled and may generate a greater volume of incoming demands and requests. Similarly, regulators and other stakeholders may make increased information and analytical demands from the financial services companies, as these institutions will also be AI-powered in the future. So, as more things change, the more they stay the same.

Risks and Barriers to Adoption of AI in Finance

Regulators will most likely subject firms in the financial services industry to strong scrutiny on their AI plans. For many policymakers, the stability of the financial system and consumer protection will likely outweigh technological innovation in financial services as a goal. In addition, banks' own legacy technology infrastructure and often conservative culture could slow the adoption of Gen AI.

Despite the benefits of AI, these new tools also pose risks to financial ecosystem. AI systems learn to make decisions based on training data, which can incorporate biased human decisions or reflect historical or social inequalities. Biases could also creep in through flawed algorithms or improper implementation. Hallucinations of AI models (in which the models generate incorrect output as if it were true) could have serious business and regulatory implications in finance.

AI models today lack transparency and explicability – a must for adoption in the financial services sector. Moreover, finance is legacy tech-ridden and highly regulated sector where commoditization of technology is the slowest.

AI models also enable cybercriminals to generate illicit code faster. As AI-generated deepfakes become increasingly convincing and ubiquitous, enterprises and

[57] Federal Reserve Bank of St. Louis, All Employees, Commercial Banking (CEU5552211001); U.S. Bureau of Labor Statistics, Employment by major industry sector. 6 September 2023

individuals are also likely to face heightened privacy, cybersecurity, and identity theft risks. Financial institutions own vast volumes of private and sensitive data, and securing these against data hacks and leaks will be crucial.

Another known constraint with AI models is the lack of transparency in the creation of many foundational models.[58] Less transparency hurts everyone – it makes it harder for users to know if they can safely rely on commercial foundation models and build applications on top of them. The highly regulated nature of the finance sector means new AI tools need to be covered by adequate risk management, transparency, and explicability criteria before deployment.

Banking is also a relatively people- and relationship-driven business, and the role of human advice and trust is often highly valued by clients in episodic transactions. While customers may be comfortable primarily relying on automation tools for basic transactional services – such as peer-to-peer (P2P) or bill payments – they often prefer human involvement in large personal (eg, mortgage, wealth and investment management) or corporate (IPOs, mergers and acquisitions) finance transactions.

Who Will Benefit?

We believe there will be alpha generation from Gen AI: Faster-moving incumbents stand to gain, as do innovative new entrants or challenger firms. As with the introduction of many new technologies over the past couple of decades, incumbent firms that empower their staff to execute technological change and transformation are likely to gain market share and/or boost profit margins at the expense of slower competitors.

As well as magnifying the gap between dynamic incumbents and conservative laggards, Gen AI could enable new entrants and scale up challenger firms in finance and other parts of the economy to gain market share.

The productivity gains in coding, content generation, and client communication, among other areas, should enable small firms to scale faster. FinTechs can use Gen AI somewhat similarly to cloud computing as a scale-up accelerator.

Systemically important banks looking to leverage Gen AI tools will likely face heightened regulatory scrutiny; we therefore believe Gen AI adoption in complex and highly regulated financial institutions is likely to be slow. We expect traditional incumbents to potentially lag in adoption. By contrast, smaller companies (eg, FinTechs) and countries that implement relatively lighter or clearer regulations are likely to see faster adoption of AI in finance and allied sectors.

The main beneficiaries will likely be firms with a comprehensive AI strategy and a clear future target operating model that considers how to attract right talent, solve for legacy technology (before overlaying AI tools), and deal with unstructured data sets across the organization (before putting AI to work). We expect Gen AI to increase the scope for new entrants to make inroads into the financial services sector of 2025-2030.

[58] Katharine Miller, “Introducing the Foundation Model Transparency Index,” Stanford University Human-Centered Artificial Intelligence, October 18, 2023.

White and Gold Hydrogen

Maggie Lin
Commodities
Citi Research

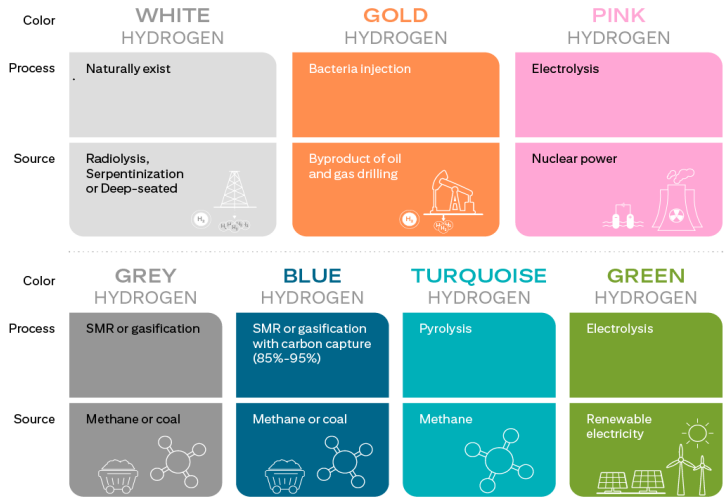
The hydrogen "rainbow" – the collection of color-related terms used to refer to different types of hydrogen fuel based on the different types of technology used to produce each – is expanding. The most prominent types of hydrogen fuel are:[59]

- **"Gray" hydrogen**, which is derived from natural gas through steam methane reforming (SMR), with CO_2 emitted in the process.
- **"Blue" hydrogen**, which is similar to gray hydrogen, but with the resulting CO_2 captured after production via caron capture, utilization, and storage (CCUS).
- **"Green" hydrogen**, which is derived carbon-neutrally via water electrolysis powered by renewable energy.

Reducing the costs of blue and green hydrogen has been a formidable task involving a lot of learning by doing and hoping economies of scale to emerge. There have also been efforts by various governments to subsidize blue and green hydrogen, with some subsidies effectively bringing the cost to zero or below.[59]

"White" hydrogen has always, in theory, been part of the rainbow – it is defined as hydrogen that occurs "naturally," trapped below the earth's surface in underground deposits. Now, some of what has previously been called "white hydrogen" is being relabeled as "gold hydrogen", which is as clean as white hydrogen but not exactly natural (it is an artifact of historical production as well as of fracking).

Figure 54. Hydrogen Rainbow



Source: Citi GPS, World Economic Forum

To this day, the National Grid website says of white hydrogen: "There are no strategies to exploit this hydrogen at present."[60] In fact, finding ways to exploit white and gold hydrogen has become an increasingly hot subject, with the number of citations on the topic exploding in 2023. Part of oil and gas companies' interest in the topic pre-dates efforts to mine white and gold hydrogen. Rather, it is a logical part of the businesses they are in, particularly refining, where upgrading requires adding hydrogen molecules to heavy oil products.

[59] Citi GPS, *Hydrogen: A Reality Check on the Hydrogen Craze*, August 2, 2023.
[60] National Grid, "The Hydrogen Colour Spectrum," accessed December 19, 2023.

These companies also face pressure to substitute gray hydrogen (ie, hydrogen created from natural gas) with blue hydrogen (ie, gray hydrogen produced using carbon capture and sequestration) or green hydrogen (ie, hydrogen produced by the electrolysis of water). Of course, all this varies based on local clean energy sources and government subsidies.[61]

The motivations to exploit underground deposits of clean hydrogen are clear. No energy is required to manufacture the hydrogen, unlike green hydrogen from water electrolysis or blue hydrogen from combining SMR processes with CCUS. For the economics to work, the main considerations are quantity and location.

What's more, given that clean hydrogen residue is a byproduct of fracking, the clean exploitation of the remnants of this process could provide the basis of a potentially lasting business that would be a natural target of the shale industry as the world moves off of oil.

Scoping Out White Hydrogen and Finding Ways to Exploit It

As pointed out in a mid-summer 2023 article in the UK's *The Telegraph*, underground hydrogen has a long history, although investigation of how much is present has historically been limited. Two instances have involved the burning of gas at Chimaera in Turkey and at Los Fuegos Eternos in the Philippines.[62]

The first major effort to scope out the extent of natural white hydrogen was published in April 2023 by the U.S. Geological Survey.[63] Without ample physical data, the scientists responsible for the study had to develop a model based on an understanding of how hydrogen might have been formed in the subsurface and the "behavior" of the hydrogen.

While they concluded that the volume of hydrogen available in the subsurface could satisfy global demand "for thousands of years," they also were quick to point out that most of the white hydrogen "is probably inaccessible." As the article stated, "in other words, hydrogen supplies are too deeply buried, or too far offshore, or in accumulations that are too small." However, the scientists also concluded that "the amount of hydrogen in the Earth's interior could potentially constitute a primary energy source."

Nonetheless, the scientific understanding of the many ways hydrogen is formed has improved tremendously, as well as of how it might be consumed, how it can enter porous rocks, and how it might be sealed. This is critically important, as it has historically been assumed that hydrogen molecules were too light and small to be readily sealed. However, double molecules of hydrogen, common in practice, are about the same size as helium atoms, which have been sealed and preserved for long periods (around 100 million years), encouraging the search to go on. Modelling is indicating favorable conditions for hydrogen accumulations in the eastern and central parts of the U.S., including the Upper Midwest and the Great Plains.

What's more, water reduction separating hydrogen and oxygen molecules looks likely to be a common factor facilitating continuous hydrogen generation in certain

[61] Noah Brenner, "Majors Start Local, Think Global in Hydrogen," Energy Intelligence, August 15, 2023.
[62] Ambrose Evans-Pritchard, "Limitless 'White' Hydrogen Under Our Feet May Soon Shatter All Energy Assumptions," *The Telegraph*, July 13, 2023.
[63] United States Geological Survey (USGS), "The Potential for Geologic Hydrogen for Next-Generation Energy," April 13, 2023.

formations. This will feature not only in the United States Geological Survey (USGS)'s efforts to continue to map out its own country's white hydrogen potential, but in also the global effort. With initial support from the U.S. Department of Energy (DoE), exploration activity will soon begin, and then the question will be whether this sets off a frenzy similar to the shale revolution.

The Drilling Has Actually Begun…

Just before the USGS published its initial report on the potential for white hydrogen development in the U.S., the journal *Science* published another influential report on white hydrogen's potential.[64] That report covered the history of an above-ground hydrogen gas explosion from a lit cigarette in 2007 in a well drilled in 1987 in Mali that had been abandoned.

A wealthy owner of a local oil and gas company bought the rights to the area and, through studies, discovered that the gas leak was 98% hydrogen. The main study, published in 2018, triggered a global trend of drilling for white hydrogen, including in Brazil, Australia, and even the U.S. (in Nebraska). Most of these experiments have found that the hydrogen is a renewable source of energy.

Indications are that private sector-financed "wildcatting" (ie, exploratory drilling) efforts are occurring around the world, including in Europe and the U.S., where recent government subsidies are aiding the growing popularity of hydrogen experiments. The main attraction is the widely held view that pumping hydrogen from underground accumulations should be significantly less expensive than using clean energy to fuel electrolyzers.

…Yet It May Be Much More Difficult than Drilling for Oil and Gas

In geochemistry, conventional wisdom suggests that hydrogen rarely exists naturally due to its reactive nature.[65] Therefore, the idea of searching for natural hydrogen accumulation in large quantities underground was long neglected despite occasional discoveries of it, such as the well drilled in Mali in 1987. Hydrogen drilling is still so overlooked that hydrogen is typically used as a carrier material to figure out the composition of gaseous samples from oil and gas drilling, making it impossible to detect hydrogen in these samples.[66]

Today, unlike oil and gas drilling, drilling for white hydrogen is still in its initial developmental stages, as exploration techniques and methodologies are still being developed and tested. Strategies for the continuous monitoring of drillholes and mines for hydrogen are likely necessary, as natural hydrogen fluxes can be variable and cyclic enough to undermine the accuracy of single grab sampling. The reactivity of hydrogen can also cause perturbed chemical signatures of hydrogen samples through interaction with casing, storage vessels, and fluid-rock interactions, affecting measurement quality.[67]

---

[64] Eric Hand, "Hidden Hydrogen: Does Earth Hold Vast Stores of a Renewable, Carbon-Free Fuel?", *Science*, February 16, 2023.
[65] U.S. Department of Energy, "Hydrogen and Fuel Cell Technology Basics," accessed December 19, 2023.
[66] David Fickling "Natural Hydrogen Could Change the World, If We Understood It," *Japan Times*, August 1, 2023.
[67] Betina Bendall, *Current Perspectives on Natural Hydrogen: A Synopsis*, South Australia State Energy Resources Division, Department of Energy and Mining, July 2022.

Despite the existence of various theories about how underground hydrogen is generated, there is no consensus yet as to how much hydrogen comes from each proposed mechanism, which is another impediment to hydrogen exploration. "Serpentinization," which involves the oxidation of ferrous minerals, has long been considered one of the main hydrogen-generating processes.

However, it generally occurs at mid-ocean ridges or in the Earth's mantle, and hydrogen extraction is economically unviable in those settings given the water depth. The good news is that scientists have found that hydrogen can also be produced through similar oxidation reactions but within onshore ironstone formations, raising the likelihood of more economically available natural hydrogen.[68]

"White" Hydrogen Is Sometimes "Gold," but Not Always

The term "gold hydrogen" has two origins. One stems from the idea that natural hydrogen is worth its weight in gold, given how inexpensive it ought to be to deliver to markets compared to energy-intensive green hydrogen. The other is its production from depleted oil wells or as an unintended consequence of fracking in shale formations.[69]

As explained in a story published in both *Science* and *Wired*, a Texas biotech company experimented with sending a preparation of bacteria and some specific nutrients down the bore of an abandoned well. The result was the generation of hydrogen and carbon dioxide, which could be sequestered.

The firm claimed that the overall cost was ultimately $1/kg, without any government support or subsidies – the same cost, more-or-less, as gray hydrogen created by steam methane reformation. It also invoked the term "gold hydrogen" as a reference to "black gold," once the proud name used for oil when it was thought of as a scarce resource not so long ago. The firm also claims that the CO_2 can be readily sequestered or re-used, something that appears as an open question.[70]

What mainly makes hydrogen exploration attractive to oil companies is the sheer number of abandoned wells, either on company books or available from state authorities. These wells are surrounded by storage and pipeline infrastructure.

The very idea that hydrogen can be produced residually after oil production is completed provides a new rationale for producing clean energy from fossil fuel wells. There are many techniques under review and experimentation, beyond biotech and nutrients. Radiolysis is a process triggered by radioactive rocks including granite, which has a similar impact on separating hydrogen from other molecules. Still another strategy is injecting oxygen that initiates chemical processes that release hydrogen and other products. In Australia, a company aptly named "Gold Hydrogen in Australia" has identified a potential 1.3 billion kg of hydrogen in the Ramsay Peninsula and Kangaroo Island in Australia.

When it comes to hydrogen and fracking, companies have two different potential strategies. One is to re-enter the fracked wells after they have stopped producing, as discussed above. Another is to frack natural gas wells using CO_2 instead of water. According to a 2020 paper on this subject, carbon dioxide could be extracted from natural gas and water via SMR, pressurized significantly, and used in place of

[68] Anna Demming, "The Hunt for Natural Hydrogen Reserves," Chemistry World, August 21, 2023.
[69] Chris Baraniuk, "'Gold Hydrogen' Is an Untapped Resource in Depleted Oil Wells," *Wired*, November 24, 2022.
[70] Cemvita, "Homepage," accessed December 19, 2023.

water for additional fracturing; hydrogen would then remain, with carbon returning into the ground.[71]

Long Shot or Near-Term Revolution?

One conclusion is certain – a lot of attention has been paid recently to white and gold hydrogen, and it appears to be rising.

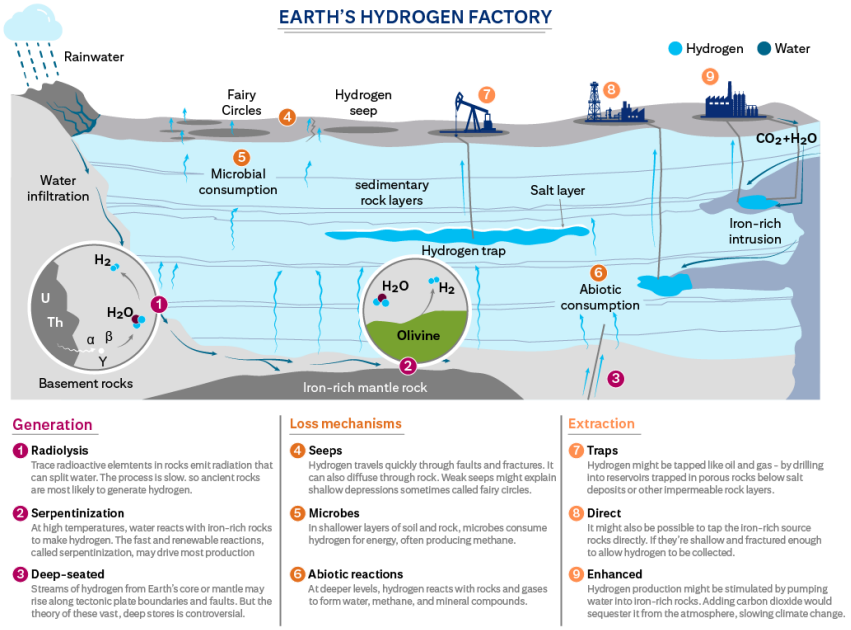
At the same time, largely private-capital efforts are underway to figure out how to implement certain proposed projects (which involve both taking advantage of known underground hydrogen deposits and experimenting with ways to extract hydrogen from older wells and fracked wells).

On top of that, the major oil and gas companies are also increasingly committing themselves to hydrogen development projects. Undoubtedly, it is wise for these companies to search for a profitable afterlife for themselves once oil demand starts falling more dramatically, as well as to find ways to use their existing infrastructure to expand into clean energy technologies.

It remains to be seen whether the hydrogen revolution via white or gold hydrogen will be as transformational as the shale revolution.

We might find out a lot sooner than we think.

Figure 55. How Hydrogen Forms Underground



Source: Science

[71] John W. Andrews, "Hydrogen Production and Carbon Sequestration by Steam Methane Reforming and Fracking with Carbon Dioxide," *International Journal of Hydrogen Energy*, Vol. 45, No. 16, March 20, 2020.

Further Reading on Innovation

Disruptive Innovations IX
Ten More Things to Stop and Think About
December 2022

Disruptive Innovations VIII
Ten More Things to Stop and Think About
October 2021

Disruptive Innovations VII
Ten More Things to Stop and Think About
February 2020

Disruptive Innovations VI
Ten More Things to Stop and Think About
August 2018

Disruptive Innovations V
Ten More Things to Stop and Think About
November 2017

Disruptive Innovations IV
Ten More Things to Stop and Think About
July 2016

Disruptive Innovations III
Ten More Things to Stop and Think About
July 2015

Disruptive Innovations II
Ten More Things to Stop and Think About
May 2014

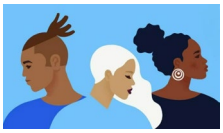
Disruptive Innovations I
Ten More Things to Stop and Think About
April 2013

Mapping the Innovation Landscape
Comparing 100 Emerging & Disruptive Technologies, Business Models, and Ideas

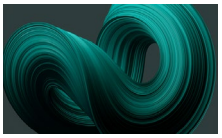
As our premier thought leadership product, **Citi Global Perspectives & Solutions (Citi GPS)** is designed to help readers navigate the most demanding challenges and greatest opportunities of the 21st century. We access the best elements of our global conversation with senior Citi professionals, academics, and corporate leaders to anticipate themes and trends in today's fast-changing and interconnected world.

All Citi GPS reports are available on our website www.citigroup.com/global/insights/citigps

[Future of Healthcare](#)
The Oxford Prescription
May 2024



[Closing the Racial Wealth Gap](#)
Restating the Economic Case
April 2024



[Supply Chain Financing](#)
Building Resilience
January 2024

[Unlocking Climate and Development Finance](#)
Creating Bankable Projects
December 2023

[Food, Water, and Climate Change](#)
Solutions for Tackling this Critical Nexus
November 2023

[APEC Economic Outlook](#)
Navigating Shifting Dynamics
November 2023

[Unlocking Climate and Development Finance Preview](#)
Creating Bankable Projects
November 2023

[Supply Chains](#)
Lessons Learned Helped by New Technology
November 2023

[Food Security](#)
Solutions for a Perfect Storm
November 2023

[Treasury Leadership](#)
Does It Matter?
November 2023

[Philanthropy and the Global Economy v3.0](#)
Perspectives on the Future of Giving
October 2023

[Supply Chain Technology](#)
Improving Efficiency, Resilience, and Transparency
October 2023

[Technology at Work v7.0](#)
The Third Phase of Globalization
October 2023

[Africa](#)
A New Growth Model
October 2023

[Skills That Pay](#)
The Returns from Specific Skills
October 2023

[Future of Cross-Border Payments](#)
Who Will Be Moving $250 Trillion in Next Five Years?
September 2023

If you are visually impaired and would like to speak to a Citi representative regarding the details of the graphics in this document, please call USA 1-888-800-5008 (TTY: 711), from outside the US +1-210-677-3788

IMPORTANT DISCLOSURES

This communication has been prepared by Citigroup Global Markets Inc. and is distributed by or through its locally authorised affiliates (collectively, the "Firm") [E6GYB6412478]. This communication is not intended to constitute "research" as that term is defined by applicable regulations. Unless otherwise indicated, any reference to a research report or research recommendation is not intended to represent the whole report and is not in itself considered a recommendation or research report. The views expressed by each author herein are his/ her personal views and do not necessarily reflect the views of his/ her employer or any affiliated entity or the other authors, may differ from the views of other personnel at such entities, and may change without notice.

You should assume the following: The Firm may be the issuer of, or may trade as principal in, the financial instruments referred to in this communication or other related financial instruments. The author of this communication may have discussed the information contained herein with others within the Firm and the author and such other Firm personnel may have already acted on the basis of this information (including by trading for the Firm's proprietary accounts or communicating the information contained herein to other customers of the Firm). The Firm performs or seeks to perform investment banking and other services for the issuer of any such financial instruments. The Firm, the Firm's personnel (including those with whom the author may have consulted in the preparation of this communication), and other customers of the Firm may be long or short the financial instruments referred to herein, may have acquired such positions at prices and market conditions that are no longer available, and may have interests different or adverse to your interests.

This communication is provided for information and discussion purposes only. It does not constitute an offer or solicitation to purchase or sell any financial instruments. The information contained in this communication is based on generally available information and, although obtained from sources believed by the Firm to be reliable, its accuracy and completeness is not guaranteed. Certain personnel or business areas of the Firm may have access to or have acquired material non-public information that may have an impact (positive or negative) on the information contained herein, but that is not available to or known by the author of this communication.

The Firm shall have no liability to the user or to third parties, for the quality, accuracy, timeliness, continued availability or completeness of the data nor for any special, direct, indirect, incidental or consequential loss or damage which may be sustained because of the use of the information in this communication or otherwise arising in connection with this communication, provided that this exclusion of liability shall not exclude or limit any liability under any law or regulation applicable to the Firm that may not be excluded or restricted.

The provision of information is not based on your individual circumstances and should not be relied upon as an assessment of suitability for you of a particular product or transaction. Even if we possess information as to your objectives in relation to any transaction, series of transactions or trading strategy, this will not be deemed sufficient for any assessment of suitability for you of any transaction, series of transactions or trading strategy.

The Firm is not acting as your advisor, fiduciary or agent and is not managing your account. The information herein does not constitute investment advice and the Firm makes no recommendation as to the suitability of any of the products or transactions mentioned. Any trading or investment decisions you take are in reliance on your own analysis and judgment and/or that of your advisors and not in reliance on us. Therefore, prior to entering into any transaction, you should determine, without reliance on the Firm, the economic risks or merits, as well as the legal, tax and accounting characteristics and consequences of the transaction and that you are able to assume these risks.

Financial instruments denominated in a foreign currency are subject to exchange rate fluctuations, which may have an adverse effect on the price or value of an investment in such products. Investments in financial instruments carry significant risk, including the possible loss of the principal amount invested. Investors should obtain advice from their own tax, financial, legal and other advisors, and only make investment decisions on the basis of the investor's own objectives, experience and resources.

This communication is not intended to forecast or predict future events. Past performance is not a guarantee or indication of future results. Any prices provided herein (other than those that are identified as being historical) are indicative only and do not represent firm quotes as to either price or size. You should contact your local representative directly if you are interested in buying or selling any financial instrument, or pursuing any trading strategy, mentioned herein. No liability is accepted by the Firm for any loss (whether direct, indirect or consequential) that may arise from any use of the information contained herein or derived herefrom.

Although the Firm is affiliated with Citibank, N.A. (together with its subsidiaries and branches worldwide, "Citibank"), you should be aware that none of the other financial instruments mentioned in this communication (unless expressly stated otherwise) are (i) insured by the Federal Deposit Insurance Corporation or any other governmental authority, or (ii) deposits or other obligations of, or guaranteed by, Citibank or any other insured depository institution. This communication contains data compilations, writings and information that are proprietary to the Firm and protected under copyright and other intellectual property laws, and may not be redistributed or otherwise transmitted by you to any other person for any purpose.

IRS Circular 230 Disclosure: Citi and its employees are not in the business of providing, and do not provide, tax or legal advice to any taxpayer outside of Citi. Any statements in this Communication to tax matters were not intended or written to be used, and cannot be used or relied upon, by any taxpayer for the purpose of avoiding tax penalties. Any such taxpayer should seek advice based on the taxpayer's particular circumstances from an independent tax advisor.