

AI Transformation Brief

Featuring AI best practices and insights to enable our members to strategize, plan, develop, deploy, manage, and govern AI-based technologies and solutions.

In This Issue → [AI in the News](#) | [AI Research Highlights](#) | [Vendor Spotlight](#) | [Upcoming Events & Resources](#)



AI IN THE NEWS

Google introduces Gemini [Read News Article](#)

On February 8, Google released Gemini Ultra, the Gemini version they used for their "demo" last December (the demo took a few liberties and gave users an impression of capabilities that were more tricks with the camera). Gemini Ultra is the version positioned as their "ChatGPT killer." A rebranding will also occur; Gemini will replace the Bard and Duet brands (for code generation).

ANALYST ANALYSIS

It should come as a surprise to no one that benchmarks cited by Google demonstrate that Gemini outperforms ChatGPT when it comes to reasoning in the areas of text, image, video, and speech. Gemini also appears to deliver better response times and provides image generation capabilities at no charge. For these reasons, I see no reason to use ChatGPT as a primary LLM.

Going forward, the real test will be comparing Gemini against GPT-4. While current benchmarks show marginally better performance in favor of Gemini, we can expect continued improvements for both models soon.

Finance worker pays out \$25 million after video call with deepfake 'chief financial officer' [Read News Article](#)

A Hong Kong branch of a UK multinational company lost \$25 million in a deepfake phishing scheme. An employee received a seemingly legitimate email from the CFO instructing the person to make secret transfers. Despite initial doubts, the employee became convinced to proceed after a Zoom call featuring a deepfake of the CFO and other employees. The only real person on the Zoom call was the targeted employee, the videos of the CFO and all other call participants were modified from existing publicly available videos.

ANALYST ANALYSIS

This incident highlights the dangers of deepfake technology and the need for improved internal security measures. The use of deepfake technology has become more common as a way to defraud individuals and organizations. In 2021, there were reports that deepfake audio technology had been used to fraudulently instruct a bank director to pay out \$35 million as part of an acquisition ([Forbes, 2021](#)).

Organizations need to continue educating employees on the latest phishing schemes and preventative actions that can be taken. For sensitive meetings, verify participants' identities through a separate channel like email or text before proceeding. For critical encounters involving finances, use secure messaging apps like Signal for additional verification and confirmation.

AI IN THE NEWS

NPHardEval Leaderboard: Unveiling the reasoning abilities of large language models through complexity classes and dynamic updates

[Read News Article](#)

The NPHardEval leaderboard serves as a testing ground for large language models (LLMs), specifically measuring their reasoning abilities. This new benchmark introduces a diverse set of 900 challenges, comprising nine categories, each with 100 questions with 10 difficulty levels. To keep the assessment fresh and prevent models from simply memorizing solutions, new challenges are added monthly.

ANALYST ANALYSIS

Complex reasoning is essential for LLMs to provide value for complex decision-making. Critics of existing benchmarks claim they fall short in differentiating the various LLMs because they are static and public tests risk overfitting, allowing LLMs to exploit specific metrics and artificially inflate performance. Thus, the introduction of a new benchmark that attempts to deliver more rigorous and dynamic assessments to improve how to measure LLM reasoning capabilities for real-world tasks. Expect more such benchmarks to follow.

Hugging Face launches open-source AI assistant maker to rival OpenAI's custom GPTs

[Read News Article](#)

During the first week of February, Hugging Face announced a free chatbot builder, taking direct aim at OpenAI's custom GPTs. This free offering, the Hugging Chat Assistant, could significantly disrupt the AI chatbot market by democratizing access to this productivity tool.

ANALYST ANALYSIS

Within a month of the official launch of OpenAI's GPT Store, a platform that enabled users to find and leverage a custom version of ChatGPT, Hugging Face has announced an alternative.

The Hugging Chat Assistant allows users to select from several open-source LLMs to power their AI assistant. In addition, a central repository reminiscent of OpenAI's GPT Store has been created for Hugging Chat Assistants developed by third parties.

The arrival of the Hugging Chat Assistants demonstrates the agility and responsiveness of the open-source community and its ability to close the current gaps with its proprietary rivals like OpenAI. Many are expecting the performance of open-source LLMs to soon rival the performance of GPT-4, Gemini, and other proprietary models.

Elon Musk's Neuralink brain chip: what scientists think of first human trial

[Read News Article](#)

On January 29, 2024, Neuralink founder Elon Musk posted that his company was successful in implanting its brain-computer interface (BCI) device into a human for the first time.

ANALYST ANALYSIS

The Neuralink chip, a coin-sized device, was recently implanted in a human for the first time. The device aims to revolutionize brain-computer interfaces (BCIs), by enabling a person to control a computer, cellphone, robotic arm, or wheelchair by thought alone. While the potential of the Neuralink device to restore movement and communication is nothing short of transformative, several hurdles require attention before widespread adoption. This emerging technology raises concerns about safety, ethical implications, and the long-term impact on the human brain. Rigorous research and open dialogue are crucial to addressing these concerns and ensuring responsible AI development before people begin integrating this technology directly into their brains.

DEFINING AI USAGE POLICIES *New!*

We identified a policy framework for AI governance

The following policy areas will be supported by new policies and updates to existing policies.



Source: [Govern the Use of AI Responsibly With a Fit-for-Purpose Structure](#)
 Info-Tech Research Group

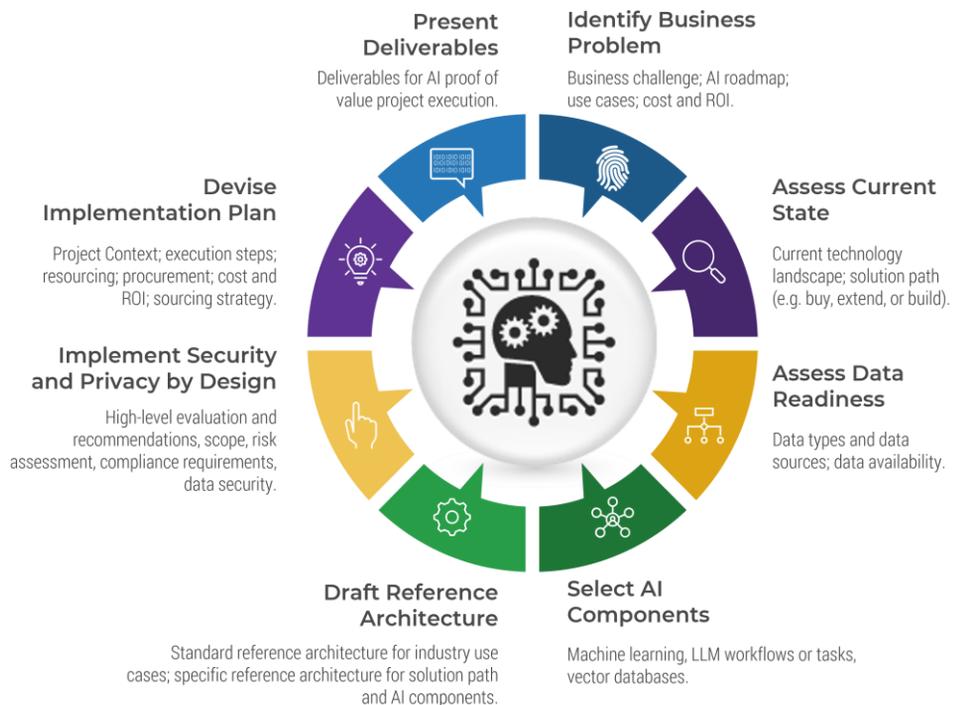
LAUNCH YOUR AI PROOF OF VALUE *New!*

Launch your AI proof of value

"We need to launch an AI proof of value to assess the viability of the technology before we invest."

Challenges

- Leaders need to show AI initiatives can work as intended and will yield tangible business benefits.
- Assessments face significant time constraints.
- If the project team takes too long to secure support, it could lead to funding delays or missed opportunities.



Adopt a formal methodology to accelerate AI vendor selection and improve outcomes.

Results

- Identify relevant capabilities to support the AI use cases you identified.
- Validate AI use cases against business requirements.
- Identify vendors to support AI use cases, devise a vendor selection model, and select products for review.

Source: [Launch Your AI Proof of Value](#)
 Info-Tech Research Group

VENDOR SPOTLIGHT – FEATURING THE AI MARKETPLACE



Custom AI Vendor Selection

Share your preferences for AI vendors and specific use cases, and let us guide you to our selection of top-notch, verified AI partners. Elevate your projects with the expertise of our dedicated analysts who are ready to assist you in making informed decisions for impactful AI integration.

[AI Marketplace | Info-Tech Research Group](#) → New!

Case Studies



AI Use Case Research



The AI Marketplace is where we house our vendor interviews and case studies for leading off-the-shelf solutions. These include AI solutions for functional areas like Application Development, IT security, and lead generation as well as vertical segments like retail, legal and many more.

UPCOMING AND RECENT EVENTS

[Info-Tech Leadership Summit](#): Tampa Bay, FL – February 21-23, 2024 →

[Info-Tech Leadership Summit](#): Nashville, TN - April 3-5, 2024 →

[CIO Priorities 2024 Webinar](#): February 21, 2024. February 23, 2024 (APAC) →

[Security Priorities 2024 Webinar](#): February 23, 2024 →

[Decoding Crime: AI Use Cases in Policing Webinar](#): March 5, 2024 →

[Unify IT in Higher Education Webinar](#): March 5, 2024 →

[Balancing Innovation and Privacy: Protecting PII Webinar](#): March 5, 2024 →

[Plan Your IT Unification Across Districts for K-12 Webinar](#): March 6, 2024 →

[Leverage Insurtech to Enhance Capabilities and Drive Returns](#): March 7, 2024 →

[Build an AI Risk Profile that Supports Sustainable Transformation in Federal Government Webinar](#): March 7, 2024 →

(and many more events...)

AI AND DATA ANALYTICS SOLUTIONS – RESOURCES

[AI Marketplace | Info-Tech Research Group](#) → New

[Artificial Intelligence Research Center](#)

Workshops

- AI Strategy Workshop
- AI Proof of Value Workshop
- AI MLOps Workshop
- AI Governance Workshop

AI EDITOR-IN-CHIEF

Bill Wong – Info-Tech AI Research Fellow <https://www.linkedin.com/in/bill-wong-the-AI-guy/>