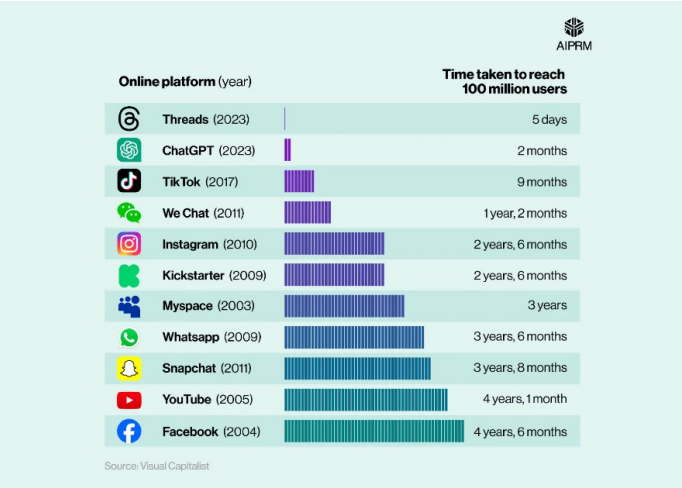
NTT DATA

Designed with the power of **GenAI**

GenAI unveiled: understanding the complexities and opportunities

GenAI is set to redefine the nature of work, becoming part of our daily lives. However, it's important to recognize that the technology is still in its early stages. We share three common pitfalls that could undermine the efforts of even the boldest pioneers seeking the success and innovation that GenAI promises – and how to address them.

The hype is real



ChatGPT reached 100 million users in just two months (Facebook: 4.5 years; TikTok: 9 months).[1]

Late in 2022, "GPT" and "ChatGPT" became buzzwords in the technology and business fields, representing the cutting edge of generative artificial intelligence (GenA), this technology had existed in the background for a little over five years[2] before it captured widespread attention and became one of the most rapidly adopted innovations in recent history.

> "In 20 years following the Internet space, we cannot recall a faster ramp in a consumer internet app.[3]

GenAI has reached a global audience, setting off a wave of applications that have sparked creativity and innovation in unprecedented ways. From creating art and melodies to producing text that closely resembles human writing, GenAI is reshaping our interaction with technology and opening a new dimension of possibilities.

With its ability to mimic human conversations and assist with critical decision-making, has significant potential to disrupt businesses and industries.

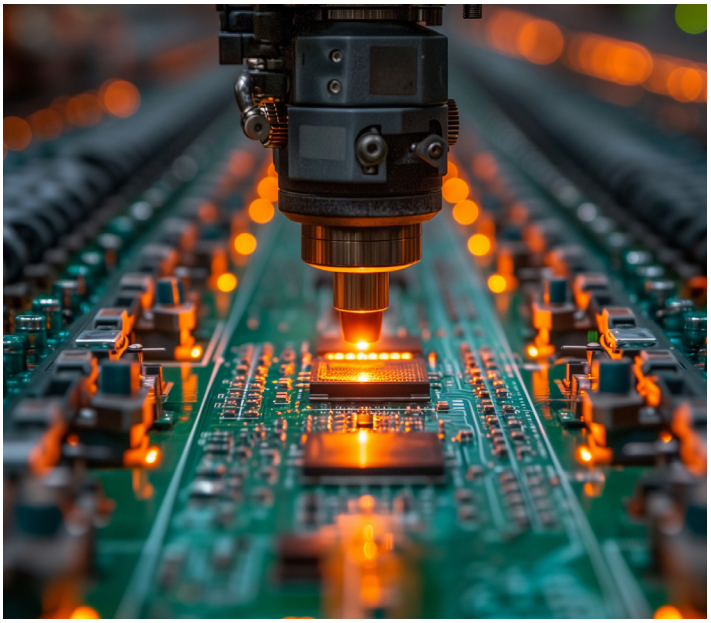
As we enter an era where humans and intelligent machines work together to achieve more than what either could alone, GenAI has set to redefine the nature of work, becoming part of our daily lives, often operating invisibly in the background.

However, it's crucial to recognize that GenAI is still in its early stages. In this guide, we share three common pitfalls that could undermine the efforts of even the boldest pioneers in the world of GenAI.

> "Generative AI (GenAI) is over the Peak of Inflated Expectations as business focus continues to shift from excitement around foundation models to use cases that drive ROI.[4]

[1]AIPRM. 100+ ChatGPT Statistics 2024. | Accessed on 9 September 2024.
[2]NTT DATA Corporation. Languages, models and business applications | Accessed on 9 September 2024.
[3]USB analysts, quoted in: Krystal Hu, Reuters (2 February 2023). ChatGPT sets record for fastest-growing user base – analyst note. | Accessed on 9 September 2024.
[4]Gartner (17 June 2024). Hype Cycle for Artificial Intelligence. | Accessed on 9 September 2024.

The evolution of GenAI



The evolution of GenAI has been nothing short of remarkable in recent years. The quality and capabilities of these AI models have undergone a profound transformation from early basic experiments to real-world applications. From healthcare and data analysis to content creation and entertainment, there are hardly any fields or sectors where GenAI has not left its mark.

With continuous advancements in machine learning and deep neural networks, GenAI promises to redefine how we interact with technology.

GenAI: 50 years in the making

1952
Christopher Strachey, a British computer scientist, writes a love-letter program for the Manchester Mark 1 computer, the first text-generating software.

1968
Terry Winograd from MIT develops SHRDLU, a natural language processing system that can understand and respond to commands in a <block world> environment.

1985
Judea Pearl, a computer scientist and philosopher, introduces Bayesian network and lays the groundwork for modeling complex systems in generative AI.

1996
Joseph Weizenbaum, an MIT professor, develops ELIZA, a computer program that simulates a conversation with a psychotherapist.

2006
Stanford University creates the ImageNet project – a large-scale visual database used for training computer vision algorithms.

2014
Ian J Goodfellow and his colleagues introduce generative adversarial networks (GANs) in a paper titled "Generative Adversarial Networks".

2018
OpenAI release GPT, a language model that generates coherent text based on a given prompt.

A new network architecture, Transformer[5], is introduced, based on attention mechanisms.

2018
The launch of BERT, which helped computers understand ambiguous language in text. Later, NVIDIA introduces StyleGAN, a GAN architecture capable of generating highly realistic images.

2020
OpenAI releases GPT-3, the largest language model at the time, capable of generating impressive human-like text across various domains.

2022
OpenAI launches an AI chatbot called ChatGPT, which quickly attracts more than 100 million users.

As of September 2024, ChatGPT has over 200 million weekly active users.

2023
Significant advancements are made in multimodal GenAI systems, which can process and generate outputs based on multiple types of data inputs, such as text, images and audio.

2024
The European Artificial Intelligence Act (AI Act), which aims to foster responsible artificial intelligence development and deployment in the EU, enters into force.

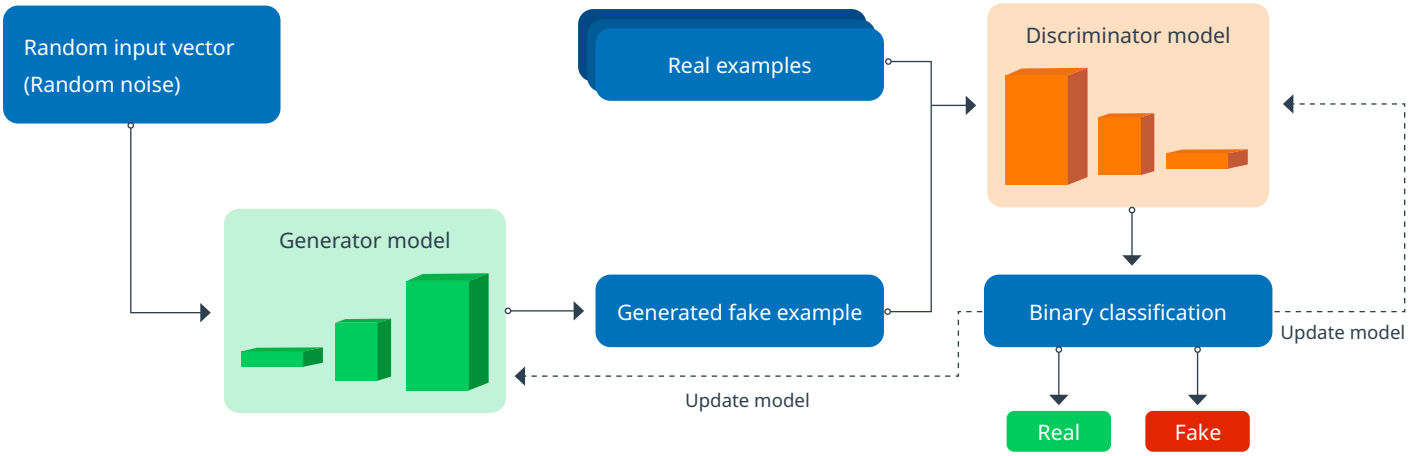
[5]Ashish Vaswani et al. (2 August 2023). Attention Is All You Need. | Accessed on 9 September 2024.

Demystifying GenAI

GenAI, at its core, is a subfield of artificial intelligence that focuses on creating algorithms and models capable of generating data that resembles human-created content. These algorithms use neural networks, with approaches like variational autoencoders (VAEs) for structured data, generative adversarial networks (GANs) for highly realistic content generation, and recurrent neural networks (RNNs) or transformers for sequential data, such as text or speech. The training process involves exposing these models to extensive datasets, allowing them to learn patterns and correlations within the data.

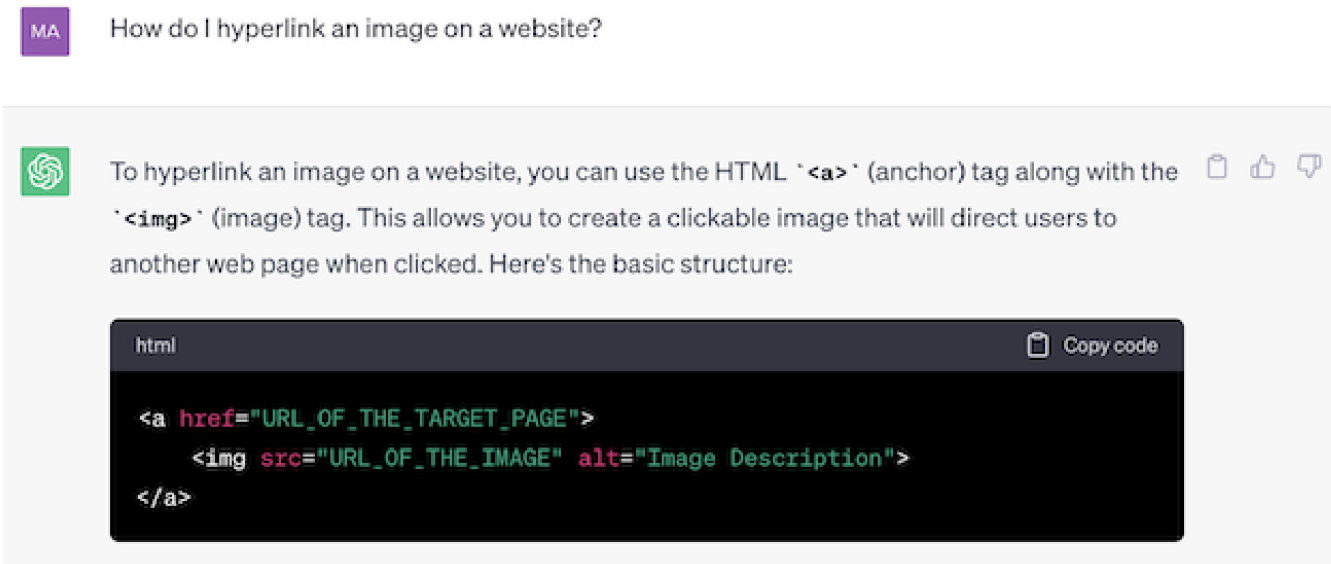
Mode collapse in GANs and ensuring high-quality training data are some of the crucial considerations in the development and ethical deployment of GenAI. GenAI finds applications in diverse fields, from natural language generation for chatbots and content creation to image synthesis for art and design. By demystifying GenAI and understanding its underlying technical principles, businesses and researchers can use it to automate content generation, improve creative processes and innovate in a wide range of areas while remaining vigilant about ethical considerations and data quality.

Generative adversarial network architecture



GenAI models use VAEs as a “recipe book” to understand the essence of various data types, blending them to create new, unique content. GANs act like a food critic (discriminator) that gives feedback on whether the results are good or bad. This ongoing interplay helps the models refine their techniques, resulting in the creation of lifelike and convincing outputs.

Where, how and why?



A simple prompt can generate code. Source: OpenAI

Key areas where GenAI is reshaping tasks and opening opportunities

Conversations and customer interactions

Picture having GenAI as your assistant in customer interactions. This smart, cognitive bot understands your customers' intent and crafts responses that elevate the quality of their experience. It's so advanced that the conversations it generates are nearly indistinguishable from human interaction. In healthcare, for example, GenAI can help medical professionals by providing real-time patient data, proposing treatment alternatives and responding to medical questions in easy-to-understand language. These exchanges can improve patient understanding, patient care and efficiency in healthcare settings.

Content creation

GenAI brings in revolutionary capabilities to empower editors, writers and content creators. By tapping into real-time global trends and knowledge, it assists creators in generating engaging content and has the potential to transform how content is created across various domains.

Rather than displacing knowledge workers and content writers, GenAI must be leveraged to enhance and boost their capabilities. It can generate text and draft articles, emails and even poems.

This capability extends to visual content and music, offering a modern and creative approach to artistic expression. DALL·E, for instance, can generate unique and visually stunning digital artworks, while Midjourney helps artists generate new ideas, refine their craft and create more impactful works.

> “ NTT DATA offers specialized services for public administrators and financial sector clients, helping them understand complex regulations and financial topics, respectively.

Knowledge management

Knowledge management had its limitations in handling extensive datasets and extracting meaningful insights from data. Traditional methods relied on manual curation and keyword-based search, often resulting in gaps in information retrieval.

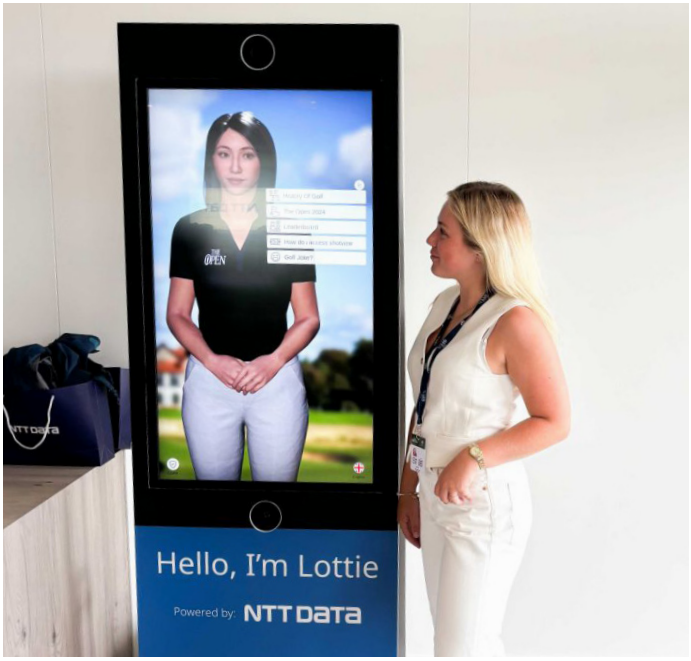
With the ability to handle large datasets, GenAI can play a pivotal role in knowledge management and help us obtain and use the right information effectively. Transformer-based models can grasp contextual nuances and quickly summarize lengthy content.

This technology helps improve search-engine capabilities and provides more accurate, context-aware results.

Content discovery

GenAI models improve content discovery by offering a nuanced and personalized experience across various platforms.

GenAI can shape content discovery through personalized recommendations, providing tailored content based on user behavior. With advanced natural language processing, GenAI refines content understanding, improving categorization and tagging for more efficient search algorithms.



NTT DATA's "digital human", Lottie, was showcased at the Open Championship 2023, which took place at Royal Liverpool.

> NTT DATA partnered with L'Oréal to create Lore, a virtual assistant based on our conversational AI platform to reimagine CX and boost online sales.

Process automation

With extensive contextual knowledge and historical data, GenAI can streamline and speed up customer interactions by automatically analyzing processes that can be automated, reducing manual effort.

NTT DATA's process automation solutions include:

- Dolffia, a multicloud GenAI solution that optimizes document processing, improves content creation and delivers data-driven insights for tailored and scalable industry-specific and cross-functional use cases.
- LITRON®, is a document-comprehension AI that can read and comprehend Japanese texts quickly and accurately. It features the LITRON® Generative Assistant, which combines a large language model (LLM) with LITRON® Generative Assistant technology to search internal documents and provide responses in a chat format. This advanced technology empowers organizations to maximize the value of their Japanese-language documents.

Software development and maintenance

GenAI is proficient in several popular programming languages and frameworks and can seamlessly convert between them. It can also comprehend human-written problem statements and instantly generate the required code, keeping pace with the latest advancements and security considerations.

GitHub Copilot is an example of GenAI used in coding applications. This programming assistant serves as an intelligent code-completion tool. It assists developers by understanding the context and generating code in real time, which significantly increases productivity.

4 steps to integrating GenAI in organizations

Incorporating GenAI into enterprise operations is much more than a technical undertaking. It's a strategic step which needs precision, alignment with organizational goals, and ongoing monitoring and refinement.

Step 1: Define clear business objectives and assess impactful use cases

Begin by articulating well-defined business goals that align with the overarching organizational vision. This will help you to focus on specific areas where GenAI can be applied effectively.

Thoroughly evaluate potential use cases for GenAI, taking into consideration factors such as ease of implementation, projected return on investment (ROI) and how these use cases align with established business objectives. Consider compliance with industry regulations and data privacy laws at every step.

This ensures that the adoption of GenAI is not just a technological advancement but a strategic move that significantly impacts organizational outcomes.

Step 2: Project discovery and precision planning

This step involves identifying technical challenges (distinct from broader business issues) and solutions. These include model choices, fine-tuning strategies and integration with external knowledge bases as well as the appropriate technology stack, encompassing cloud services, frameworks, libraries and vector databases.

It is equally important to establish both technical and nontechnical success metrics and conduct a comprehensive cost assessment.

Prompt engineering plays a vital role in achieving efficient AI systems and must be actively monitored to reduce unnecessary computations and optimize resource utilization.

Step 3: Proof of concept and pilot implementation

Before you implement a large-scale deployment, it is essential to conduct a meticulous proof of concept on a small scale, in a controlled environment, to validate the feasibility of your approach. This allows you to identify challenges early and refine your strategy as needed.

A user-friendly interface integrated with existing systems will help in obtaining feedback for ongoing improvement, fine-tuning and optimization.

Step 4: Full implementation – realizing the vision; optimization and ongoing maintenance

Based on the validated pilot and inputs received in the previous steps, we can now execute a full-scale deployment of GenAI.

This phase involves both the technological implementation and strategically incorporating the refined model into the organizational fabric.

To keep the refining the tool in order to meet changing organizational needs, a cycle of perpetual optimization, maintenance and governance is required after implementation.

Define a reference model and a standardized enterprise architecture or framework that provides a blueprint for your desired outcomes. This will help you meet critical success factors when going from proof of concept to production while enabling scalability and performance and monitoring security, privacy and governance.

Avoid the common pitfalls of GenAI implementation

Despite the numerous advantages GenAI offers, many businesses are hesitant to implement it fully. One primary reason for this is the complexity of GenAI systems. There have been blackbox issues with AI – an inability for us to see how deep-learning systems make their decisions. In addition, concerns regarding scalability, compatibility with existing systems, and the need for specialized talent can lead to hesitancy.

Ethical considerations are at the forefront of these reservations. Ensuring transparency, accountability, data privacy and the potential for harmonious human–AI collaboration are paramount.

Loss of human expertise
Over-reliance on AI can lead to a decline in human skills and expertise, as well as a potential loss of creativity and critical thinking.

Miscommunication
Unclear roles and expectations in human–AI collaboration can lead to misunderstandings and suboptimal outcomes.

Workplace disruption
Poorly integrated AI systems can disrupt established workflows and require significant adaptation, leading to resistance amongst employees.

It's crucial to address the issues that can be inadvertently overlooked during GenAI implementation.

Data privacy and security

Data privacy and security are paramount in GenAI implementation, as AI models often require access to sensitive information. Failing to prioritize data privacy and security can lead to serious repercussions.

- Implement robust data encryption, access controls and data anonymization.

To comply with regulatory frameworks such as the General Data Protection Regulation (GDPR) and California Consumer Privacy Act (CCPA), it is essential to establish who has control over the data (data sovereignty) and where the data is stored (data residency), which adds an additional layer of complexity.

- Conduct regular security audits to ensure compliance with relevant data protection laws.

Human–AI collaboration

Efficient collaboration between humans and GenAI is essential for productive and effective workflows. Neglecting to strike the right balance can lead to inefficiencies and frustration.

- Carefully design workflows that incorporate GenAI as a supportive tool rather than a replacement.
- Train and upskill employees in AI-related skills to foster more effective human–AI collaboration.

Measuring efficiency and return on AI investment

Although GenAI promises unprecedented productivity, it is essential to measure the efficiency and impact of GenAI.

Calculating your return on AI investment (ROAI) helps you understand the economic value generated by GenAI and assess the long-term benefits of your AI investment.

- Establish key performance indicators (KPIs) and validate them against business goals to determine your ROAI.

Proof of concept to production: sample selection and framework

The transition from proof of concept to production requires an appropriate sample for testing critical applications under GenAI solutions.

- Choose samples that would be relevant in real-world scenarios.
- Adopt the right framework for integrating GenAI into existing IT systems after the transition.

Understanding the limitations of GenAI

GenAI models are susceptible to generating content that may appear real but lacks grounding in reality.

GenAI models can perpetuate biases, leading to discriminatory or unfair outputs.

GenAI systems are based on statistical features. However, they lack the ability for true human reasoning.

Having a small context window of a few thousand words means the model will struggle with comprehensive context awareness.

While GenAI holds immense promise, it is not without its limitations. These limitations remind us of the need for discernment and careful consideration.

Hallucinations

Much like dreams can take unexpected and nonsensical turns, GenAI models are susceptible to generating content that is not grounded in reality. These AI "hallucinations" can result in outputs that, while creative, are factually incorrect or illogical.

Human validation and context awareness are required to interpret GenAI content.

Bias

GenAI learns from vast datasets, which may inadvertently contain biases present in society. Consequently, AI models can perpetuate these biases, leading to discriminatory or unfair outputs.

Vigilant oversight and an understanding of ethical considerations are critical to recognizing and mitigating bias in the development and deployment of GenAI systems.

Lack of human reasoning

While GenAI excels at pattern recognition and content generation, it lacks the ability for true human reasoning. It cannot comprehend concepts, make judgments based on ethical or moral principles, or engage in complex decision-making as humans do.

Human supervision and guidance are necessary when using GenAI in critical applications.

Limited context overview

A retrieval augmented generation (RAG) approach can parse relevant information from a large corpus of text to generate high-quality text that is both coherent and diverse. This leads to more accurate and informative responses, making GenAI ideal for applications such as text summarization, machine translation, chatbots and content creation.

However, we must bear in mind that GenAI operates within the constraints of the data it has been trained on. It may struggle with providing comprehensive context awareness, especially in situations where the information required extends beyond its training data. This can result in responses that lack nuance and fail to consider broader contextual factors.

Understanding and navigating these limitations is essential for using GenAI effectively. By acknowledging these boundaries, organizations can make informed decisions about the use of GenAI and ensure that its applications align with both ethical and practical considerations.

Regulation – risks and complexities

The regulation of GenAI presents a complex challenge. On the one hand, it can mitigate risks like harmful content generation and deepfake manipulation. On the other, excessive regulation may stifle innovation and hinder the development of beneficial applications. Striking the right balance is crucial to harness the potential of GenAI while safeguarding against potential misuse. Finding a middle ground that encourages responsible research and application is essential in navigating this risk.

Economic impact of GenAI[6]

- GenAI's impact on productivity could add trillions of dollars in value to the global economy.
- About 75% of the value that GenAI use cases could deliver falls across four areas: customer operations, marketing and sales, software engineering, and R&D.
- GenAI will have a significant impact on all industry sectors and could deliver value equal to an additional $200 billion to $340 billion annually.
- GenAI has the potential to change the way we work, augmenting the capabilities of individual workers by automating some of their individual activities.
- GenAI can substantially increase labor productivity across the economy, but this will require investments to support workers as they shift work activities or change jobs.

[6]McKinsey (14 June 2023). The economic potential of generative AI: The next productivity frontier. | Accessed on 10 September 2024.

Generative AI at NTT DATA

NTT DATA has been conducting research and developing services and products utilizing technologies focused on natural language processing and groundbreaking models such as GPT since 2018. We have published several white papers about LLMs (including GPT-2 and GPT-3)[7].

We are continuing our journey of development with foundational Japanese-language models, which laid the groundwork for our continued innovation. NTT's proprietary Japanese LLM, tsuzumi, has lightweight yet world-class Japanese-language processing capabilities, significantly reducing the costs required for training, inference and tuning.

Our vision is firmly rooted in the belief that GenAI, particularly LLMs, will continue to shape the future of the AI landscape. Our long-standing commitment to natural language processing has uniquely positioned us to leverage our technological expertise and meet the growing demand for GenAI.

To make our AI offerings scalable and efficient and reduce adoption time, we follow an asset-based consulting (ABC) approach. This enables us to offer the best scalable services and fast implementation to our clients in combination with our top talent in various offerings:

Intelligent document processing and knowledge management

We meet client needs in areas such as intelligent document processing (IDP), knowledge management and process automation.

Conversational AI

Our solutions leverage the code-generation capabilities of foundational models to automate several tasks in the end-to-end development and maintenance of software.

Code generation and transformation

Leveraging GenAI, we have efficiently modernized legacy apps, improving their performance, scalability and maintainability.

Intelligent managed services

We leverage GenAI to automate enterprise outcomes, revolutionizing IT environment management through streamlined resource provisioning, workflow optimization and infrastructure efficiency.

Our mission is to bolster the practical applications of language-based AI by enhancing its fundamental attributes with efficiency, reliability, ethical standards and customizability. We are committed to bridging the gap between language-based AI and the real world, and aspire to create AI systems that go beyond computational algorithms to produce real-world outcomes. In empowering AI to enact positive change in our world, we can realize our vision for a future where technology seamlessly augments our everyday lives.

Advantages of our solutions

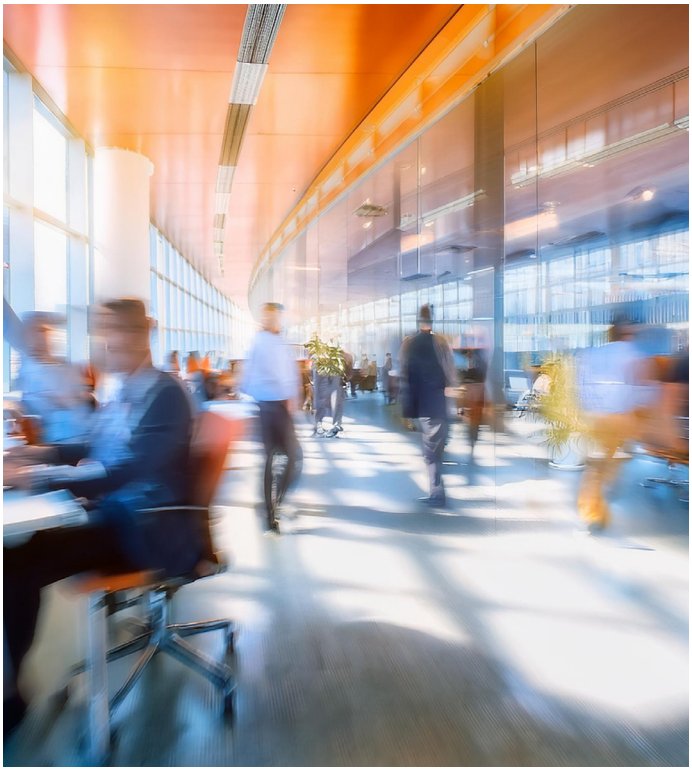
Our assets are reusable.

We address common challenges and pain points.

Our solutions are aligned with customer feedback.

Clients realize value from their investment faster.

We offer a clear development roadmap.



[7]NTT DATA Corporation. Languages, models and business applications. | Accessed on 9 September 2024.

The bigger picture of leveraging GenAI

With its potential to reshape industries and unlock new possibilities, GenAI is a remarkable tool. However, it is not the only solution in our toolkit. We see it as a versatile instrument that complements the portfolio of solutions we offer to clients to help them address complex challenges.

Before implementing GenAI, it's essential to conduct a cost-benefit analysis. GenAI's power comes at a price, in terms of both graphics processing unit (GPU) processing and electricity consumption. It's important to assess whether the ROI justifies its use, particularly when a task may be more efficiently handled by human expertise or other AI techniques. Infrastructure considerations, such as the need for substantial computing resources, must be weighed against the potential benefits.

Therefore, a smart tweak would be employing smaller LLMs to strike a balance between performance and practicality. This cost-effective way of enhancing efficiency and transparency makes these models attractive alternatives to larger LLMs for a wide range of applications.

For instance, a media company using GenAI to create content may discover they can produce a high volume of articles efficiently. However, if not carefully managed, the cost of GPU clusters and electricity may erode the potential cost savings of this approach.

In this scenario, adopting smaller LLMs would be more cost-effective, efficient and easier to deploy.

Conversely, a financial institution using GenAI for data analysis may find that the speed and accuracy of AI-driven insights justify the computational expenses.

> NTT's tsuzumi is available in two versions: a lightweight version, with 7 billion parameters, and an ultra-lightweight version, with 600 million parameters. With top-level Japanese-language processing capabilities, it is approximately 300 times lighter than OpenAI's GPT-3 and is claimed to surpass GPT-3.5 and top domestic LLMs as per Rakuda, a benchmark for GenAI.[8]

[8]NTT. (25 April 2024). NTT's Large Language Models 'tsuzumi'. | Accessed on 10 September 2024.

GenAI, with its creative abilities and data-driven insights, can undoubtedly drive innovation and efficiency. However, it should be viewed as a valuable addition to our portfolio of solutions, rather than a one-size-fits-all remedy.

All technology, including GenAI, is a means to an end, not the end itself. The true measure of success lies in using these tools effectively to solve real-world problems, while being mindful of ethical considerations and the cost-effectiveness of our choices. By embracing the change thoughtfully and holistically, we can navigate the evolving landscape of AI with confidence and innovate for a brighter future.

Let's get started

See what NTT DATA can do for you

- Deep industry expertise and market-leading technologies
- Tailored capabilities with your objectives in mind
- Partnerships to help you build and realize your vision

All images in this report were generated using Azure Open AI, Adobe Firefly and Microsoft Co-Pilot.


NTT DATA