How Generative AI Will Transform Businesses

Nomura Research Institute, Ltd.
Center for Strategic Management & Innovation

Takashi Sagimori, Expert Researcher
Yoshiaki Nagaya, Expert Strategist

October 2023
The Adoption of Gen AI Varies Across Industries in Japan

While Gen AI’s adoption rate is still low, uptake is rapidly growing. Industries like IT and communications, education, construction and transportation are leading the way.

The online questionnaire surveyed 2,421 full-time professionals (ages 20 to 69) living in Japan equally distributed across sex and age groups.

Source: "Questionnaire Survey on AI Adoption" (NRI, May 2023)
## In-house Gen AI adoption among private-sector companies in Japan

<table>
<thead>
<tr>
<th>Company</th>
<th>Business</th>
<th>Gen AI Adoption</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daiwa Securities</td>
<td>Financial</td>
<td>Document creation</td>
</tr>
<tr>
<td>Hitachi</td>
<td>Electronics</td>
<td>Launched “Generative AI Center”</td>
</tr>
<tr>
<td>KDDI</td>
<td>IT/Communication</td>
<td>Research and document creation</td>
</tr>
<tr>
<td>Meiji Yasuda Life</td>
<td>Insurance</td>
<td>ChatGPT PoC for in-house operations</td>
</tr>
<tr>
<td>Mercari</td>
<td>E-commerce</td>
<td>Dedicated team to study Gen AI and large language models</td>
</tr>
<tr>
<td>Mitsui Sumitomo Insurance</td>
<td>Insurance</td>
<td>PoC to streamline in-house inquiries and those from insurance agencies</td>
</tr>
<tr>
<td>Mitsubishi Research Institute</td>
<td>Research, Consulting</td>
<td>ChatGPT for in-house report creation</td>
</tr>
<tr>
<td>Panasonic</td>
<td>Electronics</td>
<td>Using ChatGPT via intranet</td>
</tr>
<tr>
<td>Softbank</td>
<td>IT/Communication</td>
<td>Launched new company for utilizing Gen AI</td>
</tr>
</tbody>
</table>
AI adoption beyond in-house use among Japanese companies

- Dentsu: Innovating ad creative production processes using AI
- Hakuhodo: Introduced automatic ad copy generation, AI image service “H-AI NARRATIVE”
- Kokuyo: Using Gen AI in office designs
- Nomura Securities: Aims to provide easy-to-understand information to individual investors
- Recruit: Incorporated travel reservation site “Azure OpenAI Service”

“The Ultimate Office Space to be unveiled by Kokuyo in 2030”
Companies are also using their own Gen AI to offer new services

**Amadeus Code Inc.**
- Creates original songs
- Applies natural language processing technology to music
- Suggests melody lines to music creators, with aim to also enhance productivity

**Plug Inc.**
- Uses Gen AI to suggest product package designs
- Includes feature for scoring customer design satisfaction
- Saves consumer goods makers time in their product development operations

**OPT Inc.**
- Automatically generates and evaluates ad content
- Combines in-house Gen AI with GPT-4 to create highly effective advertising content
- Reported time savings of 75% for creating ads

Companies are also using their own Gen AI to offer new services.
Gen AI is expected to see the greatest use in replacing mundane tasks like creating meeting minutes to also supporting creative work like music, video, and artwork generation.

Source: “Questionnaire Survey on AI Adoption” (NRI, May 2023)
Gen AI challenges and threats

Ethical issues like data privacy and brand integrity are emerging as the biggest concerns in the evolution of Gen AI.

Existing Challenges

- Data biases
- Black boxes (AI accountability)
- System verification difficulties

New Threats

- Copyrights
- Hallucinations (AI illusions)
- Deep fakes
Gen AI challenge #1: Copyrights

Data
- Training data collected from the web may contain copyrighted material
- It’s difficult to distinguish between Gen AI output and copyrighted work

Model
- There’s risk that models used are restricted by licences
Gen AI challenge #2: Hallucinations

Chatbots are still evolving and often give inaccurate responses.

Example:

NRI was founded April 1, 1965. A quick search in ChatGPT in May 2023 uncovered the wrong answer. The response was corrected by August 2023.

|   |  
|---|---
| 🧐 Tell me the date that NRI was founded. |
| 💲 NRI (Nomura Research Institute, Ltd.) was founded on December 22, 1965. |
| 🤔 December 22 is not right, is it? |
| 💲 I’m sorry, I provided the wrong information. NRI (Nomura Research Institute, Ltd.) was founded on November 16, 1959. As you said, it was not December 22. I apologize for the error. |
Gen AI challenge #3: Deep fakes

The spread of false information by malicious actors is becoming a major problem as Gen AI becomes more accessible to the public.

A fake video posted on Facebook in 2022 shows Ukraine’s President Zelenskyy calling Ukrainians to surrender to Russia.

A fake video showing a submerged district in Shizuoka City circulated on social media.
Solutions to combat deep fakes are starting to emerge

Pindrop and NABLAS can now identify impersonated audio “spoofs” created through the misuse of Gen AI helping prevent phone scams and other crimes.

Pindrop Security Inc.

NABLAS
Using Gen AI as a brainstorming partner

Aria, Head of Operations at Zinly asks ChatGPT for help to envision the opening of a new café at her company:

I’ve been thinking about opening a new café for staff. It needs to be remote-work friendly.

I envision the café’s interior to be made entirely of renewable materials, use energy efficient electronics, and keep waste to a minimum, which I think would encourage employees to use it.

Are there any ways this idea can be improved?

ChatGPT says…

That’s an excellent idea! Initiatives that meet today’s needs like environmental awareness and providing remote work environments are quite fascinating. Below, I’ve given some ways to further improve your idea and some points to consider.

1. Noise measures:
   - It’s important to have a calm environment when doing remote work. Outside noise can cause problems, particularly when you’re having a conference call or video call. Consider taking appropriate noise prevention measures, or providing quiet individual booths.

2. High-speed internet connections:
   - If you’re doing video calls or other kinds of data-heavy work, a fast and stable internet connection is essential. Think about Wi-Fi quality and bands, the number of connection points, etc.

3. Convenience and accessibility:
   - It’s important to have the café in a location that’s easy for company workers to get to, and that is easily accessible by public transportation.

4. Providing food and drinks:
   - There should be healthy snacks and drinks that people can grab while they’re working, lunch options, etc.
“Data to Text” will support output that can be understood by everyone

Traditionally

Customer data, daily work reports, etc.

Going forward

Data

101011
01101
10110

Analytical experts interpret the results

AI generates text from data

Final insights can be understood by everyone
Statement on Gen AI from the University of Tokyo:

"""
We are in a major period of change, and University members cannot remain on the sidelines; we must instead continue to discover and use generative AI and other new technologies with integrity, respecting legal frameworks and considering the impacts on social/economic systems.
"""

Guidelines on Gen AI from Musashino Art University:

- Always fact check Gen AI content
- Do not input personal information, confidential information, or malicious material
- Submitting work that uses Gen AI responses without citation is considered plagiarism
Short-term: Costs and processing time are higher than conventional AI

Fast development as a result of intense competition has led Gen AI creators to prioritize rapidly increasing scale and performance, resulting in a growing number of technology issues.

Data quality: Datasets are so immense that the content cannot be fully examined.

Please note: legal and other rights-related issues not included.
Current Gen AI systems are difficult to maintain

Language processing systems have practical use, but knowledge is incomplete and under-developed
The drive toward larger-scale systems, specialization and research into power saving is advancing

Power saving solutions and improved practical utility through functional specialization could help reduce the cost of Gen AI development

Example:
In June 2023, Microsoft Research announced phi-1\(^*1\), a new model offering equivalent performance to GPT-3.5 while being smaller in scale

\(^*1\): Has just 1.3 billion parameters, or 1/100\(^{th}\) that of GPT-3
Having AI organize updated information and continually train on a neural network is currently inefficient.

We are already seeing a method being used known as Retrieval Augmented Generation, which involves grounding LLMs on data obtained from an external knowledge base to perform reasoning.
Prospects

Moving toward autonomy to enable AI to “create action plans on its own”

- In November 2022, Google unveiled a framework that combines reasoning and acting called ReAct*1
- When a user sets up a task goal, AI solves the task on its own (AI independently decides actions for achieving the goal)
Prospects

Moving toward autonomy to enable AI to “create action plans on its own”

User

“I want to know the state of competing funds”

Current: Generative AI only

User Generated

Text Generation

Create task plan
1. Obtain sources
2. Analyze text
3. ...

Perform the tasks
1. Execute source obtainment
2. Execute text analysis
3. ...

Check the results

Finished

Pre-supplied tools
Tool A explanation,
Tool B explanation,
Tool C explanation,
Tool D explanation

Create task plan
1. Identify specific tasks based on the goal
2. Find tools helpful for tasks
3. Create action plan

Perform the tasks

Check the results

Going forward: Autonomous AI agent
Hypothesis – 1

With brain development, LLMs will produce a “division”

- LLM architecture design may grow more advanced, such as system and data separation

- Stable diffusion achieved through adjustment of trained model (CLIP*1) and diffusion model

- Where data and system was previously intermixed, they will now be separated

- In the old model, knowledge updates affect the entirety. A system boundary means knowledge updates will have no system-side effects

*1 Contrastive Language-Image Pre-Training
Hypothesis – 2

A new age of knowledge will emerge, with AI spawning new discoveries

New knowledge may be discovered through frameworks such as CLIP, which connects images and text.

In practice
MIT and Harvard researchers published a Nature paper in May 2023 on Geneformer, a model that makes predictions about cell behavior from gene sequences.

If image and text pairs are vectorized and vectors that achieve high cosine similarity can be generated, images and text are considered to have captured the same expression.

Identifies candidate therapeutic targets for cardiomyopathy.
Hypothesis – 3

Foundation models will spread across industries

- Companies will invest in AI foundation models to enhance their employee capabilities
- Post-development, AI will be further customized by departments for deployment

<table>
<thead>
<tr>
<th>Foundation model</th>
<th>Industry specific</th>
<th>Company specific</th>
<th>Department specific</th>
</tr>
</thead>
<tbody>
<tr>
<td>AI trained on a massive amount of data</td>
<td>Retail-oriented responses</td>
<td>Corporate culture responses</td>
<td>Job function responses</td>
</tr>
<tr>
<td>Financial services-oriented responses</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

AI developers / government  System developers / trade groups  User companies and departments
Artificial general intelligence (AGI) will no longer be just a dream

In 2019, Prof. Bengio from the University of Montreal predicted AI's deep learning evolution from System 1 to System 2.
The 4\textsuperscript{th} AI boom and beyond

2023 may well have marked the start of the 4\textsuperscript{th} AI boom and we are only at the start of the next phase in AI evolution.

As Gen AI rapidly becomes more sophisticated, what previously took one year is happening in just one week. Both industry and research are leading this charge.

It is crucial to align on the definition of what Gen AI is and isn’t, as the terminology can be misleading. Gen AI manages language and text; it does not produce interesting images or text.

Language processing systems are already creating tremendous value at the practical level, but they still haven't reached their full capacity. The current models are low-maintainability systems that have data written into their programs. Language processing (the systems) will one day be separated from the data, helping to expand AI system capacity.

Challenges with Gen AI abound but so do opportunities. Businesses must fully understand its characteristics before leveraging AI for their operations.
Envision the value, Empower the change