

A horizontal bar with a gradient from yellow to green to blue, positioned above the main title.

# The introduction to Azure Open AI in Financial Services industry

Connie Leung  
Senior Director,  
Financial Services Business Lead - Asia

Artificial Intelligence

Machine Learning

Deep Learning

Generative AI



## Artificial Intelligence

The field of computer science that seeks to create intelligent machines that can replicate or exceed human intelligence

---



## Machine Learning

Subset of AI that enables machines to learn from existing data and improve upon that data to make decisions or predictions

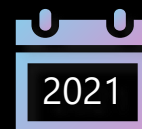
---



## Deep Learning

A machine learning technique in which layers of neural networks are used to process data and make decisions

---



## Generative AI

Create new written, visual, and auditory content given prompts or existing data



*Ensure that artificial  
general intelligence (AGI)  
benefits humanity*



*Empower every person and  
organization on the planet  
to achieve more*

---

# Microsoft Cloud Runs on trust

Your data is your data

Your data from any fine-tuning is not used to train the foundation AI models

Your data is protected by the most comprehensive enterprise compliance and security controls



The next  
revolution in  
computing

# Azure OpenAI

## Top 4 capabilities



### Content generation

**Call center analytics:**  
automatically generate responses to customer inquiries

Generate personalized UI for your website



### Summarization

**Call center analytics:**  
summary of customer support conversation logs

**Subject matter expert document:**  
summarization (e.g., Financial reporting, analyst articles)

Social media trends summarization



### Code generation

Convert natural language to SQL (or vice versa) for telemetry data

Convert natural language to query proprietary data models

Code documentation



### Semantic search

Search reviews for a specific product/service

Information discovery and knowledge mining

## Examples of multiple model use cases

**End to end call center analytics:** classification, sentiment, entity extraction, summarization and email generation

**Customer 360:** hyper-personalisation using timely summarization of customer queries & trends, search, and content generation

**Business process automation:** search through structured & unstructured documentation, generate code to query data models, content generation

# Choosing a use case in Capital Markets

## Use cases to avoid



- As a Generative AI model, it is not designed to process numerical calculations/real-time information or replace sophisticated risk models
- Sending automated AI Advice or generatively developed content directly to a customer without human review
- Avoid use cases that have regulatory impact or oversight
- Examples include
  - Risk modelling
  - Realtime Trading and Transaction Management
  - Reconciliation

## Use cases to evaluate



- Start small and build from there
- Azure OpenAI is a Generative Large Language Model
- Use cases involving text analysis, summarization, forms or content generation are the sweet spot
- Information discovery and knowledge mining

# Emerging OpenAI Deployment Patterns

Start Inside the Organization first..

## Horizon 1



Use Cases  
Inside the  
Organization

Internal Copilot  
deployments

## Horizon 2



Copilot agents that  
directly interact  
with clients

## Horizon 3



New product  
offerings

Human 'in the  
loop' review of  
generative content



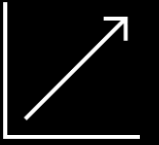
Human Supervision of  
multiple Copilot agents



Increased  
Automation



# Capital Markets use cases

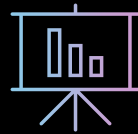


**Client  
engagement**



**Market  
research**

**Investor report  
summarization**



**Pitchbook  
generation**



**Know your  
Product**



**Accessibility  
and  
language  
translation**



# Use cases

## Internal Virtual agents



Copilots to  
augment  
Contact  
Center and  
Agents

### Traditional agent

- Q and A pairs
- Process or decision tree based
- Limited ability to answer ad hoc questions



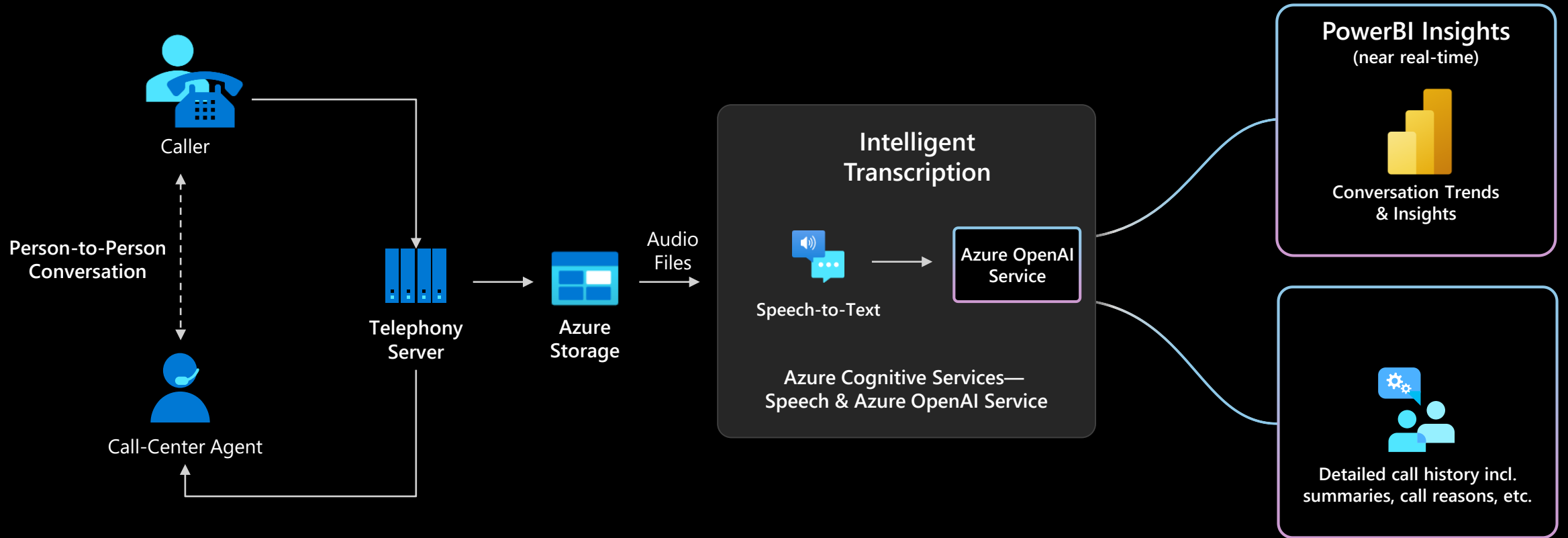
### Azure OpenAI powered agent

- Search structured and unstructured data
- Contextual awareness
- Assist customer in new ways
- Personalization



# Contact center Analytics – Client Engagement

Extract rich insights from call transcripts



# Use cases

## Contact center insights



### Contact Center Insights

**Playground**

Deployments: text-davinci-002 | Examples: Load an example | Code View

**Customer conversation**

Hi there I just had a car accident and wanted to report it. OK, I hope you're alright and in safe place to have this conversation. Yes, I'm fine thank you. Can you please describe to me what happened? I was driving on the M23 and I hit another car. Are you OK? Yeah, I'm just a little shaken up. That's understandable. Can you give me your full name? Sure, it's Sarah Standl. Do you know what caused the accident? I think I might have hit a pothole. OK, where did the accident take place? On the M23 near junction 10. Was anyone else injured? I don't think so. But I'm not sure. OK, well we'll need to do an investigation. Can you give me the other drivers information? Sure, his name is John Radley. And your insurance policy number. OK. Give me a minute. OK, it's 546452. OK, is there any damage to your car? Yes, Headlights are broken and the airbags went off. Are you going to be able to drive it? I don't think so. I'm going to have to have it towed. Well, we'll need to get it inspected. I'll go ahead and order you a tow van. I'll also start the claim and we'll get everything sorted out. Thank you.

**Azure OpenAI instruction**

Extract the following from the conversation:

1. Main reason of the conversation
2. Sentiment of the customer
4. How did the agent handle the conversation?
5. what was the Final outcome of the conversation
6. create a short summary of the conversation

**Output**

1. The main reason of the conversation was to report a car accident.
2. The sentiment of the customer was shaken up but OK.
4. The agent handled the conversation well and was empathetic.
5. The final outcome of the conversation was that the agent started a claim and ordered a tow truck.
6. The customer was in a car accident and the agent started a claim and ordered a tow truck. The customer was shaken up but OK.

# Azure OpenAI Service

## Sample use cases

Call center  
automation  
Example:  
OpenAI prompt

**Playground**

Deployments: text-davinci-002 | Examples: Summarize Text | Code View

**Azure OpenAI instruction in English**

You must extract the following information from the phone conversation below:

1. Call reason (key: reason)
2. Cause of the incident (key: cause)
3. Names of all drivers as an array (key: driver\_names)
4. Insurance number (key: insurance\_number)
5. Accident location (key: location)
6. Car damages as an array (key: damages)
7. A short, yet detailed summary (key: summary)

Make sure fields 1 to 6 are answered very short, e.g. for location just say the location name. Please answer in JSON machine-readable format, using the keys from above. Format the output as JSON object called "results". Pretty print the JSON and make sure that is properly closed at the end.

**Customer conversation in English (or other language)**

Phone conversation:  
Hi I just had a car accident and wanted to report it. OK, I hope you're alright, what happened? I was driving on the I-18 and I hit another car. Are you OK? Yeah, I'm just a little shaken up. That's understandable. Can you give me your full name? Sure, it's Sarah Standl. Do you know what caused the accident? I think I might have hit a pothole. OK, where did the accident take place? On the I-18 freeway. Was anyone else injured? I don't think so. But I'm not sure. OK, well we'll need to do an investigation. Can you give me the other drivers information? Sure, his name is John Radley. And your insurance number. OK. Give me a minute. OK, it's 546452. OK, what type of damages has the car? Headlights are broken and the airbags went off. Are you going to be able to drive it? I don't know. I'm going to have to have it towed. Well, we'll need to get it inspected. I'll go ahead and start the claim and we'll get everything sorted out. Thank you.

**Output as JSON**

```
results = {
  "reason": "car accident",
  "cause": "hit a pothole",
  "driver_names": ["Sarah Standl", "John Radley"],
  "insurance_number": 546452,
  "location": "I-18 freeway",
  "damages": ["headlights broken", "airbags went off"],
  "summary": "Sarah Standl was driving on the I-18 when she hit another car. She thinks she hit a pothole. John Radley was the other driver. There were no injuries but both cars have damage."
}
```

# Azure OpenAI Service

## Sample use cases (cont.)

Example:  
generate  
auto text  
in any  
language

**Playground**

Deployments: text-davinci-002 | Examples: Load an example | Code View

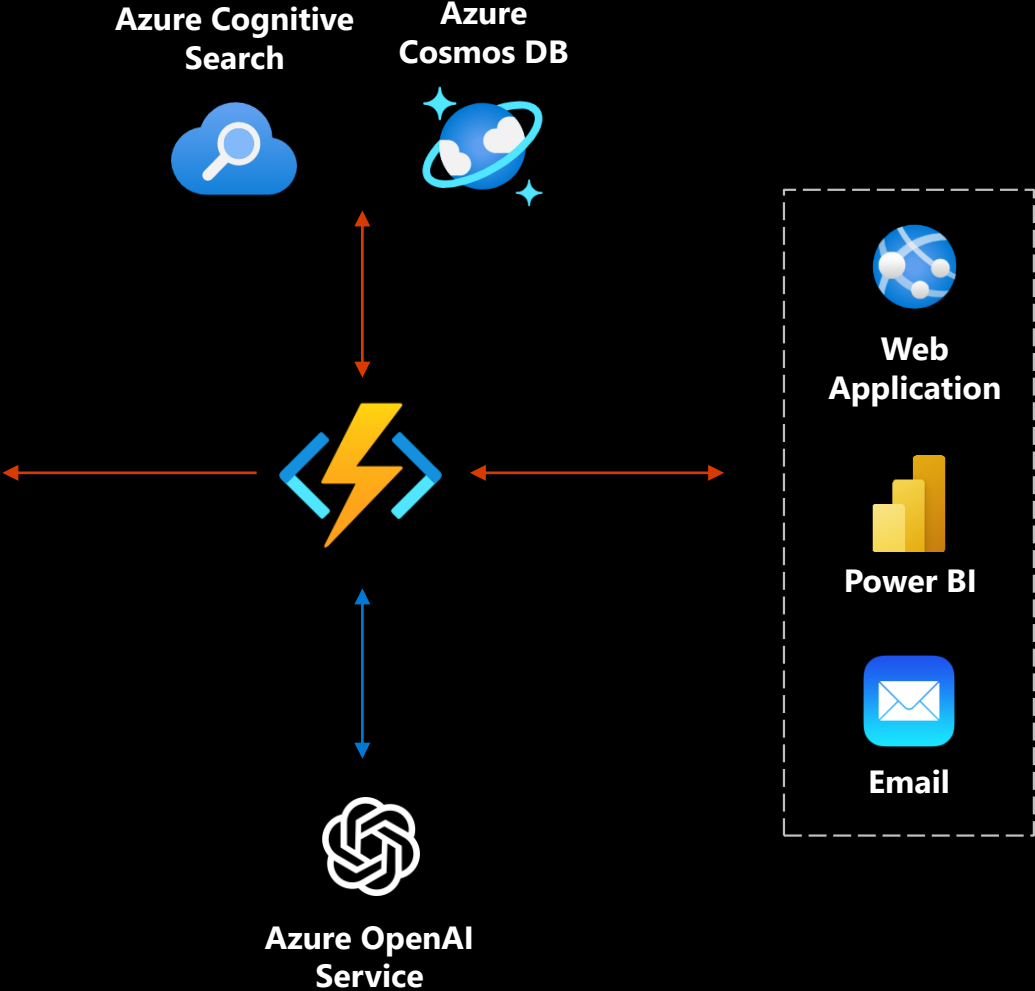
**Input—German**  
Wartungsfenster für den 14. Januar von 14 bis 15 Uhr geplant, erwarten Sie Ausfallzeiten.

**Azure OpenAI instruction—English**  
Generate a formal customer communication in English  
- offer help at info@operations.com  
- be apologetic where possible  
- use 0:00 as time format  
Please output in a letter format.

**Output—English**  
Dear Customer,  
We are writing to inform you that we have scheduled maintenance for January 14th from 14:00 to 15:00. We apologize for any inconvenience this may cause and offer help at info@operations.com should you have any questions.  
Sincerely,  
Your Operations Team

# News & Social listening to the listed stock companies on HKEX

Perform summarization for each company's news & entire sector or market by GPT



# From 12,308 words content to 1,500 words valuable insight summary email

```
In [6]: # Concatenate all the top 10 news record's text content into a single string.
news_text_content_string = news_data_with_text_df.to_string(columns=['Text'], header=False, index=False)
print(news_text_content_string)
```

Science & technology | Generative AI Large, creative AI models will transform lives and labour markets They br  
st of three special articles we explain how they work\n\nImage: George Wylesol\n\nS ince November 2022, when O  
, first opened the chatbot to the public, there has been little else that the tech elite has wanted to talk ab  
e founder of a London technology company messaged your correspondent unprompted to say that this kind of AI is  
hese daysâ€. He says he is in the process of redesigning his company, valued at many hundreds of millions of d  
embodies TSMC , a Taiwanese semiconductor firm that finds itself in the geopolitical crosshairs. GPT -4, the a  
GPT , has aced exams that serve as gateways for people to enter careers in law and medicine in America. It can  
generative AI â€ models can churn out digital photos, drawings and animations. Chatembodies more knowledge tl  
rse cogently about mineral extraction in Papua New Guinea, or about, a Taiwanese semiconductor firm that finds  
the artificial neural network which powers Chat, has aced exams that serve as gateways for people to enter care  
n generate songs, poems and essays. Other â€generativeâ€ models can churn out digital photos, drawings and an  
too quickly. GPT -4 is a type of generative AI called a large language model ( LLM ). Tech giants like Alphabet  
r own LLM s, and given them names like P a LM , Megatron, Titan and Chinchilla. Running alongside this excitem  
stry and beyond, that generativemodels are being developed too quickly.-4 is a type of generativecalled a large  
abet, Amazon and Nvidia have all trained their owns, and given them names like, Megatron, Titan and Chinchilla.  
tech boss says he is â€incredibly nervous about the existential threatâ€ posed by AI , even as he pursues it,  
about it dailyâ€. Governments in America, Europe and China have all started mulling new regulations. Prominent  
artificial intelligence to be paused, lest the software somehow run out of control and damage, or even destroy,  
or excited you should be about this technoloev. it helps first to understand where it came from. how it works :



```
In [7]: # Count the number of words in the top 10 news record's text content.
len(re.findall(r'\w+', news_text_content_string))
```

Out[7]: 12308



[https://github.com/easonlai/google\\_news\\_content\\_scrape\\_and\\_analyze\\_with\\_gpt](https://github.com/easonlai/google_news_content_scrape_and_analyze_with_gpt)

```
In [28]: # Display the final summary from the top 10 news record's text content.
news_content_summary = response["choices"][0]["text"].strip()
print(news_content_summary)
```

The amount of text available on the internet is growing at an exponential rate. A large portion of this text is in the form of news content. As the amount of news content grows, so does the need for automated methods of summarizing this content.


One such method is to use large language models, or LLMs. LLMs are a type of generative AI that are becoming increasingly popular. These models are able to generate text, songs, poems, essays, and even pass exams. However, there is deep concern within the tech industry that these models are being developed too quickly and that they may someday run out of control and damage or destroy human society.


This concern was renewed when OpenAI, a research lab founded by student Jack Hughes, released the ChatGPT chatbot. This chatbot is able to generate responses to questions that are similar to those of a human. The release of this chatbot has caused a lot of concern within the tech industry, with many people calling for a moratorium on the development of LLMs.

Despite these concerns, the development of LLMs is continuing. Hughes has co-founded a new company, Auto-GPT, which is developing a new LLM that is even more powerful than ChatGPT. This new LLM, called Auto-GPT, is able to generate responses to questions without needing to be prompted. This is a significant improvement over ChatGPT, which required the user to prompt it again to keep going.

With Auto-GPT, you can set up the AI to automatically generate the next prompt based on the previous response. This is really powerful because it lets you build up long strings of data. So if you are trying to generate a list of all the restaurants in a city, for example, you can have Auto-GPT automatically generate the next restaurant name based on the last one it gave you.

Auto-GPT is still in beta, so it is not as user-friendly as some of the other options on this list. But it is really powerful and could be a great option for those with the coding skills to use it.

GPT News Summary of Today 

Eason  [bric3.maj@jet.com](mailto:bric3.maj@jet.com) Sun, Apr 23, 1:15 AM (1 day ago) ☆ ↶ ↷

to me

Here is the news summary of GPT for today.

The amount of text available on the internet is growing at an exponential rate. A large portion of this text is in the form of news content. As the amount of news content grows, so does the need for automated methods of summarizing this content. One such method is to use large language models, or LLMs. LLMs are a type of generative AI that are becoming increasingly popular. These models are able to generate text, songs, poems, essays, and even pass exams. However, there is deep concern within the tech industry that these models are being developed too quickly and that they may someday run out of control and damage or destroy human society. This concern was renewed when OpenAI, a research lab founded by student Jack Hughes, released the ChatGPT chatbot. This chatbot is able to generate responses to questions that are similar to those of a human. The release of this chatbot has caused a lot of concern within the tech industry, with many people calling for a moratorium on the development of LLMs. Despite these concerns, the development of LLMs is continuing. Hughes has co-founded a new company, Auto-GPT, which is developing a new LLM that is even more powerful than ChatGPT. This new LLM, called Auto-GPT, is able to generate responses to questions without needing to be prompted. This is a significant improvement over ChatGPT, which required the user to prompt it again to keep going. With Auto-GPT, you can set up the AI to automatically generate the next prompt based on the previous response. This is really powerful because it lets you build up long strings of data. So if you are trying to generate a list of all the restaurants in a city, for example, you can have Auto-GPT automatically generate the next restaurant name based on the last one it gave you. Auto-GPT is still in beta, so it is not as user-friendly as some of the other options on this list. But it is really powerful and could be a great option for those with the coding skills to use it.

GPT News Summary Sources

Title	Media	Timestamp	Description	Link
Large, creative AI models will transform lives and labour markets	The Economist	2023-04-22 23:55:39.052920	GPT-3 is able to process a maximum of 2,048 tokens at a time, which is around the length of a long article in The Economist. GPT-4, by contrast...	<a href="https://www.economist.com/interactive/science-and-technology/2023/04/22/large-creative-ai-models-will-transform-how-we-live-and-work">https://www.economist.com/interactive/science-and-technology/2023/04/22/large-creative-ai-models-will-transform-how-we-live-and-work</a>
Student Who Never Attended Classes Claims To Have Scored 94% in Exam With ChatGPT's Help	NDTV	2023-04-22 23:03:39.869218	In the latest viral claim, ChatGPT helped a student prepare for his exam and score a stunning 94 per cent.	<a href="https://www.ndtv.com/features/student-who-never-attended-classes-claims-to-have-scored-94-in-exam-with-chatgpt-help-3069730">https://www.ndtv.com/features/student-who-never-attended-classes-claims-to-have-scored-94-in-exam-with-chatgpt-help-3069730</a>
ChatGPT sparks AI investment bonanza	DW	2023-04-22 23:03:39.851599	The launch of a new branch of artificial intelligence (AI) has reenergized the global tech sector. As investors pour billions into AI startups...	<a href="https://www.dw.com/en/chatgpt-sparks-ai-investment-bonanza/a-65368393">https://www.dw.com/en/chatgpt-sparks-ai-investment-bonanza/a-65368393</a>



# Market Research use cases



The screenshot shows a web browser window with a document from Gap Inc. on the left and a chat interface on the right. The document is titled "Gap Inc. GAP INC. REPORTS THIRD QUARTER FISCAL 2022 RESULTS" and contains text about financial results for the third quarter of fiscal 2022. The chat interface on the right is titled "Discover" and shows a search for "Lululemon third quarter fiscal 2022 results". It includes a comparison table between Gap Inc. and Lululemon, a "Stop Responding" button, and a text input field for asking questions.

**Gap Inc.**

**GAP INC. REPORTS THIRD QUARTER FISCAL 2022 RESULTS**

**SAN FRANCISCO – November 17, 2022** – [Gap Inc.](#) (NYSE: GPS), a portfolio of purpose-led, billion-dollar lifestyle brands including Old Navy, Gap, Banana Republic, and Athleta, and the largest specialty apparel company in the U.S., today reported financial results for its third quarter ended October 29, 2022.

“I have deep conviction that we have a portfolio of iconic brands that our customers love, increased confidence in our platform to drive leverage and economies of scale, and belief in the team’s ability to deliver. We have sharpened our focus on execution to optimize profitability and cash flow, are bringing more rigor to our operations, and balancing our assortments in response to what our customers are telling us. While our efforts show early signs of improvement, we are clear that there is work to be done to deliver what our customers, employees and shareholders expect from Gap Inc.” said Bob Martin, Executive Chairman and Interim CEO, Gap Inc.

**Third Quarter Fiscal 2022 - Financial Results**

- Net sales of \$4.04 billion, up 2% compared to last year. Comparable sales were up 1% year-over-year.
  - Online sales increased 5% compared to last year and represented 39% of total net sales.
  - Store sales increased 1% compared to last year. The company ended the quarter with 3,380 store locations in over 40 countries, of which 2,743 were company operated.
- Reported gross margin was 37.4%; adjusted gross margin, excluding \$53 million in impairment charges related to Yeezy Gap, was 38.7%, deleveraging 320 basis points versus last year.
  - On a reported basis, merchandise margin declined 480 basis points versus last year; adjusted for the impairment charge, merchandise margin declined 370 basis points. Merchandise margins were negatively impacted by higher discounting and inflationary commodity price increases and partially offset by lapping last year’s higher air freight expense.
  - Rent, occupancy, and depreciation (ROD) leveraged 10 basis points versus last year.

**Discover**

**Chat** Compose Overview Site Info

[Compare this with Lululemon in a table](#)

✓ Searching for: **Lululemon third quarter fiscal 2022 results**

✓ Generating answers for you...

According to the web page and the web results, here is a comparison of Gap Inc. and Lululemon in the third quarter of fiscal 2022:

Metric	Gap Inc.	Lululemon
Net sales	\$4.04 billion <sup>1</sup>	\$1.9 billion <sup>2</sup>
Net sales growth	2% <sup>1</sup>	28

Learn more: [click to view](#)

[Stop Responding](#)

Ask me anything...

0/2000

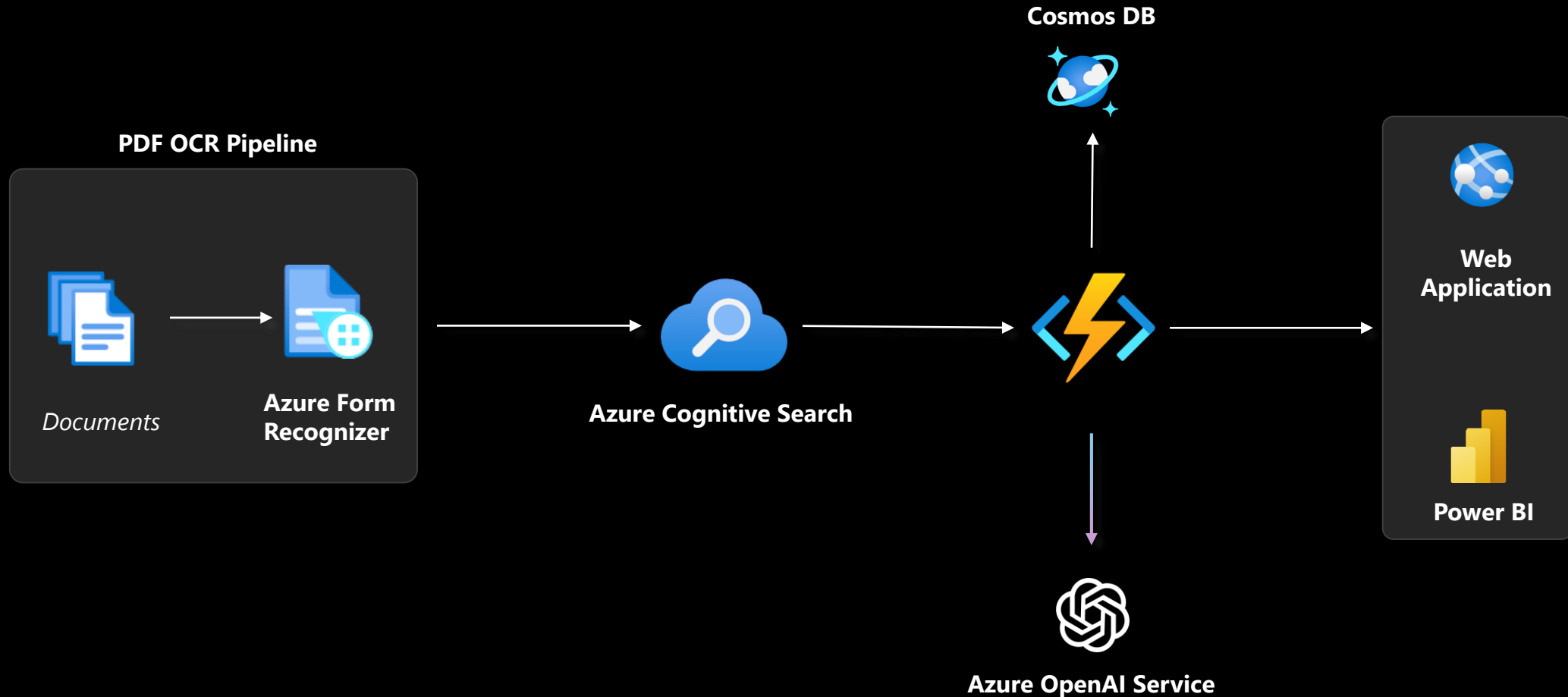
# Market Research use cases



News and Research		Target	Outcome
<b>BUILD</b> (customized knowledge)	Research's own	Enable intelligent news feed based on user's live interests in providing access to one's: <ul style="list-style-type: none"> <li>◦ emails,</li> <li>◦ meetings,</li> <li>◦ chats,</li> <li>◦ web navigation,</li> <li>◦ Research content metrics (e.g. time spent on topics)...</li> </ul>	Live feedback metrics collection and readjustment
	internal clients		Speech to text feeding CRM and other relevant desks
	external customers		Combined with Research platform to easily mine received data
<b>PUSH</b> (user fed information passively)	Research's own	Users deciding to opt for: <ul style="list-style-type: none"> <li>◦ live feed or</li> <li>◦ customizable frequency and delivery time</li> </ul> and means: <ul style="list-style-type: none"> <li>◦ Email summary,</li> <li>◦ Digital room (for better customization)</li> <li>◦ Teams channel (news channel),</li> <li>◦ during live client conversations when a specific topic is mentioned...</li> </ul>	Less information
	internal clients		More relevance
	external customers		At the right time
<b>PULL</b> (user actively looking for information)	Research's own	AI powered assistant to: <ul style="list-style-type: none"> <li>◦ suggest headlines and topics to user investigating Research content</li> <li>◦ propose cross references to users writing Research content</li> </ul>	Enhanced user navigation
	internal clients	AI powered chatbot for users to extract client information from: <ul style="list-style-type: none"> <li>◦ CRM: stats on deals and interests (Research info sent, subscriptions, time spent reading it...), interactions (last meeting)</li> <li>◦ Live feedback on client navigation and searches by type</li> <li>◦ LinkedIn and other publicly available sources to summarize interests and profile</li> </ul>	Faster and richer meeting preparation
	external customers	Chatbot for users to: <ul style="list-style-type: none"> <li>◦ Navigate faster in Research content database</li> <li>◦ Facilitate value extraction (desired format, summary or deep dive...)</li> </ul>	Improved customer experience

# Document process automation

Extract rich insights from documents and summarizing them



# Copilot in Excel



Microsoft



# Capital Markets - Top 4 Capabilities and Sell-Side Use Cases



## Content Generation

Pitchbook Generation, Assist in the Drafting of Research Paper

Automatic feed of CRM and dispatch to relevant desks

Live assistance while in client interaction (suggest Research paper, speaker coach...)

Call Center Analytics: Automatically generate responses to customer inquiries



## Summarization

Summary of Client interactions to prepare for meetings, analyze, score and improve relationships

Media, Financial Reports and Research Papers Summarization

Sentiment Analyzes to Detect and Assess (topic relevance, insights on position...)

Compliance monitoring, Error tracking and Fraud detection



## Semantic Search

Search Research Database, CRM, public and private sources

Information Discovery and Knowledge Mining

Offer customized research and information feed capabilities to clients

Pre-trade checklists based on Compliance Rulebook



## Code generation

Convert Natural Language to Code, Translate Code

Convert Natural Language to Query Proprietary Data Models, and to Excel macros

Code Documentation and Explanation (salvage individual developments)

# Azure OpenAI

## Benefits

Increased efficiency and productivity by getting more done in lesser time through rapid prototyping and quicker time to market\*

Faster time to realize value

Enhanced customer experience with a greater focus on customer-centric services and products, whilst utilizing feedback and trends better

Easy to use: even entry level data scientists or business users can use Azure OpenAI Service Playground

Ensure data privacy & security, and implement it in a responsible manner using a Trusted Cloud Provider

Does not require a long tedious annotation process\*

Less training data is required for many relevant use cases\*

Ability to perform text analytics and generation tasks that up until now were reserved only to humans

\* in most relevant use cases observed



Thank you