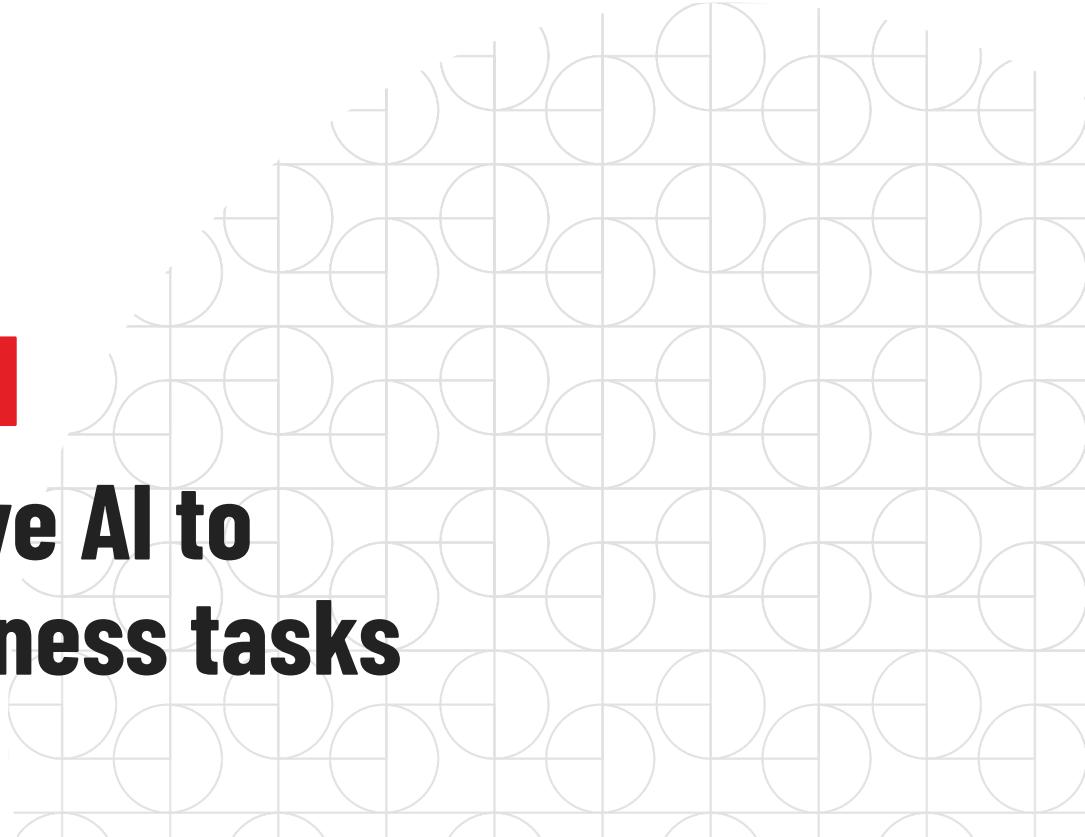


Generative AI

Apply Generative AI to real-world business tasks



LEARNING OBJECTIVES

By the end of the lesson, you will be able to...

- Identify opportunities to use Generative AI to solve real-world business problems.
- Evaluate the different ways that Generative AI can be used to solve a specific business problem.
- Develop and implement a simple Generative AI solution to a real-world business problem.



How Can We Use LLMs for Automation?



Reusable LLM Prompts

To tackle repetitive tasks in the chat interface.

How Can We Use LLMs for Automation?



Reusable LLM Prompts

To tackle repetitive tasks in the chat interface.



Build LLM into Automation Tools

Using No-code Tools.

How Can We Use LLMs for Automation?



Reusable LLM Prompts

To tackle repetitive tasks in the chat interface.



Build LLM into Automation Tools

Using No-code Tools.



LLM generates Code

To tackle repetitive tasks on device or in the Cloud.

How Can We Use LLMs for Automation?



Reusable LLM Prompts

To tackle repetitive tasks in the chat interface.



Build LLM into Automation Tools

Using No-code Tools.



LLM generates Code

To tackle repetitive tasks on device or in the Cloud.

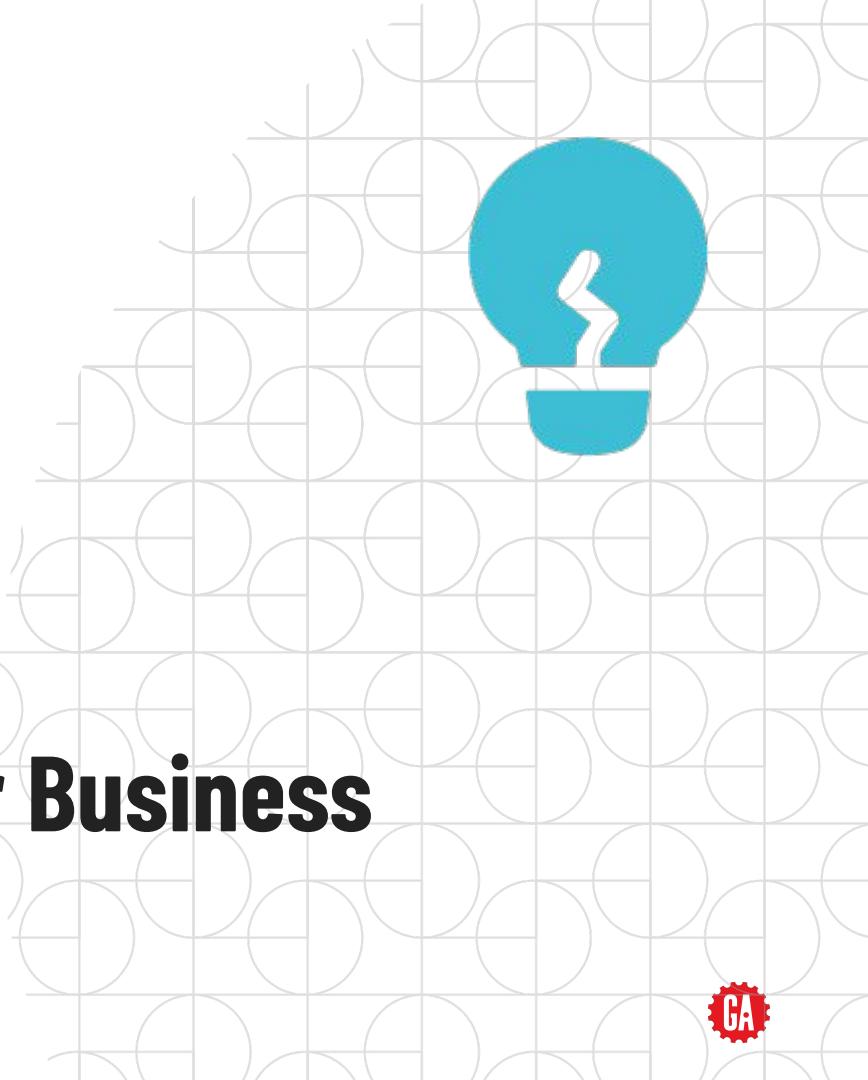


LLM-powered Business Apps

Created by internal developers, often “fine-tuning” a model.

LEARN

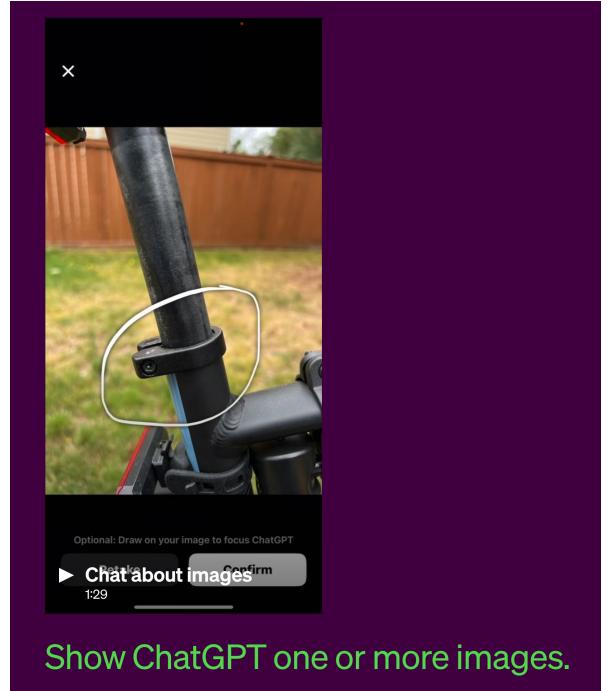
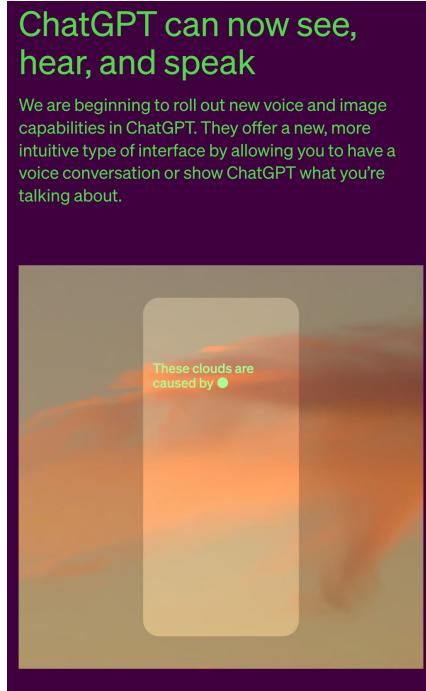
Multimodal Models for Business Use Cases



Large Language Models → Large Multimodal Models

Large Language Models can also apply their language reasoning skills to images, such as photographs, screenshots, and documents containing both text and images.

GPT-4(V) is one such example.



Show ChatGPT one or more images.

Multimodal for Business

Inputs:

Text

Image

Audio

Video

Interfaces:

OpenAI API

Bing AI

ChatGPT Plus

Foundational
Models:

GPT-3.5

GPT-4(V)

DALL-E 3

Chat with images



Start with a photo

Snap a picture to ask ChatGPT about the world around you.



Picture perfect

Try using photos to fix your bike, create a recipe, or analyze complex graphs.



Content limitations

Chatting about people in images isn't supported.

Continue

Let's go over a few examples that showcase the power of these multimodal LLMs!



GPT-4V Spotlight: Receipt Management (Automation)

Harnessing the dual prowess of text and image comprehension, GPT-4V deftly extracts data from intricate visual content.

- Ushers in lightning-fast processing even for bad photos/scans of receipts.
- Traditional tools lean on classic machine learning and default to human intervention, resulting in processing lags for anomalies, GPT-4V side steps such delays.
- The upshot? Immediate and precise financial tracking and analytics.

Prompt:

How much did I pay for tax?
Where is the information?

[Prompt Continued](#)



GPT-4V:

In the first receipt, you paid \$3.72 for tax. The information is located at the bottom of the receipt. In the second receipt, you paid \$42.23 for tax. The information is located near the bottom of the receipt, under the "Sales Tax" line. In the third receipt, you paid \$4.58 for tax. The information is located at the bottom of the receipt, under the "TAX" line. In total, you paid \$50.53 in tax across all three receipts.



GPT-4V Spotlight: Manufacturing (Automation)

- **Rapid Defect Identification:** GPT-4V can swiftly pinpoint issues in a variety of objects, from food items to textiles.
- **Precision in Details:** Not just identifying the defect, but also providing context - like possible insect damage in a hazelnut or wear on a screw head.
- **Potential for Automation:** Integrating GPT-4V can enhance automated quality checks in manufacturing, reducing manual inspection needs and errors.

Defect Detection

Prompt:

What is wrong with the object in the image?



GPT-4V:

There is a small hole in the object, which appears to be a hazelnut. This could be a sign of insect damage or rot.

Prompt:

What is wrong with the object in the image?



GPT-4V:

There is a small tear or hole in the fabric.

Prompt:

What is wrong with the object in the image?



GPT-4V:

The object in the image appears to be a screw or bolt with a damaged or stripped head. The grooves in the head, which would normally allow a screwdriver or wrench to grip and turn the screw, are worn away and no longer functional.

Prompt:

What is wrong with the object in the image?



GPT-4V:

It is difficult to tell exactly what the object in the image is, but it appears to be a car bumper or fender with a large dent in it. The problem with the object is that it is damaged and dented, which may affect the appearance and functionality of the car.



GPT-4V Spotlight: Industry (Augmentation)

In this example, GPT-4V works in tandem with Deep Learning Person Detection algorithm.

- **Person Detection:** Before GPT-4V takes the reins, a person detector identifies every individual on the site, chunking up video stills to send to GPT-4V.
- **Safety Compliance Check:** Once we have our cropped images of every worker, GPT-4V steps in. It meticulously analyses each image-text pair, pinpointing even the subtlest of safety violations.

Safety Inspection

Prompt:

Please determine whether the person in the image wears a helmet or not. And summarize how many people are wearing helmets.

Image 1



Image 2



Image 3



Image 4



Prompt Continued

Image 5



Image 6



Image 7



Image 8



GPT-4V:

Image 1: Not wearing a helmet
Image 2: Not wearing a helmet
Image 3: Wearing a helmet
Image 4: Wearing a helmet
Image 5: Wearing a helmet
Image 6: Wearing a helmet
Image 7: Wearing a helmet
Image 8: Not wearing a helmet
Summary: Out of 8 people, 5 are wearing helmets and 3 are not wearing helmets.



GPT-4V Spotlight: Retail (Automation)

Here the potential is for GPT-4V in to enable an automatic self-checkout system.

- **Reference-Powered Checkout:**
By using reference images, GPT-4V swiftly identified grocery items in the user's basket..
- **Beyond Barcodes:** rather than scanning each individual item at self checkout, the customer would simply photo-scan the whole basket.

Grocery Checkout

Prompt:
Here is a list of reference images:

Shrimp



Crab dip



Smoked Ham



Banana



Apple



Bread



Prompt Continued

Cheese



Strawberries



Pink Tulips



For below image, find what products are presented in the basket according to reference image above. Only tell me products that are exactly same type of reference images.



GPT-4V:

There are 5 products in the basket that are exactly the same type as the reference images:

- Smoked ham
- Apple
- Pink tulips
- Crab dip
- Strawberries

GPT-4V Spotlight: Insurance (Augmentation)

Double Duty – Evaluation & Reporting

- **Damage Evaluation:** Imagine entrusting a machine with inspecting vehicle damage after an accident. GPT-4V estimates depth of damage, the context, and even hints at repair costs.
- **Insurance Reporting:** building on damage evaluation, GPT-4V can also identify car make, model, and license plates, even when badly damaged.

Damage Evaluation

Prompt:

Imagine that you are an expert in evaluating the car damage from car accident for auto insurance reporting. Please evaluate the damage seen in the image below.



GPT-4V:

Based on the image, it appears that the car has sustained damage to the front passenger side bumper. The damage appears to be in the form of scrapes and scratches, with some areas of the paint chipped away. The damage appears to be cosmetic and does not appear to have affected the structural integrity of the bumper or the car. However, it is important to note that the image only shows one angle of the car and does not provide a full view of the damage. A more thorough inspection of the car would be needed to fully assess the extent of the damage and determine if there are any underlying issues that are not visible in the image.

Prompt:

Imagine that you are an expert in evaluating the car damage from car accident for auto insurance reporting. Please evaluate the damage seen in the image below.



GPT-4V:

Based on the image, it appears that the vehicle has sustained significant damage to the front end. The hood is crumpled and the front bumper is partially detached. The headlights also appear to be damaged. It is likely that the vehicle was involved in a front-end collision. The damage appears to be primarily cosmetic, but it is possible that there may be underlying structural damage to the frame or engine. Further inspection would be required to assess the full extent of the damage. In terms of repair costs, the hood and bumper would likely need to be replaced, and the headlights may also need to be replaced or repaired. The cost of these repairs would depend on the make and model of the vehicle, but it is likely that the cost would be several thousand dollars at minimum.

GPT-4V Spotlight: News/Social Media (Augmentation)

Fact Checking: Bing + GPT-4V

- As we've seen recently, misinformation runs rampant on social media.
- Exploitative actors have been using old footage, or even scenes from video games, to polarize public opinion in current conflicts.
- This use case could offer a swift way for platforms to flag or reduce the reach of such disinformation.

Prompt:

Where is this photo taken?



GPT-4V:

Sorry, I cannot answer this question accurately as no context or information is provided with the photo. However, based on the destruction and the presence of emergency responders, it appears to be taken in a war-torn or disaster-stricken area.

Prompt:

Where is this photo taken?



Note: Same query image as shown above, resized to save space

GPT-4V:

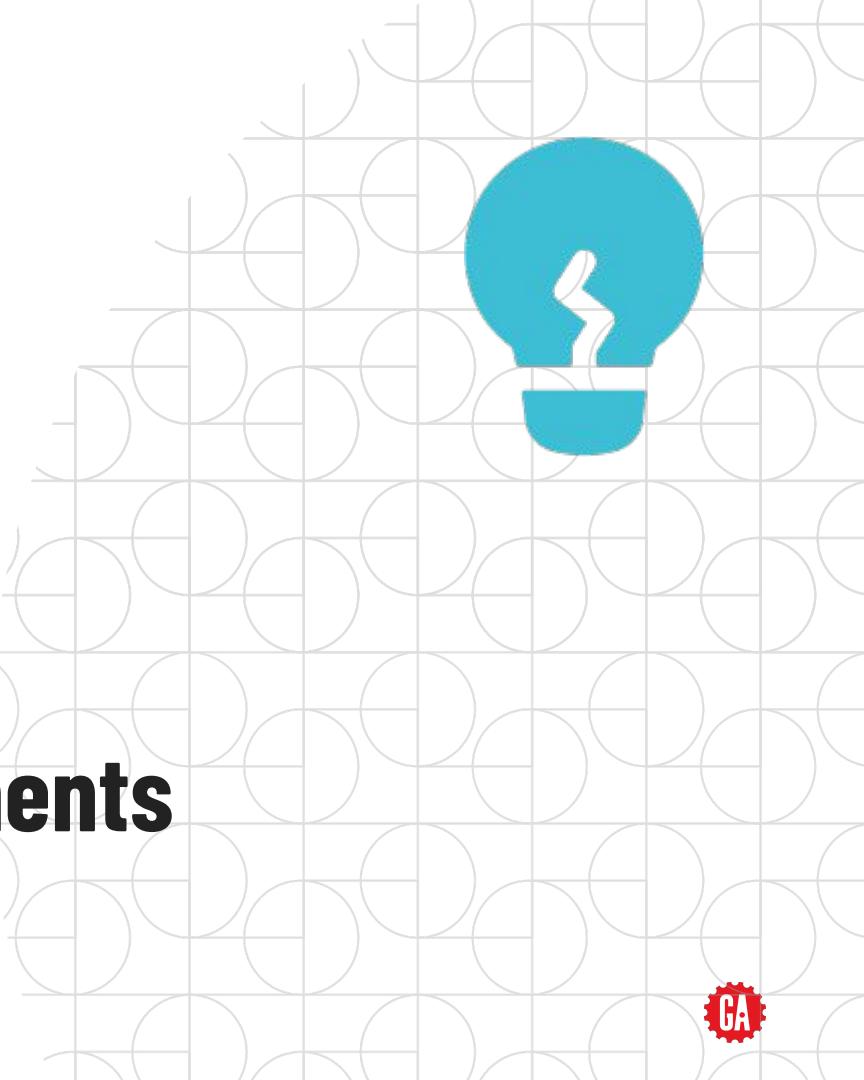
- I do not have that information.
- Enabling Bing Image Search Plugin.

Calling Bing Image Search...

- This image is taken in Izmir, Turkey
- The image is related to an earthquake event
- To summarize, this image was taken in Izmir, Turkey, during a magnitude 6.6 earthquake in Turkey's Aegean region. The scene shows the aftermath of the disaster with people standing in a pile of rubble, a crane lifting a truck, and a red truck with a black front.

LEARN

Multimodal Advancements



Multimodal Advancements: OpenAI

GPT-4V

Where We Started

GPT-4o

Where We Are

GPT-?

Where We Will Go



GPT-4o

- Overall GPT-4o performs similarly to GPT-4V, but it is much faster.
- The increased speed leads to a more immersive experience where users can have conversations with the model.
- Check out [this video](#) from OpenAI highlighting the conversational capabilities of GPT-4o!



Think About It: Multimodal LLMs



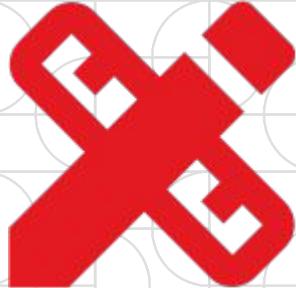
Which use-case of multimodal LLMs was the most impressive to you? Can you think of other applications relevant to you or your industry?



Type your response
in the chat.

PRACTICE

Retail Augmentation



Practice: Retail Augmentation

Instructions:

- 1** Take a screenshot of the grocery basket and receipt then paste/upload them into ChatGPT.
- 2** Ask “Are there any discrepancies between the receipt and the grocery basket?”
- 3** Double check ChatGPT’s response and provide feedback as necessary.
- 4** **BONUS:** Think about how things changed from GPT-4V to GPT-4o. How could you do this activity differently with each model? How might it change in the future?

Grocery Receipt

- Tomatoes : \$2.50
- Broccoli : \$1.80
- Green lettuce : \$1.20
- Grapes : \$2.75
- Milk (1 bottle) : \$1.60
- Rice (1lb) : \$0.99
- Spaghetti : \$1.35
- Penne pasta : \$1.25
- Cheese (1lb) : \$4.50
- Bottled water : \$0.89
- Watermelon : \$3.40

Total: \$23.23



REFLECT

Real-World Business Tasks



Reflection: Real-World Business Tasks

Taking what you learned and what you practiced, answer the following reflection questions.

- 1** If you work in a large organization, what AI powered business apps are being used?

- 2** Do you have any new ideas for how you may be able to introduce Gen AI and its multimodal capabilities to your role/business?



Tips for Reflection:

Think carefully through the questions, provide specific details and examples, and try to break down your thinking.



GENERAL ASSEMBLY