

NOVA

IMS

Information
Management
School

AI

Artificial Intelligence

Vitor Santos

vsantos@novaims.unl.pt

Marco António Silva

masilva@novaims.unl.pt

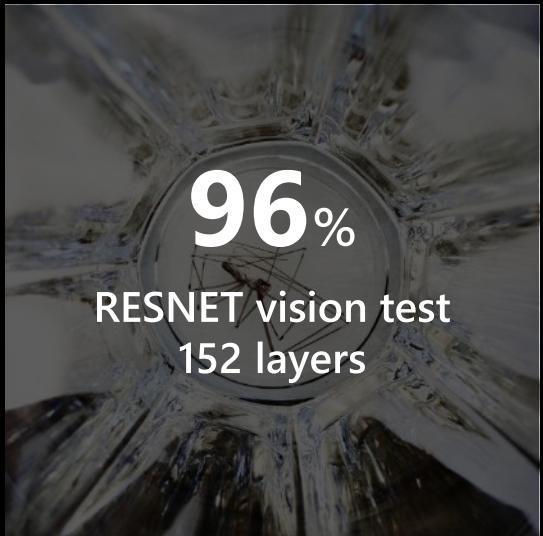
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

A close-up photograph of a man with dark hair and glasses, wearing a dark jacket over a collared shirt. He is gesturing with his right hand, pointing his index finger upwards, while holding a white remote control in his left hand. The background is dark.

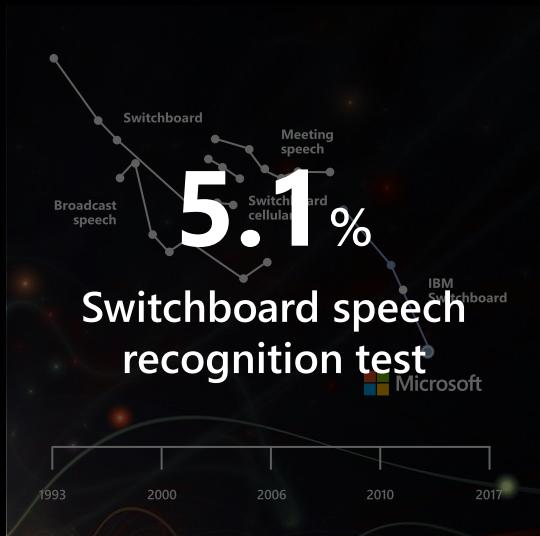
Our strategy is to build best-in-class platforms and productivity services for an intelligent cloud and an intelligent edge infused with artificial intelligence (“AI”).

Microsoft AI breakthroughs

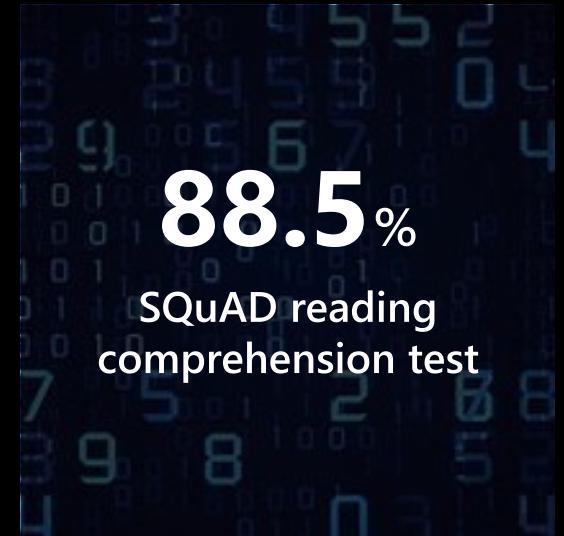
Vision



Speech



Language



2016

Object recognition
Human parity

2017

Speech recognition
Human parity

March 2018

Machine translation
Human parity

January 2018

Machine reading comprehension
Human parity

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



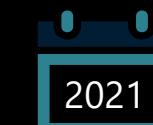
Machine Learning

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



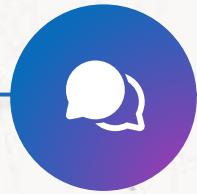
Generative AI

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

AI-driven business transformation



Enrich
employee
experiences



Reinvent
customer
engagement



Reshape
business
processes



Bend the
curve on
innovation



Azure AI Foundry



Copilot Studio



Visual Studio



GitHub



Azure AI
Foundry SDK

Model Catalog

Foundational models

Open-source models

Task models

Industry models

Azure
OpenAI Service

Azure AI
Search

Azure AI
Agent Service

Azure AI
Content Safety

Azure
Machine Learning

Evaluations

Customization

Governance

Monitoring

Observability

Azure AI Services

Build cutting-edge, market-ready AI applications with out-of-the-box and customizable APIs and multi-modal models

The screenshot shows the Azure AI Foundry interface with the project 'contoso_a1' selected. The left sidebar has a navigation menu with 'Overview', 'Model catalog', 'Playgrounds', 'AI Services' (selected), 'Build and customize', 'Code', 'Fine-tuning', 'Prompt flow', 'Assess and improve', 'Tracing', 'Evaluation', 'Safety + security', and 'My assets' sections for 'Models + endpoints', 'Data + indexes', and 'Web apps'. The main content area is titled 'Azure AI Services' and describes creating intelligent apps with task-specific models. It features a 'Speech playground' section with a video player showing a yellow biplane over mountains, a transcription, and a pronunciation score of 92%. Below this are sections for 'Content Understanding', 'Document field extraction', and 'Infuse your solutions with AI capabilities' for Speech, Language + Translator, Vision + Document, and Content Safety. The 'What's new' section highlights 'Document translation', 'Ensure content safety for generative AI', and 'Extract PII'. The 'Learning resources' section includes links to Documentation, Watch a video, Get started with AI on Azure, and Microsoft Q&A.

Azure AI Services

Create intelligent apps with these small, task-specific models—created responsibly by Microsoft—that are accurate, cost-efficient, and ready to use with a single key.

Speech playground

Enhance customer experiences through speech to text, text to speech, and speech translation features.

Try speech playground

Translated video Original
Translated video Original
Today was a [beautiful] day we ha
outside. [] In [] the morning t
yet the air was crisp and cold tow

Pronunciation score
Score breakdown
Accuracy 92%
Fluency 92%

Content Understanding Preview
Transform content of any modality into task specific structured data using Generative AI.

Document field extraction Preview
Extract fields from documents and forms using a custom generative extraction model.

Infuse your solutions with AI capabilities

Speech
Enhance customer experiences through speech to text, text to speech, and speech translation features.
View all Speech capabilities

Language + Translator
Analyze, summarize and translate using LLM-powered natural language processing capabilities.
View all Language + Translator capabilities

Vision + Document
Discover information and insights from documents, images and video with OCR and multi-modal AI.
View all Vision + Document capabilities

Content Safety
Detect harmful, offensive, or inappropriate user-generated or AI-generated content in your app including text, image, and multi-modal APIs.
View all Content Safety capabilities

What's new

Document translation
Translate documents from source language to target language from file types such as .docx, .pptx, .xlsx, .txt, .html and more.

Ensure content safety for generative AI
Detect harmful, offensive, or inappropriate AI-generated content in your application.

Extract PII
Identify and redact sensitive entities that are associated with an individual.

Learning resources

Documentation
What are Azure AI services? Learn how each service can help you meet your development goals.
Read the documentation

Watch a video
The AI Show Live showcases the amazing work happening in AI at Microsoft.
Watch a video

Get started with AI on Azure
Learn about the kinds of solutions AI can make possible.
Microsoft Learn

Microsoft Q&A
Find it on Q&A — the home for technical questions and answers at Microsoft.
Microsoft Q&A

All hubs + projects Project contoso_a1

Help

Management center

Image may not reflect actual user interface.

Design intelligent apps with Azure AI services

Leverage out-of-the-box and customizable APIs and multimodal models

Azure OpenAI Service

- Access to powerful AI models
- Scalable development
- Compliance & security
- Integration with other Azure Services

Azure AI Search

- AI enrichment & semantic ranking
- Generative AI content creation
- Vector search for data organization

Azure AI Speech

- Speech to text (including the Whisper model on Azure OpenAI Service)
- Text to speech
- Speech translation
- Speaker recognition

Azure AI Vision

- Image and face analysis
- Custom model training
- Face detection and recognition
- Document text extraction

Azure AI Content Safety

- AI-driven content moderation for enhanced safety
- Customize safety thresholds for diverse user types
- Detect and prevent Jailbreak Risk from XPIA attacks

Azure AI Document Intelligence

- Automated documentation generation
- Documentation quality analysis
- Interactive documentation experiences
- Natural language understanding for documentation

Azure AI Language

- Task-optimized AI models for text analytics
- Custom industry-specific AI for healthcare
- Custom, industry-specific models

Azure AI Translator

- Multilingual text and speech translation
- Synchronous and asynchronous translation request support
- Native translation of documents and manuals



Azure AI Speech

Speech-to-text

- Contact center – post call analytics and agent assist
- Meeting transcription & insights
- Audio/video transcription & insights
- Dictation
- Caption & subtitles for online and in-person meetings
- Audio and video captions & subtitles

Text-to-speech

- Context-aware, highly expressive HD voices
- Professional voice/brand voice
- Avatars

Speech translation

- Interpreter agent in Microsoft Teams
- Video translation API
- Real-time captions and subtitles
- Audio and video caption & subtitle
- Batch video

The screenshot displays the Azure AI Foundry interface, specifically the Speech Playground section. On the left, a sidebar provides navigation and management options. The main area features a grid of cards representing different speech scenarios:

- Real-time transcription:** Live transcription capabilities on your own audio without writing any code.
- Fast transcription:** Leveraging advanced speech recognition technology for rapid analysis and recognition.
- Batch transcription:** Transcribe a large amount of audio in storage and get results asynchronously.
- Speech translation:** Translate speech into other languages of your choice with low latency.
- Custom speech:** Fine-tuning. Customize speech to text with your data to adapt to specific speaking styles, vocabulary, and more.
- Voice gallery:** Quickly pick from a variety of prebuilt voices to use in your text-to-speech apps.

A modal window titled "Real-time transcription - Sample code" is open in the foreground, showing fields for "Endpoint" (set to "https://.westus.preview.app.vision.azure.com"), "Resource key" (redacted), and "Region" (set to "eastus"). Below the modal, a note states: "You should use environment variables or a secret management tool like Azure Key Vault to prevent accidental exposure of your key in applications. Learn more".

Available in ai.azure.com



Azure AI Speech-to-text Offerings

Realtime Transcription

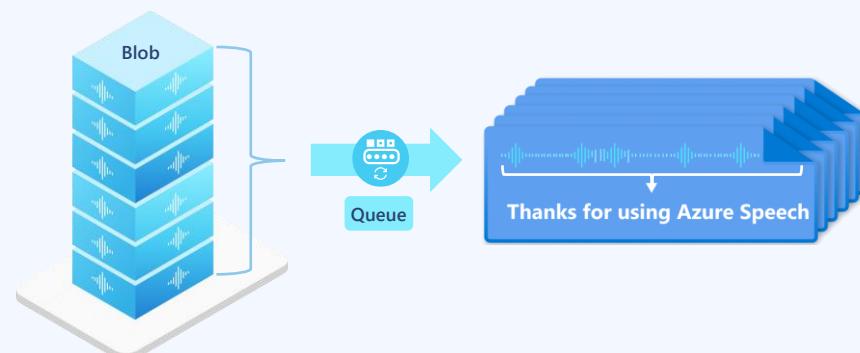
Transcribe from **microphone / streaming audio**, returning transcripts in **real time** with **interim results**.



Realtime Factor ~1
First response in a few hundred milliseconds

Batch Transcription

Transcribe **large amount of audio files** in storage, returning transcripts **asynchronously** with **low time urgency**.



Realtime Factor not guaranteed
(Task queueing, low urgency)

Fast Transcription (new)

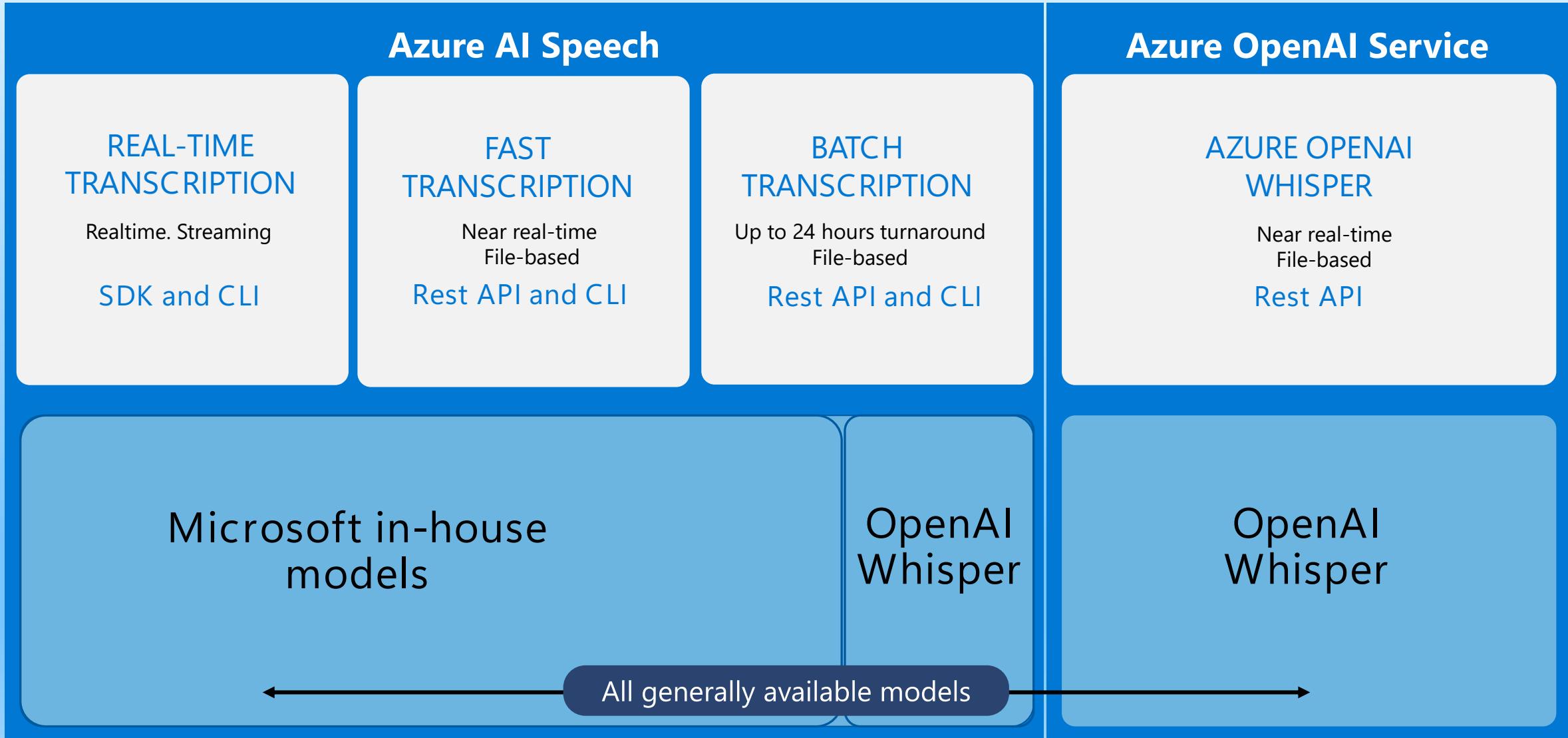
Transcribe **one audio file per request**, returning transcripts **synchronously** and **quickly**.



Realtime Factor << 1
(e.g., 20 mins audio done in 1 min)

Azure AI has three differentiated speech to text offerings

All of them with generally available models



Azure AI Speech

Fast transcription

**Ultra-fast audio file to text conversion for post call analytics,
voicemail, video editing, and more**

Long, context-aware native display models

Speaker diarization

Multi-channel audio input

Near real-time response

[Available in ai.azure.com](https://ai.azure.com)



Azure AI Speech

Custom text-to-speech avatar

Create a unique and realistic avatar or digital twin using your own video data

Custom avatar model training

Full control of your data,
model, and endpoints

API for video and live chatbot
creation tool

[Available in ai.azure.com](https://ai.azure.com)



Azure AI Speech Translation capabilities

Multi-lingual translation

- Automatic language detection
- Real-time translated captions
- Use text to speech platform voices
- Multiple languages in the same session

Video translation

- Batch service for e2e video dubbing
- Speaker diarization and voice matching
- Uses platform voices or personal voice
- Human in the loop for content editing
- GPT reformulation for translation
- Powering video dubbing in Edge Canary



Azure AI Speech offers transcription technologies for every

	Real-time Transcription	Batch Transcription (include Batch Whisper)	Fast Transcription (new)	Whisper in Azure OpenAI Service
Description	Transcribe live audio from microphone / stream , returning transcripts in real time with interim results .	Transcribe large amount of audio files in storage, returning transcripts asynchronously with low time-urgency .	Transcribe one audio file per request , returning transcripts synchronously and quickly .	A replication of OpenAI Whisper API hosted in Azure.
Input Form	Streaming audio	Audio file (many files via a blob storage URL)	Audio file (one audio file per request)	Audio file (one audio file per request)
Recommended Using Scenarios	<ul style="list-style-type: none"> Online Meeting Caption/ Subtitle (translation) In-person Meeting Caption/ Subtitle (translation) Contact Center - Voice Agent Contact Center – Agent Assist (streaming input with partial results) Voice Assistant/Chatbot Dictation Voice Commanding 	<ul style="list-style-type: none"> Contact Center – Post Call Transcription and Analytics (lower urgency) Meeting Transcription – Post Meeting Summarization and Analytics (lower urgency) Recorded Audio/Video transcription (lower urgency) 	<ul style="list-style-type: none"> Quick Audio / Video edit Contact Center – Agent Assist (file input) Contact Center – Post Call Transcription and Analysis (higher urgency) Meeting Transcription – Post Meeting Summarization and Analytics (higher urgency) Recorded Audio/Video transcription (higher urgency) Voice Mail Video Dubbing 	<ul style="list-style-type: none"> Offline Audio/Video captioning/subtitling Audio/Video Transcription Speech translation to English only
Use if you need	<ul style="list-style-type: none"> Streaming input with microphone; Intermediate recognizing results Improve quality with Phrase Lists Improve quality with Custom Speech and deploy endpoints 	<ul style="list-style-type: none"> Bulk audio files in a storage in a request Extreme large audio files with >2hrs audio length Large selection of model versions Audio formats/codec support in service side Multiple audio channels Best quality 	<ul style="list-style-type: none"> Simple config via one HTTP post Transcribe much faster than real audio length One audio file per request Audio formats/codec support in service side Good quality. 	<ul style="list-style-type: none"> Good readability and disfluency Multilanguage recognition
Don't Use if you need	<ul style="list-style-type: none"> Get transcript much faster than real audio time; Audio formats/codec support in service side 	<ul style="list-style-type: none"> Streaming input; Get transcript quickly and synchronously 	<ul style="list-style-type: none"> Streaming input; Lexical format result (e.g., "one two three" instead of "123") 	<ul style="list-style-type: none"> Streaming input; Lexical format result (e.g., "one two three" instead of "123") Large audio files Customization with your data
RTF (Realtime Factor) and Latency	~ 1 (realtime), first-byte latency response in a few hundred milliseconds	Not guaranteed due to task queueing (low time urgency)	« 1 (e.g., 20 mins audio done in 1 min)	« 1

DEMO

Text To Speech– Speech API
Automatic Speech Recognition – Speech API





Future of Order Taking

drivethru.wav

Start Recognition

Neural TTS Samples

A palpable chill passes through the room at this maladroit refusal .



Synthesized



Recorded

You can't know how long Tara would have been mad , because that was taken away from her.



Synthesized



Recorded

In Georgia , governor Zell Miller has declared a state of emergency in four counties.



Recorded



Synthesized

Nostalgia could still be evoked after only twenty years , and many remembered when.



Recorded



Synthesized

The forecast for next week shows sunshine with highs in the mid forties.



Synthesized



Recorded

Azure AI Language and Translation

Build conversational interfaces, summarize documents, and analyze text using prebuilt AI-powered features

The screenshot shows the Azure AI Foundry interface for the 'Language' service. The left sidebar includes sections for Overview, Model catalog, Playgrounds, AI Services (selected), Build and customize (Code PREVIEW, Fine-tuning, Prompt flow), Assess and improve (Tracing PREVIEW, Evaluation), Safety + security, My assets (Models + endpoints, Data + indexes, Web apps), and a navigation bar at the top.

The main content area is titled 'Language + Translation' and describes integrating natural language into apps, bots, and IoT devices. It features a 'Language Playground' section with a button to 'Try the Language playground'. A sample interaction is shown: 'Hello, my name is Mateo Gomez. I los' (with 'Person' highlighted) followed by 'August 17th, and I would like to requ' (with 'DateTime' highlighted) and 'purchase I made was of a Chicken par' (with 'Text' highlighted). Configuration options include API version (2024-05-01 preview), Model version (Latest ("GA")), and Text language.

The 'What's new from Language' section highlights three capabilities: Extract PII from conversation, Extract health information, and Summarize for call center, each with a 'Try it out' button.

The 'Explore Language capabilities' section lists four categories: Extract Information (Classify Text, Summarize Information, Translation), Extract PII from text, Extract PII from conversation, Extract health information, Extract named entities, and Other language capabilities (View capabilities (10+)).

The 'Learning resources' section provides links to Documentation, Watch a video, Get started with Azure AI Language, and Microsoft Q&A.

At the bottom, a note states 'Image may not reflect actual user interface.'



Azure AI Language Overview

- A **unified service** with one Azure resource, one user experience, one set of APIs
- **Multilingual** for +100 languages via Z-Code and Z-Code++.
- Powered by large language models
- **Pre-built**, task-optimized models for immediate value
- **Customization** to adapt models to customer specific needs
- **Flexible deployment** through Azure hosted service or as customer-hosted containers



Regulatory
Compliance

CSA

FedRAMP

GDPR

HIPAA

HITRUST

ISO

PCI DSS
Certified

AICPA
SOC
Formerly SAS 70 Report

US DoD

Personally Identifiable Information (PII) Detection

- Identifies, categorizes, and redacts sensitive info in unstructured text
- Supports a wide range of PII info, including region-specific PII, PHI and PCI
- Supports conversational scenario with a specialized model (aka. Conversation PII)
- 79 languages supported

Typical use cases:

- Mask sensitive info to avoid the leak, inappropriate handling or data bias
- Auto label and classify docs based on the sensitivity

Sample text:

The food is great, and the place was impeccably clean. You can order from their online menu at www.contosohouse.com, call 312-555-0176 or send email to order@contosohouse.com!

Recognized entity categories:

URL

PhoneNumber

Email

Redacted text:

*The food is great, and the place was impeccably clean. You can order from their online menu at www.contosohouse.com, call *****-***** or send email to *****!**

PII Detection – *Supported Categories*

Basic PII Info

- Person
- Person Type
- Phone Number
- Organization – *incl. types:*
 - Medical
 - Stock Exchange
 - Sports
- Address
- Email
- URL
- IP Address
- Date Time – *incl. type:*
 - Date
- Quantity – *incl. type:*
 - Age

Government & Country/Region-Specific

- U.S. Social Security Number (SSN)
- U.S. Driver's License Number
- U.S. Passport Number
- U.S. Individual Taxpayer Identification Number (ITIN)
- U.S. Drug Enforcement Agency (DEA) Number
- U.S. Bank Account Number
- + *more than 90 PII entities across 31 other countries/regions*

Financial Account

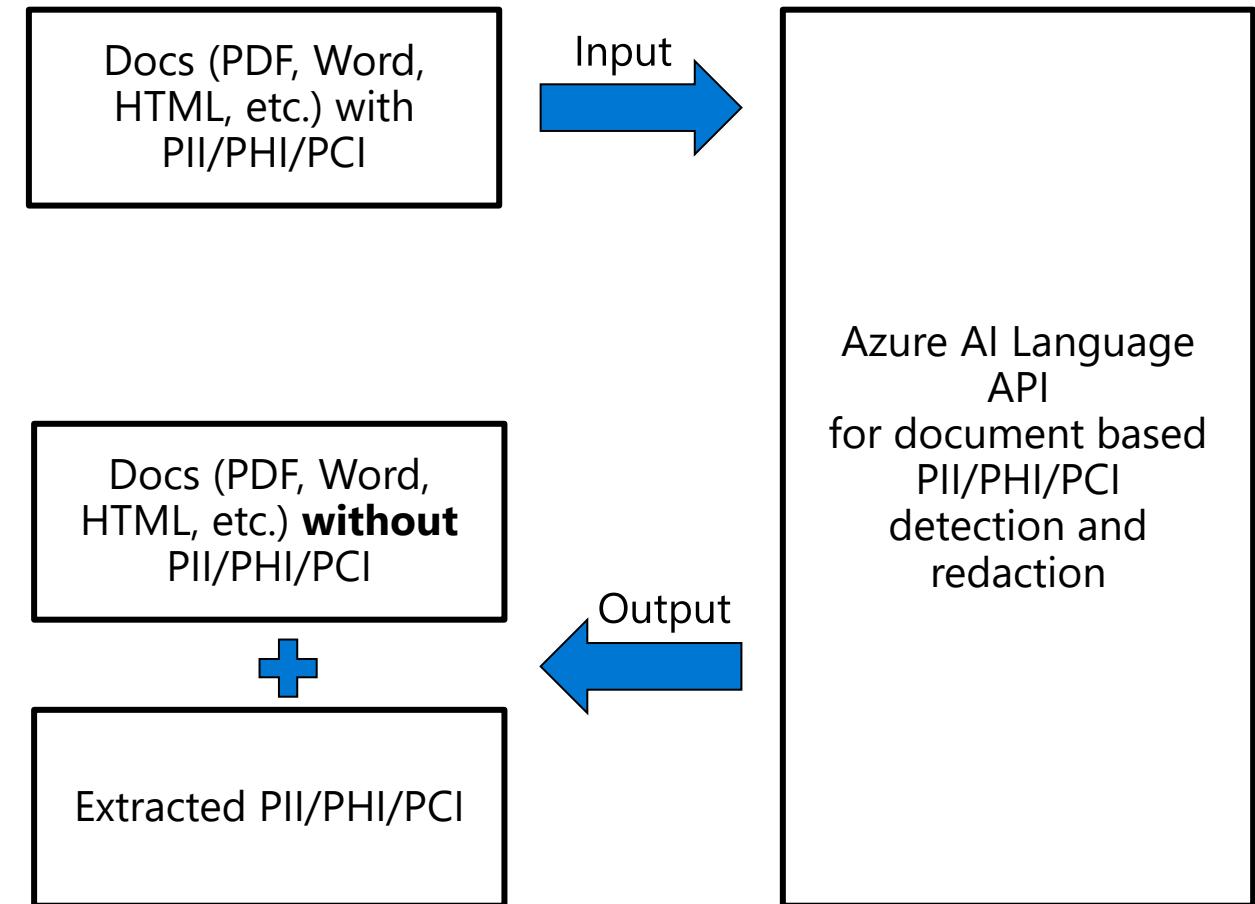
- ABA Routing Number
- SWIFT Code
- Credit Card Number
- IBAN Code

Azure Information

- Azure DocumentDB Auth Key
- Azure IAAS DB Connection String
- Azure SQL Connection String
- Azure IoT Connection String
- Azure Publish Setting Password
- Azure Redis Cache Connection String
- Azure SAS
- Azure Service Bus Connection String
- Azure Storage Account Key
- SQL Server Connection String

PII Redaction on Native Document Formats (preview)

- Identifies and redacts sensitive PII info directly in native documents –
No more doc pre-processing and post-processing needed!
- PII entities and values extracted from documents are included in the API response payload, which can be stored separated and used for later



Document Summarization

Auto-generate a concise summary from both **abstractive** and **extractive** approaches to improve the productivity of information analytics and processing.

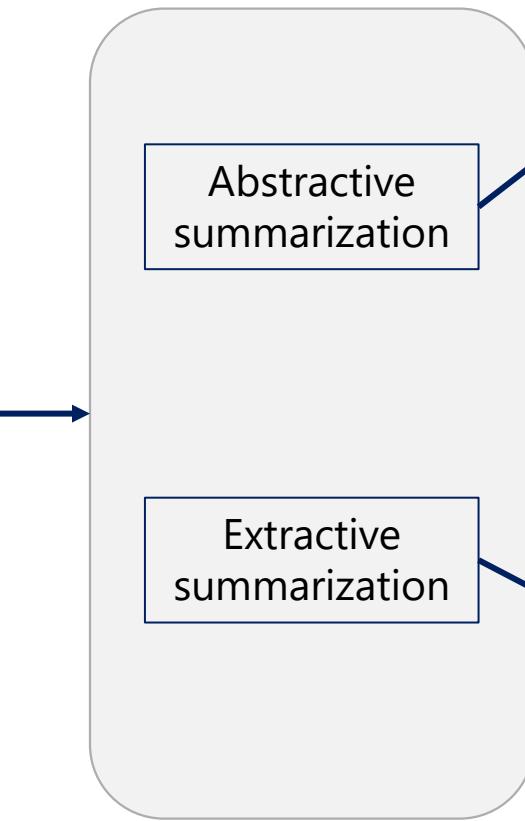
Input:

REDMOND, Wash. and Santa Monica, Calif. – Jan. 18, 2022 – With three billion people actively playing games today, and fueled by a new generation steeped in the joys of interactive entertainment, gaming is now the largest and fastest-growing form of entertainment. Today, Microsoft Corp. (Nasdaq: MSFT) announced plans to acquire Activision Blizzard Inc. (Nasdaq: ATVI), a leader in game development and interactive entertainment content publisher. This acquisition will accelerate the growth in Microsoft's gaming business across mobile, PC, console and cloud and will provide building blocks for the metaverse.

Microsoft will acquire Activision Blizzard for \$95.00 per share, in an all-cash transaction valued at \$68.7 billion, inclusive of Activision Blizzard's net cash. When the transaction closes, Microsoft will become the world's third-largest gaming company by revenue, behind Tencent and Sony. The planned acquisition includes iconic franchises from the Activision, Blizzard and King studios like "Warcraft," "Diablo," "Overwatch," "Call of Duty" and "Candy Crush," in addition to global eSports activities through Major League Gaming. The company has studios around the world with nearly 10,000 employees.

Bobby Kotick will continue to serve as CEO of Activision Blizzard, and he and his team will maintain their focus on driving efforts to further strengthen the company's culture and accelerate business growth. Once the deal closes, the Activision Blizzard business will report to Phil Spencer, CEO, Microsoft Gaming.

Mobile is the largest segment in gaming, with nearly 95% of all players globally enjoying games on mobile. Through great teams and great technology, Microsoft and Activision Blizzard will empower players to enjoy the most-immersive franchises, like "Halo" and "Warcraft," virtually anywhere they want. And with games like "Candy Crush," Activision Blizzard's mobile business represents a significant presence and opportunity for Microsoft in this fast-growing segment.



Abstract of an article
(general or on your interest):

Abstract Summary:

Jan. 18, 2022 Microsoft Corp announced plans to acquire Activision Blizzard Inc. This acquisition will accelerate the growth in Microsoft's gaming business and will provide building blocks for the metaverse. The acquisition price is \$95.00 per share, in an all-cash transaction valued at \$68.7 billion. Microsoft will become the world's third-largest gaming company by revenue, behind Tencent and Sony.

Highlights key sentences:

Extract Summary:

This acquisition will accelerate the growth in Microsoft's gaming business across mobile, PC, console and cloud and will provide building blocks for the metaverse

Microsoft will acquire Activision Blizzard for \$95.00 per share, in an all-cash transaction valued at \$68.7 billion, inclusive of Activision Blizzard's net cash.

When the transaction closes, Microsoft will become the world's third-largest gaming company by revenue, behind Tencent and Sony.

Named Entity Recognition (NER)

- Identifies and categorizes entities in unstructured text
- 30+ pre-defined entity categories (includes subcategories)
- 79 languages supported

Typical use cases:

- Search indexing optimization
- Business process automation based on entity categories and values
- Customer analysis to highlight relevant topics

Available in containers

Sample text:	Recognized entity categories:
As dozens of heads of state convene for the annual U.N. General Assembly in New York this week, the lingering conflict in Syria is taking a back seat while tensions in the Persian Gulf and global trade wars take center stage.	PersonName
Now in its ninth year, many Syrians fear the unresolved war has become a footnote in a long list of world crises, with weary leaders resigned to live with President Bashar Assad ruling over a wrecked and divided country for the foreseeable future.	PersonType
On the eve of the global gathering in New York, U.N. Secretary-General Antonio Guterres announced that a long-awaited committee that would draft a new Syrian constitution has been finalized — a step the U.N. hopes will put the war-ravaged country on track for a political solution.	Location
	Organization
	Date
	Quantity
	Event

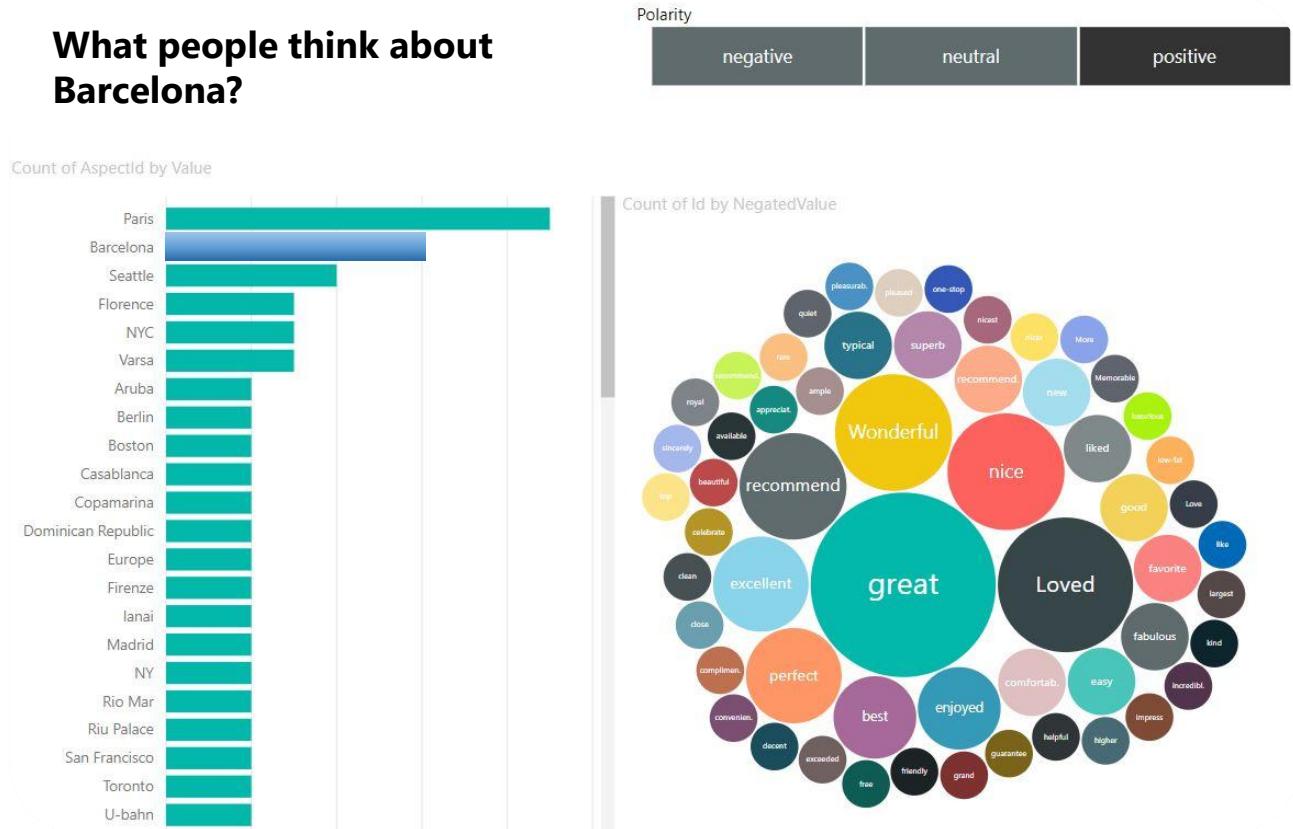
Sentiment Analysis and Opinion Mining

- Mines text for clues about positive or negative sentiment
- Associate sentiment with specific aspects of the text (i.e., targets & their assessment)
- 90+ languages supported

Typical use cases:

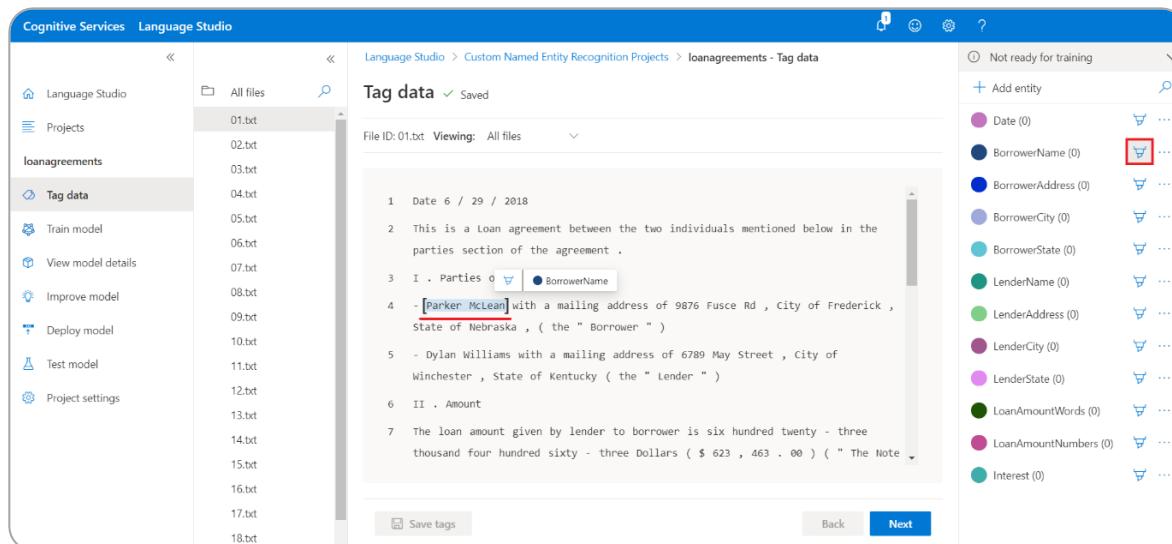
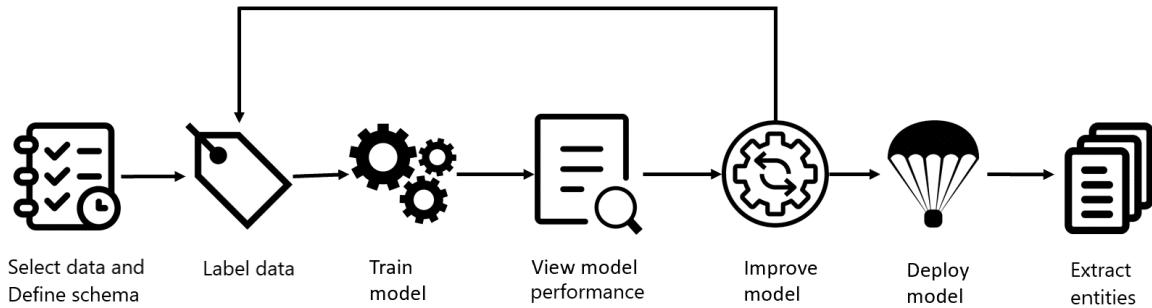
- Social media analysis to find out what people think of your brand, products, or competitors
- Customer call analysis to identify the calls in risk and opportunities of agent coaching

What people think about Barcelona?



Custom Named Entity Recognition

- Your own custom models to extract domain-specific entities from unstructured text
- 95+ languages supported
- Iteratively label data, train, evaluate, and improve model performance in **Language Studio** before making it available for consumption



Key Phrase Extraction

- Extracts key words and phrases
- Powered by transformer-based model
- Tailored for both **short texts** (without much topic information) and **long texts** (where it's easier to track topic)
- Supports 90+ languages

Typical use cases:

- Search indexing optimization
- Article tags for topic filtering
- Word cloud

Sample text:

Around a third of the global population in the world's poorest countries have been pinning their hopes on India to deliver their Covid-19 vaccines. The Covid-19 outbreak is sad. Then the virus overwhelmed India itself. The outbreak has caused widespread misery for the country's population of 1.3 billion. Hospitals are overwhelmed, oxygen is scarce and the death toll of 200,000 is thought to be a huge undercount. Covid-19 is a dangerous disease.

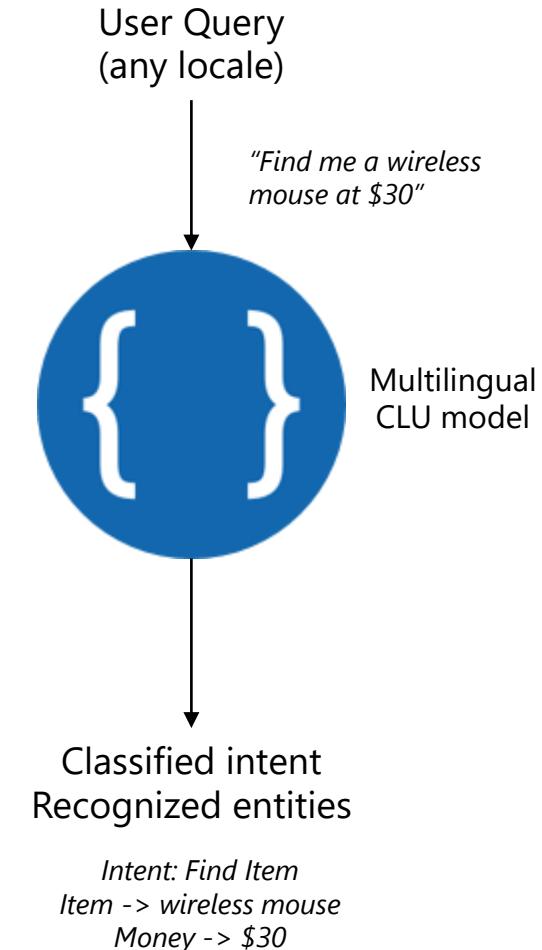
Extracted key phrases:

The Covid-19 outbreak, Covid-19 vaccines, poorest countries, widespread misery, death toll, huge undercount, dangerous disease, global population, third world, hopes, India, virus, country, Hospitals, oxygen

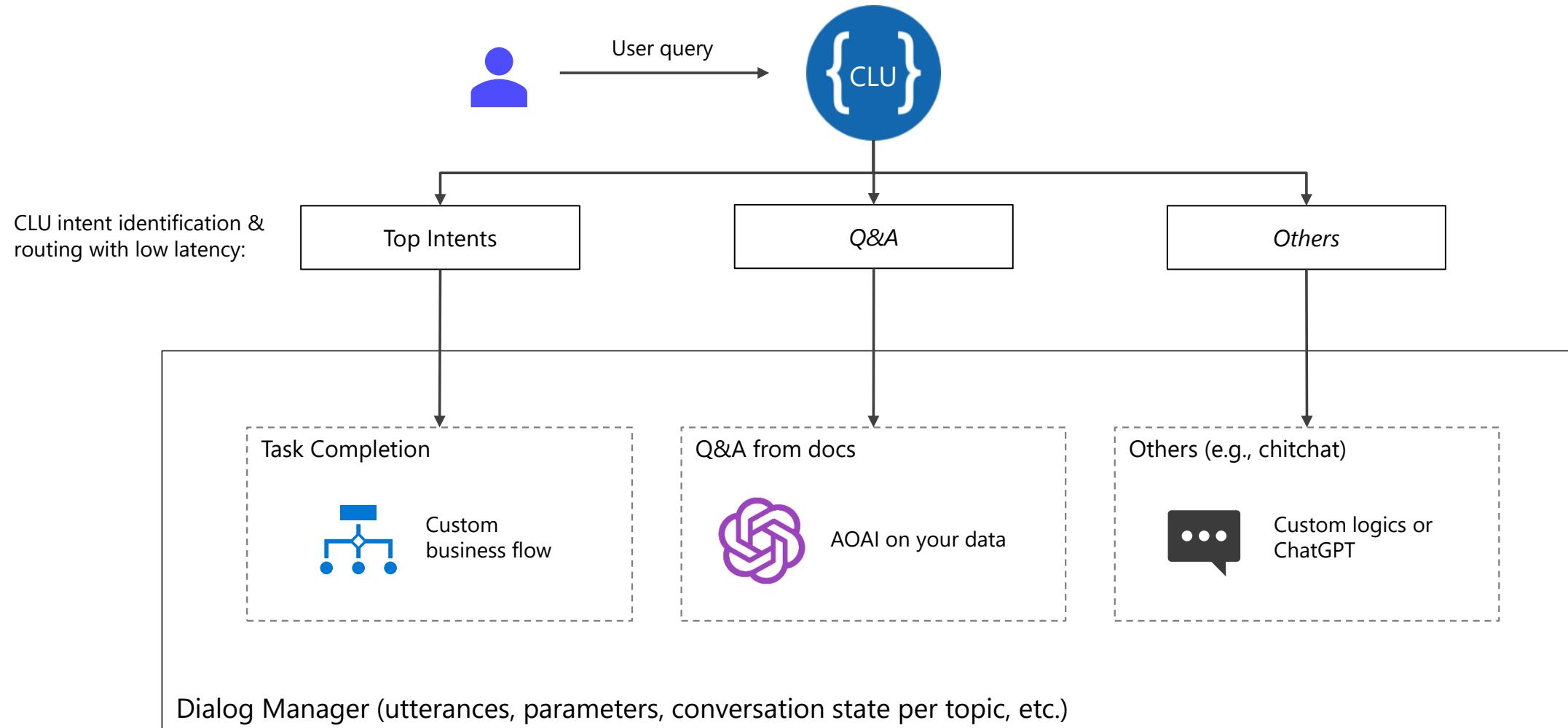
Conversational Language Understanding (CLU)

Train and deploy your custom CLU model to predict intents and extract entities using Language Studio

- Understands the semantics meaning, not just text patterns
- Native multilingual support: Train in one language, predict in all other 96+ languages
- Azure OpenAI-powered data labeling to accelerate the custom model building
- Orchestrate with Azure OpenAI to enable your conversational AI experience with higher quality, lower cost and latency



Orchestrate CLU with Azure OpenAI Service



Translator API overview

Translate Text

Into multiple languages

Transliterate from one script to another

Translate Documents

Retaining structure and layout

Into multiple languages

Large volumes of text data

Real time or batch processing for large documents

Custom Models

Build custom models with easy-to-use Custom Translator

Uses general & custom models to translate Text and Documents

Security and Compliance

Enterprise ready with Azure AD authentication & Managed Identity

Regulatory Compliance



CSA



FedRAMP



GDPR



HIPAA



HITRUST



ISO



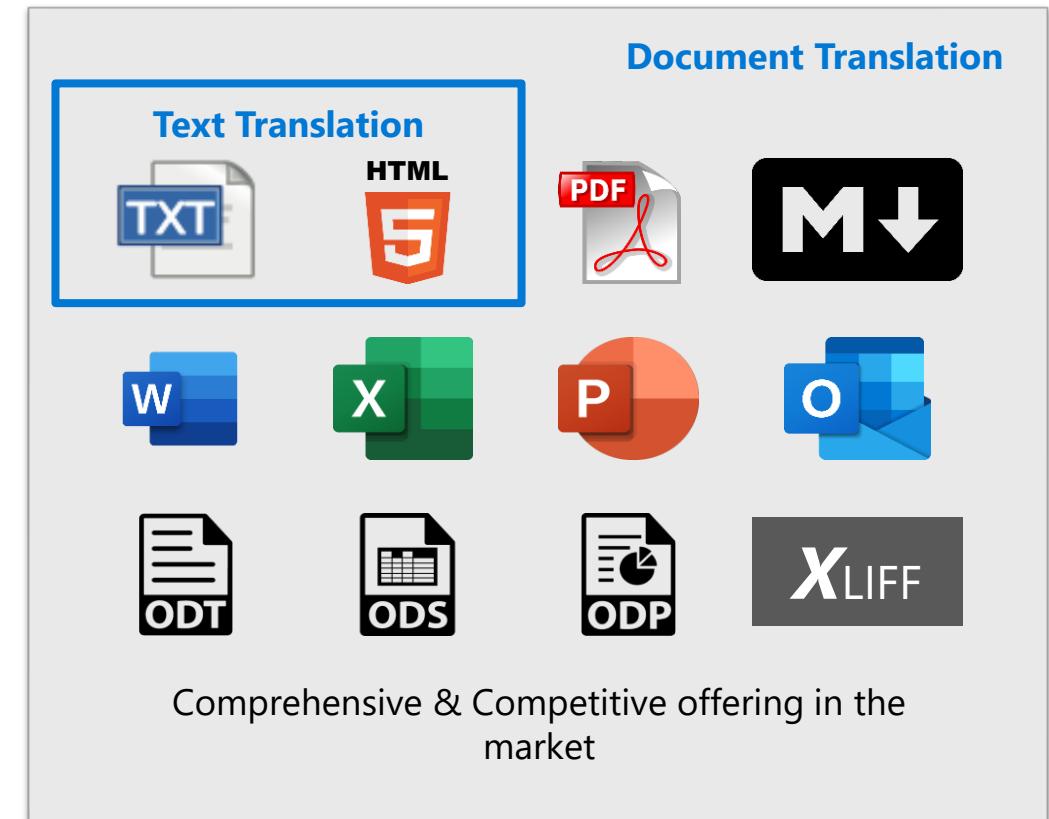
PCI



SOC



US DoD



Document Translation overview

Translates rich documents retaining structure and layout

Translates documents into multiple languages

Translates batch of large documents

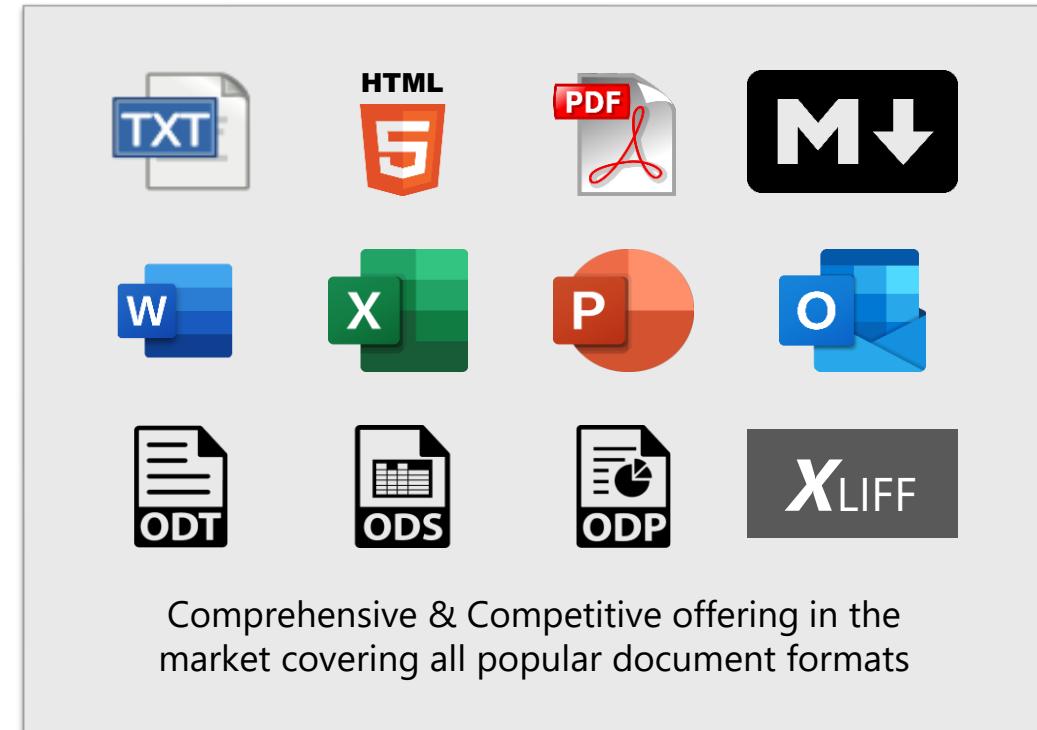
Translates multi-lingual documents

Uses general & custom models

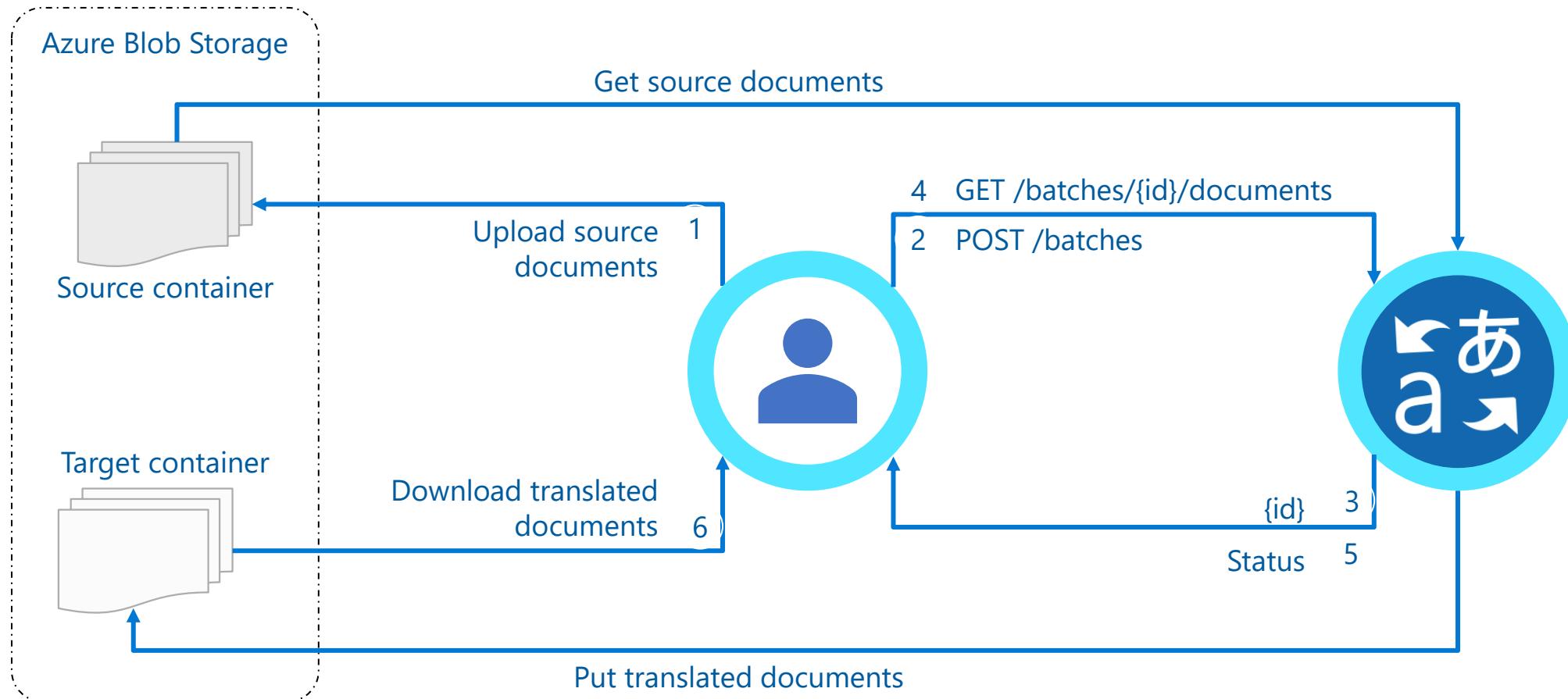
Applies custom glossaries

Azure AD authentication

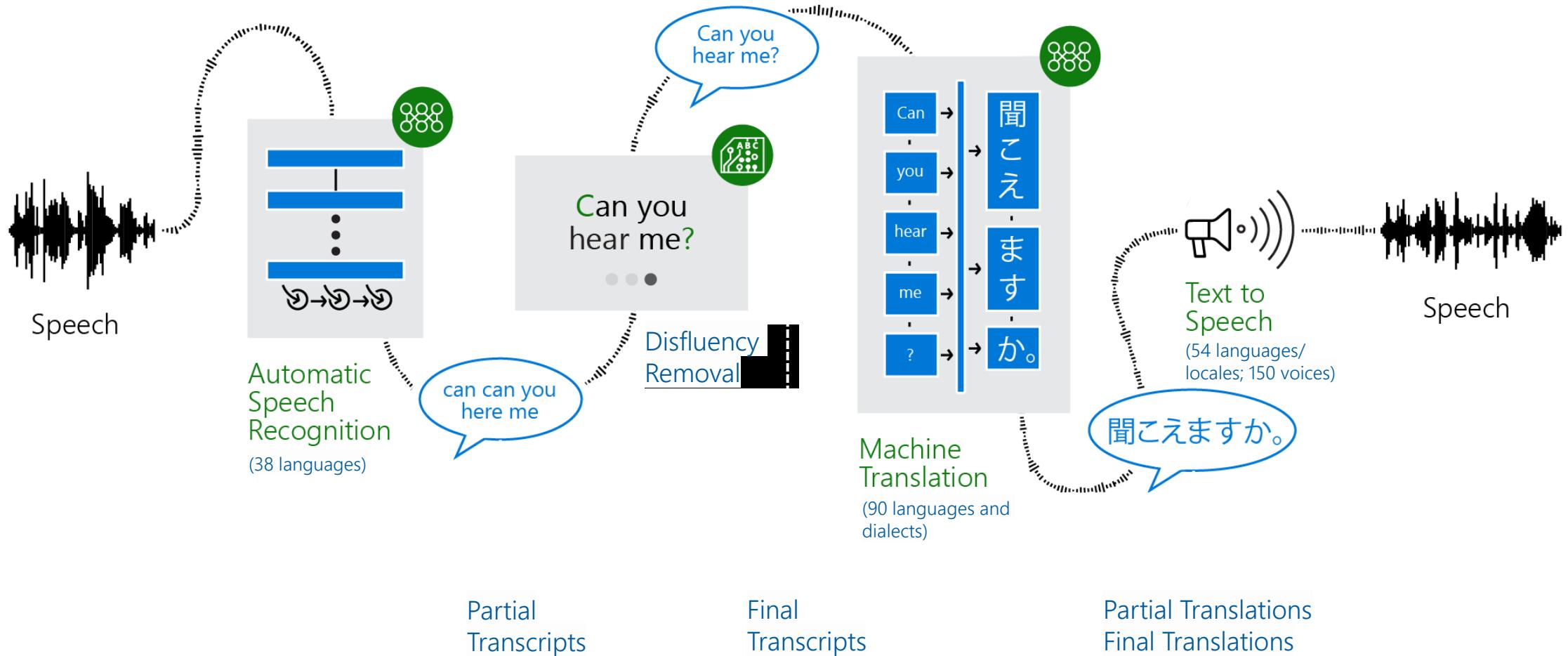
Secured access to storage using managed identity



Document Translation simple flow



Speech Translation SDK



Using Custom Translator

The screenshot shows the Microsoft Language Studio Custom Translator interface. On the left, a sidebar lists workspaces: 'maeWS' (selected) and 'kor-enu'. Under 'kor-enu', options include 'Manage documents', 'Train model' (selected), 'Model details', 'Test model', and 'Publish model'. 'Workspace settings' is also listed. The main area displays 'Model details > kor-enu'. A green success message 'Training succeeded' is shown with a BLEU score of **33.64** (↑ 33.64). Below this, a table shows 'Document sets used for training' with 8 rows of data. At the bottom, a note says 'Learn how to evaluate a model's BLEU score'.

Name	Created date	Type	Korean sentences	English sentences	Aligned sentences	Used sentences
korenabstractsdata100	06/21/2021	Training	925	967	635	
korenabstractsdata101	06/21/2021	Training	843	859	620	
korenabstractsdata102	06/21/2021	Training	1,325	1,260	880	877
korenabstractsdata103	06/21/2021	Training	1,228	1,209	764	
korenabstractsdata104	06/21/2021	Training	231	233	161	
korenabstractsdata105	06/21/2021	Training	870	851	676	
korenabstractsdata106	06/21/2021	Training	1,045	1,011	726	

1
Upload
your language training
data

2
Train
your new custom
model

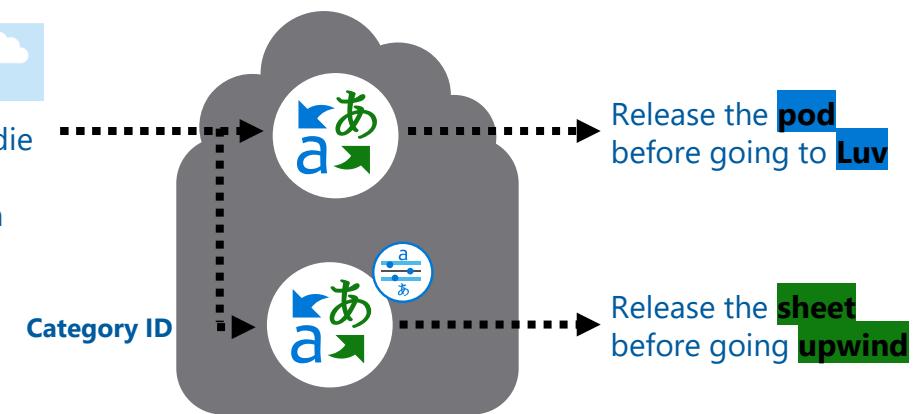
3
Test
to compare your new
custom translations to
standard translations

4
Deploy
to translate live via
Translator REST API

All popular document formats



Lassen Sie die
Schot los,
bevor Sie in
Luv gehen



Azure AI Vision and Content Understanding

Read text, analyze images, and detect faces with OCR and machine learning

The screenshot shows the Azure AI Foundry interface with the 'contoso_a1' project selected. The left sidebar navigation includes 'Overview', 'Model catalog', 'Playgrounds', 'AI Services' (selected), 'Build and customize', 'Code PREVIEW', 'Fine-tuning', 'Prompt flow', 'Assess and improve', 'Tracing PREVIEW', 'Evaluation', and 'Safety + security'. Under 'My assets', there are sections for 'Models + endpoints', 'Data + indexes', and 'Web apps'. The main content area is titled 'Vision + Document' and describes integrating vision with generative AI for reading text, analyzing images, and processing documents. It features sections for 'Content understanding' (with a 'Try demo' button) and 'Document field extraction' (with a 'Try demo' button). Below these are sections for 'View all other vision capabilities' (Document, Face, Image tabs, currently Document is selected), 'Prebuilt models for specific documents' (Invoices, Receipts, Identity documents, Health insurance cards), 'General document analysis models' (OCR/Read, Layout), and 'Build document field extraction models from your own data' (Document field extraction, Document field extraction - neural and template, Document classification model). Each card includes a 'Try demo' button. The bottom right corner features a blue and purple abstract graphic.

Azure AI Foundry / contoso_a1 / AI Services / Vision + Document

All hubs + projects Project contoso_a1

← Vision + Document

Give your apps the ability to read text, analyze images, process documents and detect faces with technology like optical character recognition (OCR) and machine learning.

Integrate vision with generative AI

Content understanding Explore how you can transform content of any modality into task specific structured data. Try demo

Document field extraction Extract fields from documents and forms using a custom generative extraction model. Try demo

View all other vision capabilities

Document

Prebuilt models for specific documents

Invoices
Extract invoice ID, customer details, vendor details, ship to, bill to, total tax, subtotal, line items and more.
Try demo

Receipts
Extract time and date of the transaction, merchant information, amounts of taxes, totals and more.
Try demo

Identity documents
Extract expiration name, dates, machine readable zone, and more from passports and US driver licenses.
Try demo

Health insurance cards
Extract information such as, medical network, member name, and deductible.
Try demo

US Tax forms
Extract key information, etc. from US Tax forms. Supported US Tax forms are W-2, 1040, 1098, and 1099.
Try demo

US Mortgage forms
Extract key information from US Mortgage forms. Supported US Mortgage forms are 1003, 1008, and closing disclosure.
Try demo

Marriage certificates
Extract marriage certificate information, such as, couple's names, date of birth/marriage, occupations, addresses, nationality, parents, and officiant.
Try demo

Credit cards
Extract information such as, card number, cardholder name, due date, and bank information.
Try demo

Contracts
Extract title, signatory parties' information (including names, reference names, and address), and more from a contract.
Try demo

General document analysis models

OCR/Read
Extract printed and handwritten text from images and documents.
Try demo

Layout
Extract tables, check boxes, and text from forms and documents.
Try demo

Build document field extraction models from your own data

Document field extraction Extract fields from documents and forms using a custom generative extraction model.
Get started

Document field extraction - neural and template
Build a custom model to analyze forms of a specific type with same structure and fields.
Get started

Document classification model
Build a document classification model to efficiently and accurately split and classify various forms and documents.
Get started

Image may not reflect actual user interface.

Azure AI Vision

GPT-4 TURBO ENHANCEMENTS

Video prompting

Object grounding

Dense text accuracy

HIGH QUALITY PRE-TRAINED MODELS

OCR

FACE

IMAGE ANALYSIS

SPATIAL ANALYSIS

Printed Text Extraction

Detection

Tagging & Captioning

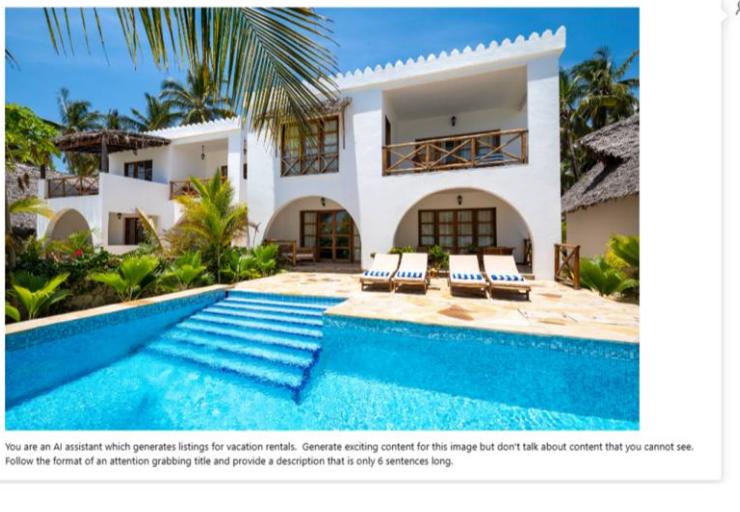
Person Tracking

Handwritten Extraction

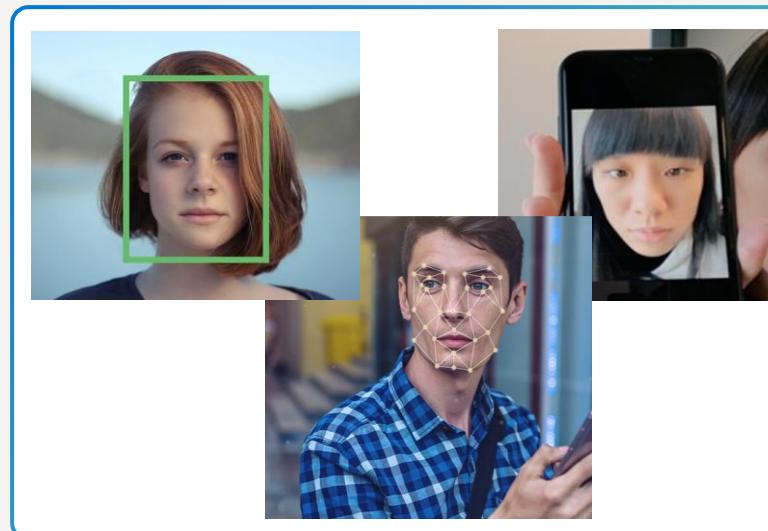
Recognition & Liveness

Content Moderation

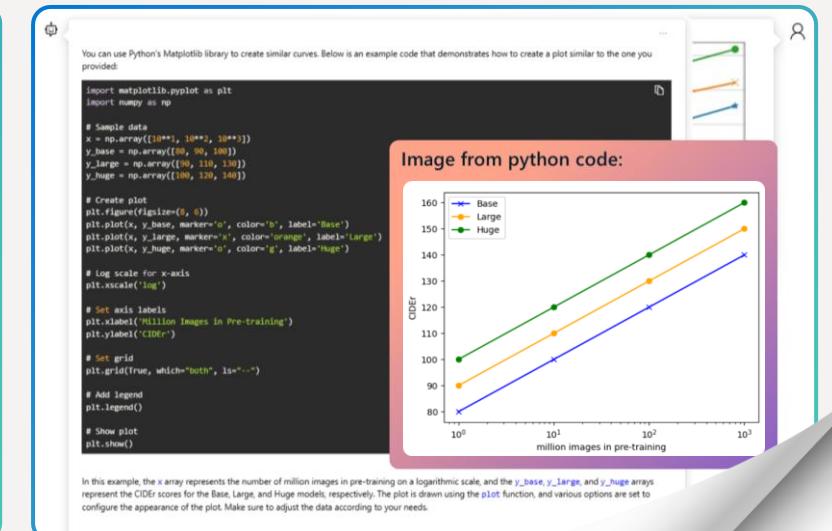
Object Tracking



Analyze objects in images and videos and provide detailed descriptions



Detect and recognize facial liveness



Convert images into code





Public Preview

Azure AI Content Understanding

Build with documents, audio, images, video and text

Streamline multi-modal
app development

Task specific structured
output for agentic workflows

Pre-built templates
to speed development

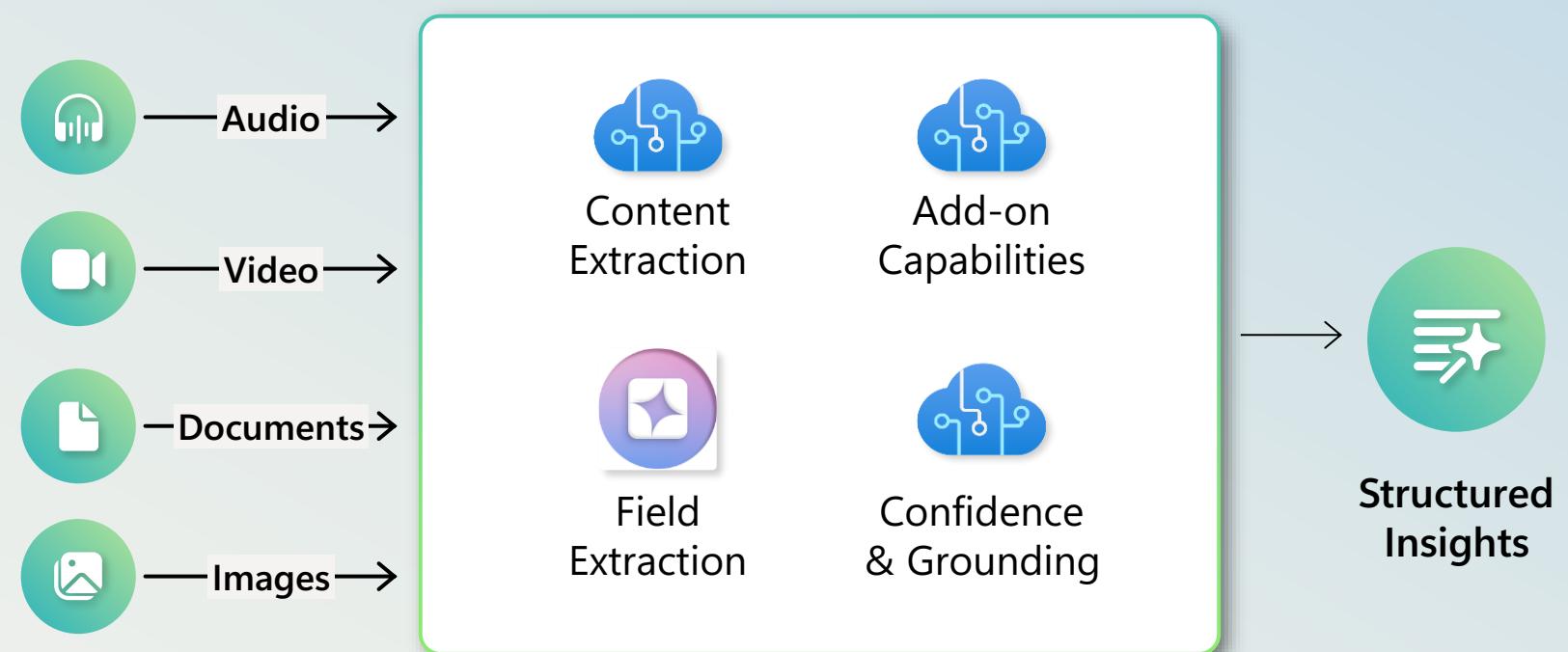
AI.Azure.com

Generate insights from documents, text, audio, images, and video

Differentiators:

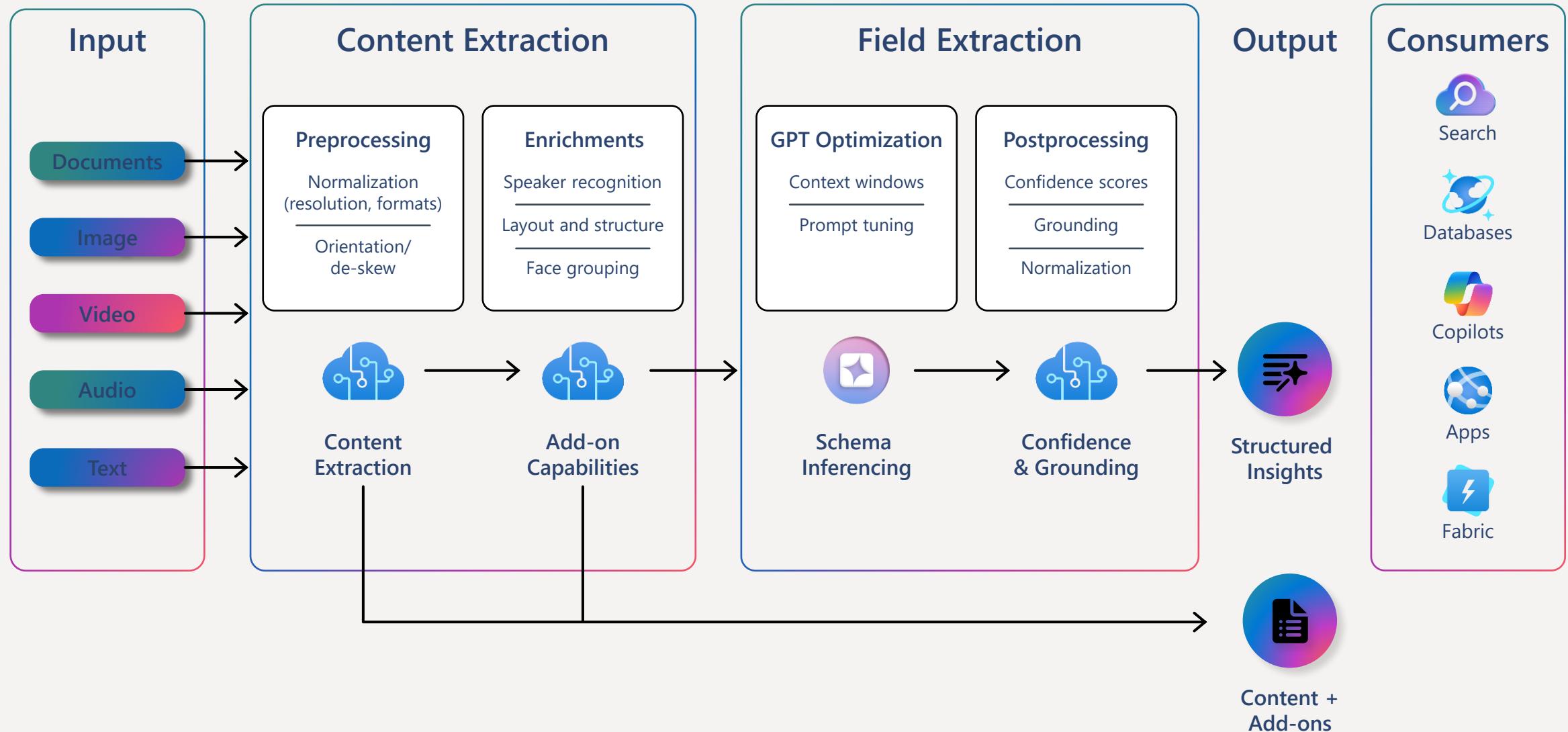
- Streamline multi-modal app development
- Task specific structured output for agentic workflows
- Pre-built templates to speed development

Azure AI Content Understanding



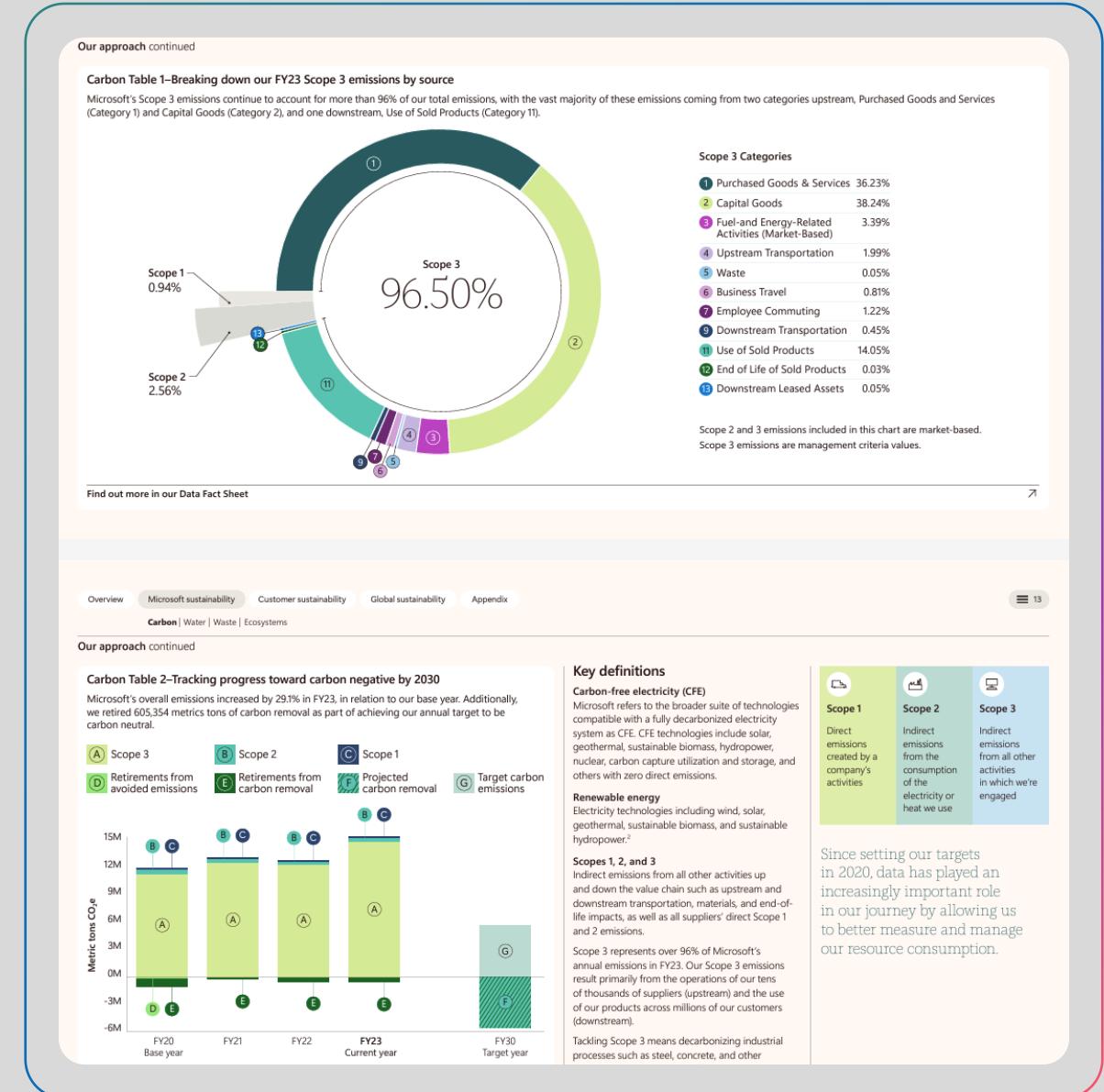


Components of Content Understanding



Understanding content from tables, charts and graphs

Content extraction for Document chat with Azure AI Search



Computer Vision

Visual Intelligence Made Easy

Easily customize your own state-of-the-art computer vision models that fit perfectly with your unique use case. Just bring a few examples of labeled images and let Custom Vision do the hard work.



Main Types of Computer Vision Algorithms

IMAGE CLASSIFICATION

What are my images about?



OBJECT DETECTION

Locate rectangular areas containing known objects in an image



SEMANTIC SEGMENTATION

Locate known objects in an image, at pixel level

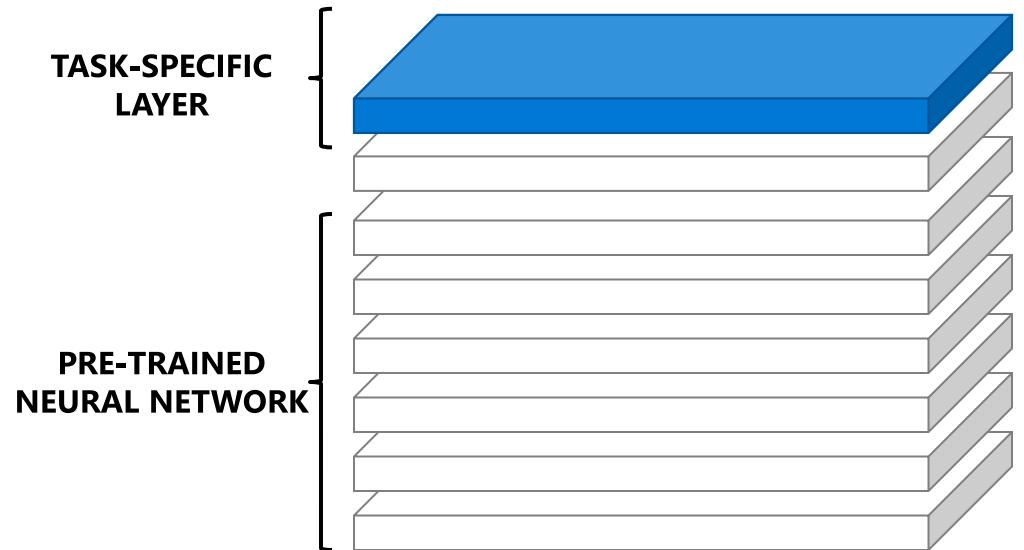


Standing on the Shoulder of Giants – Transfer Learning

By **leveraging existing pre-trained models**

it becomes possible to train highly performant models using a relatively small task-specific data set.

Microsoft Cognitive Services for Computer Vision are ready-to-use and designed around that approach. Custom models trained on ImageNet (AlexNet, Resnet...) are also good examples of foundations for transfer learning.



Azure OpenAI Service (AOAI)

Azure OpenAI Service empowers developers and businesses to harness the power of advanced AI technologies without the need for extensive AI expertise or infrastructure investment.

- Latest innovations with same-day access - Offering the latest models in the Azure OpenAI Service on the same-day with the same features.
- Widest Range of Deployment offerings - Standard, Provisioned and Batch available across 28 different regions, data zones and Global
- Your data is private and secure. Data residency at rest and at processing, and data protection from exfiltration.
- Safety is built-in with tools to responsibly build and implement AI.
- Enterprise Azure promises with private networking, managed identity, customer managed keys, azure monitoring, cost management, and more.
- Developer-Friendly Integration. Seamless integration with GitHub, Visual Studio and Copilot Studio.
- Choice of Models - Over 1800 models available to integrate into your application.

Discover models for your use case

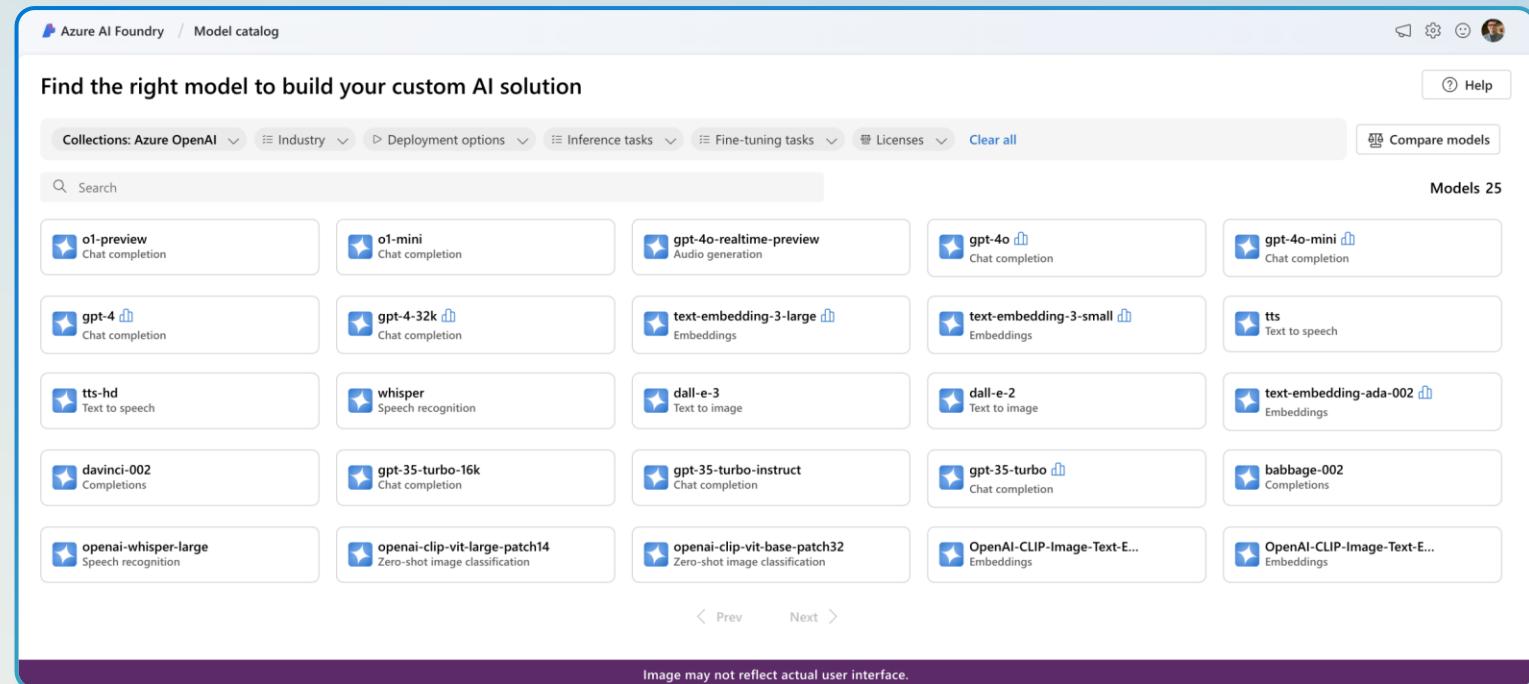


Image may not reflect actual user interface.



Leverage OpenAI's state-of-the-art models and Azure's cloud platform to build innovative AI-driven applications and solutions across various industries and domains.





OpenAI

Cutting-edge AI models
Pioneering research and advancements
Unlocking new creative possibilities



Azure AI Foundry

Secure data and built-in responsible AI
Global availability and flexibility
Developer-friendly integrations

Azure OpenAI Service

Connect your data | Customize your agents | Manage quota & deploy

Reasoning and
multimodal
text, image, audio

o1
o1-mini
GPT-4o
GPT-4o mini

Fine-tuning

GPT-4o
GPT-4o mini

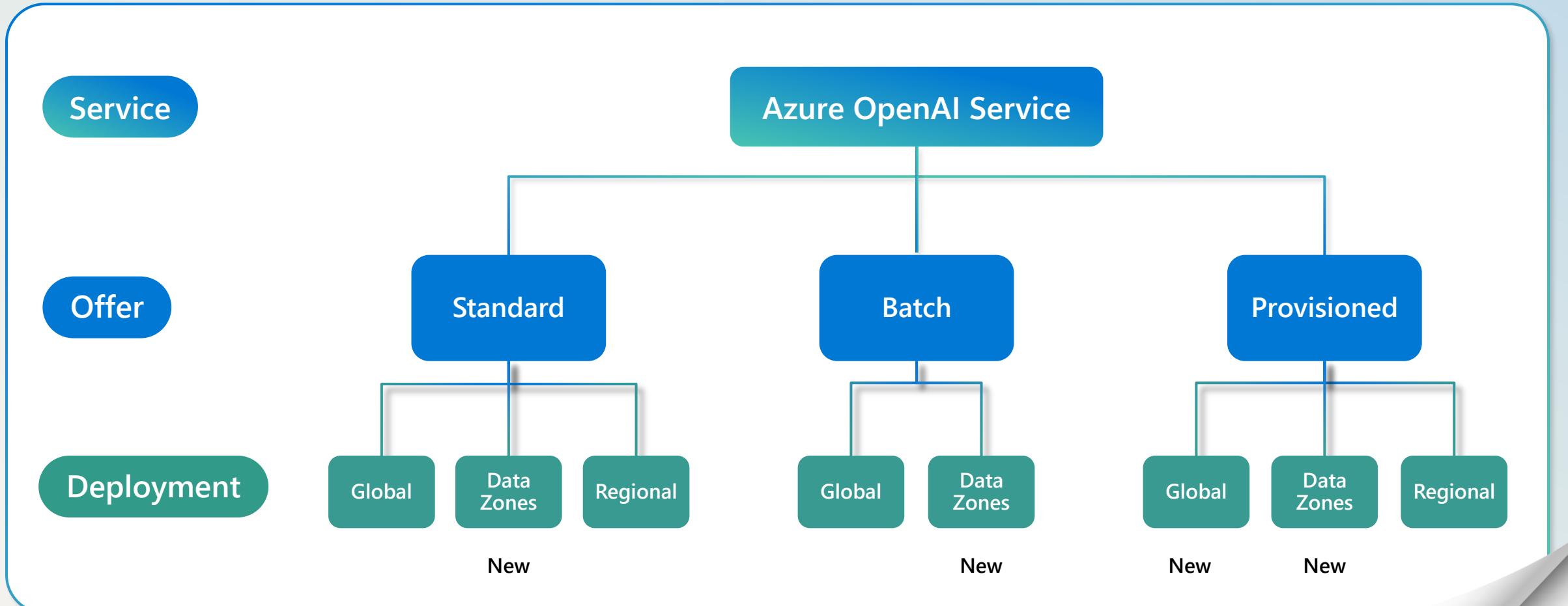
Images

DALL·E 3

Transcription &
translation

Whisper

Azure OpenAI Service Offerings



Azure OpenAI Service

GPT-4o-Realtime-Preview

Enable low-latency, multilingual speech-to-speech conversations

Real-time Audio
Interactions

Natural Speech
Generation

Multilingual
Capabilities

[Available in ai.azure.com](https://ai.azure.com)



Azure OpenAI Service

GPT-4o-Audio-Preview

Enable advanced audio processing and content generation

Native sentiment
analysis

Immersive Speech
Generation

Multilingual
Capabilities

Available in ai.azure.com



o1

Advanced reasoning

Solves complex problems

Enhances LLM capabilities:

- Logical problem-solving
- Complex math
- Scientific inquiries
- Extended conversations

o1-mini (in preview)

Smaller, faster model

Excels at:

- Instruction following
- Small context operations
- Code debugging

80% cheaper than o1-preview

**Text Based
Input/Output Only**

**Generates 6-10x
more tokens**

**Support for 128K
Context window**

**Longer response
time for reasoning**

Available in ai.azure.com



Azure OpenAI Service model guidance

o1

Multimodal model focused on advanced reasoning and solving complex problems, including math and science tasks.

o1-mini

Smaller, faster, and more cost-efficient version of o1. Performs well at instruction following and small context operations.

GPT-4o

A versatile, multimodal model for text and image processing, non-English languages, and vision tasks.

GPT-4o-mini

A cost-effective version of GPT-4o for lightweight applications.





Demo

Azure OpenAI



Safeguard with Trustworthy AI

Generative AI introduces new risks



Ungrounded
outputs & errors



Jailbreaks &
prompt injection
attacks



Harmful content
& code



Copyright
infringement



Manipulation
& human-like
behavior



Safeguard with Trustworthy AI

Secure your data, AI apps, and models

Build end-to-end verifiable privacy into your AI apps

Utilize built-in AI safety tooling

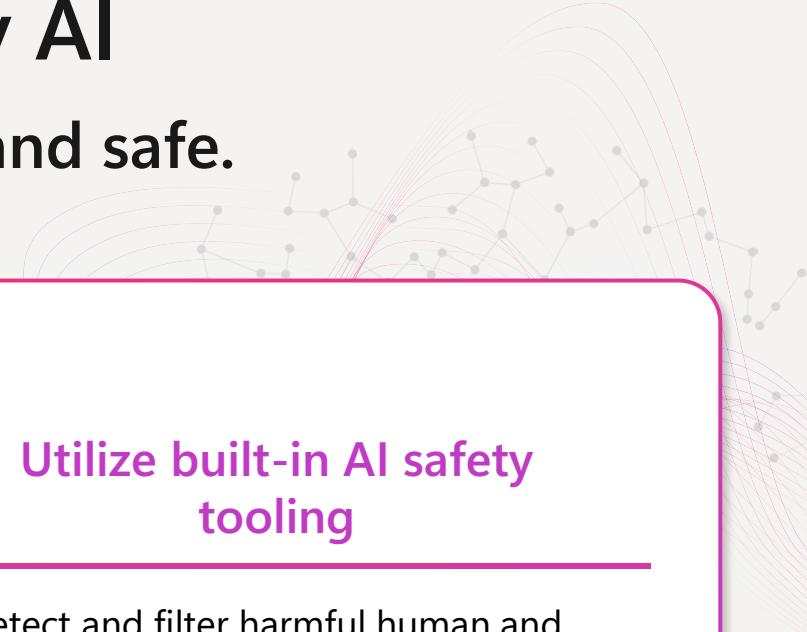
The screenshot shows the Azure AI Foundry interface with a network graph background. The main page title is "Create a new evaluation". On the left, there's a sidebar with navigation links like Home, Model catalog, Playgrounds, AI Services, Build and customize, Code, Prompt flow, Fine-tuning, Monitor performance, Tracing, Evaluation (which is selected), and Content filters. Under "Evaluation", there's a "My assets" section with options for Models + endpoints, Prompts, Data + indexes, and Web apps. The main content area has two sections: "Select metrics" and "Risk and safety metrics curated by Microsoft". The "Select metrics" section contains several checkboxes for metrics like Groundedness, Relevance, Coherence, Fluency, GPT similarity, F1 score, etc. The "Risk and safety metrics curated by Microsoft" section also contains checkboxes for metrics like Self-harm-related content, Hateful and unfair content, Violent content, Sexual content, Protected material, and Indirect attack. At the bottom, it says "Image may not reflect actual user interface."





Safeguard with Trustworthy AI

Use and build AI that is secure, private, and safe.



Secure your data, AI, apps, and models

Confidently design apps with technologies, templates, and governance best practices to help manage security risk and compliance.

Build end-to-end verifiable privacy into your AI apps

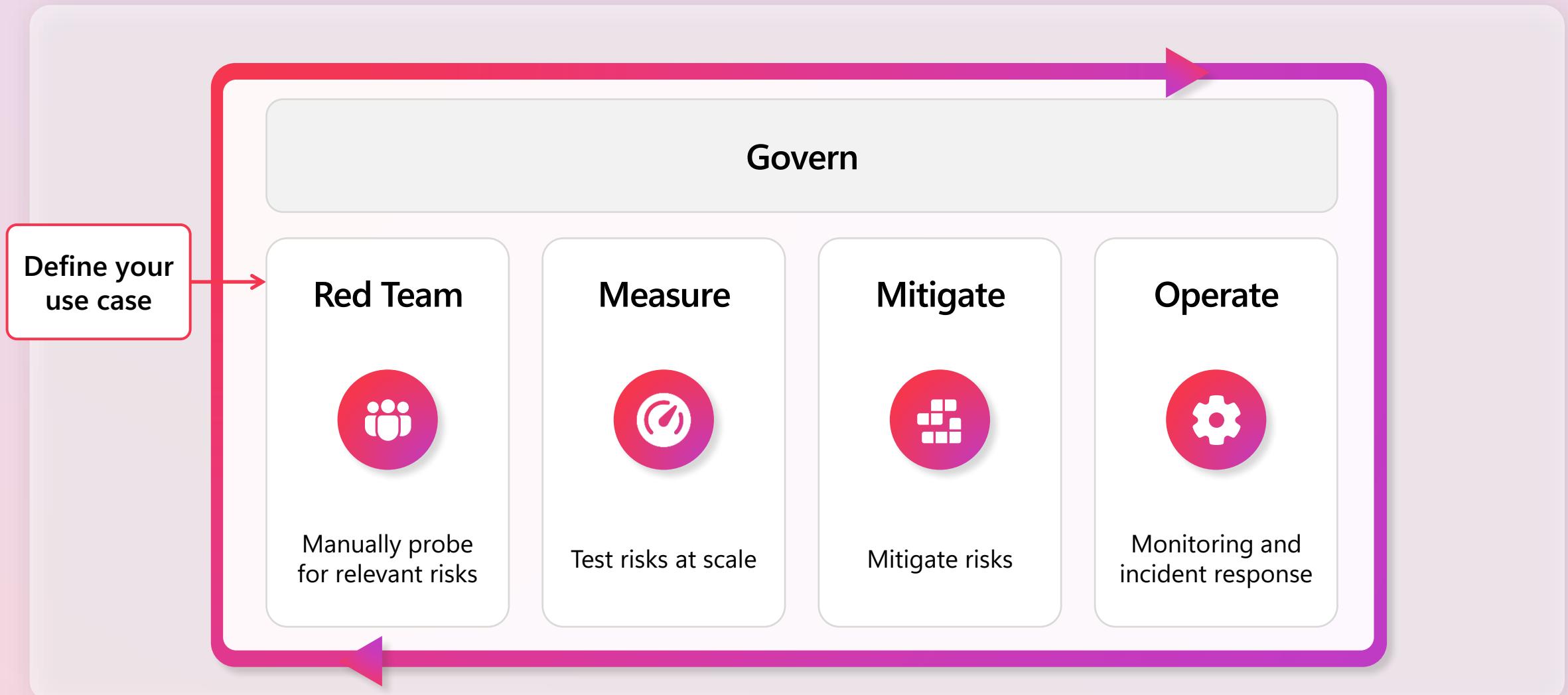
Protect your data at rest, in transit, and in use with confidential AI.

Utilize built-in AI safety tooling

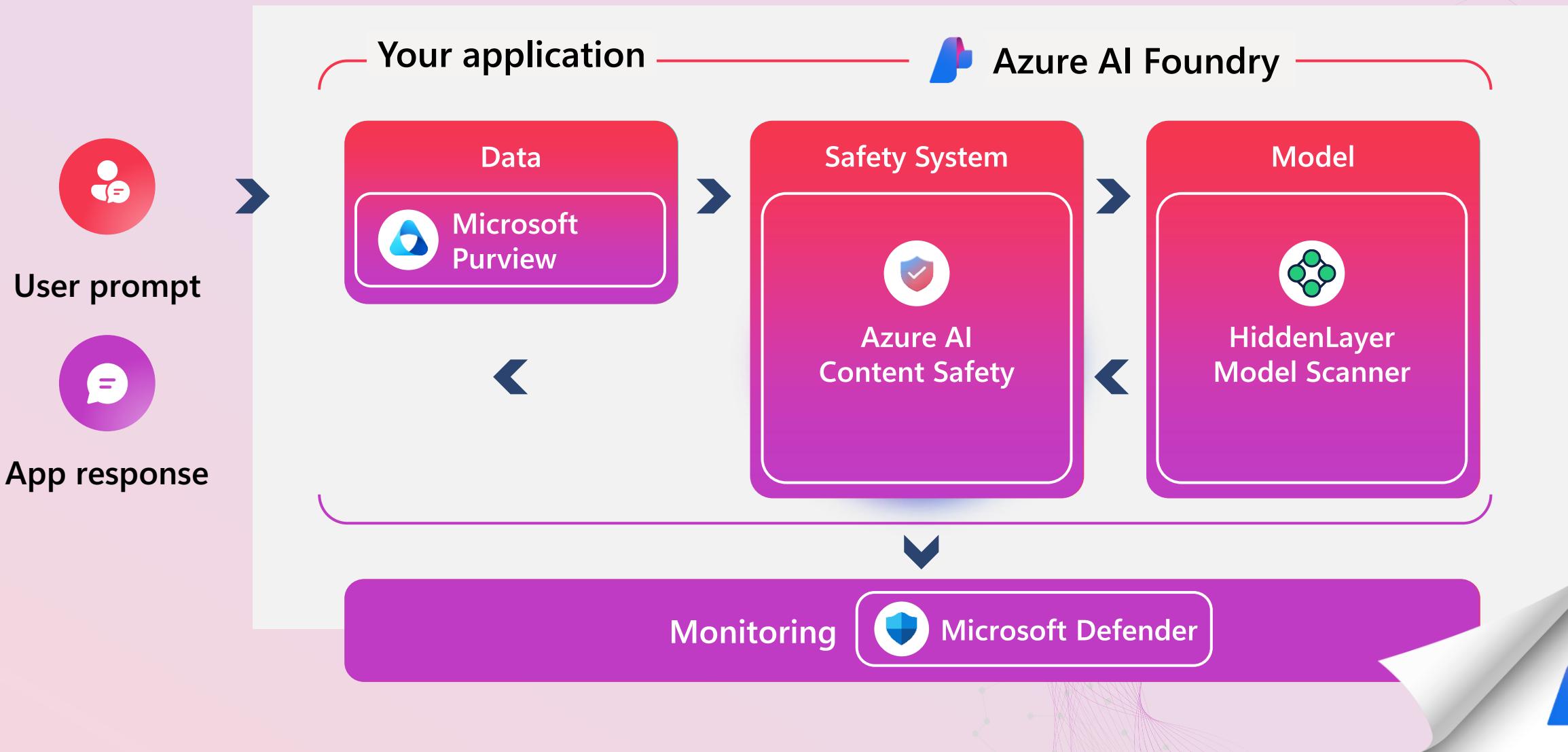
Detect and filter harmful human and AI-generated content, protect PII, and safeguard applications against prompt attacks, hallucinations, protected and owned materials, and bias.



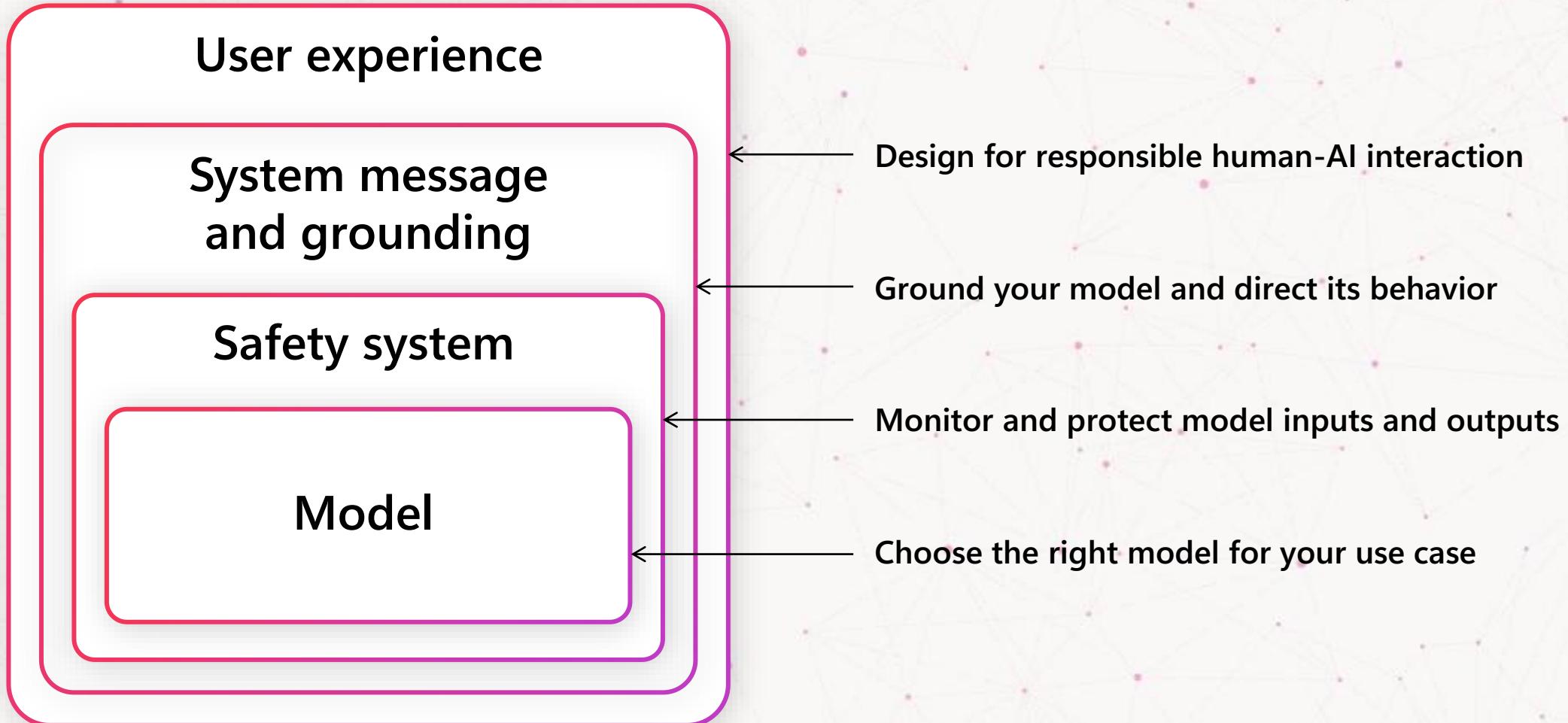
Responsible innovation is iterative



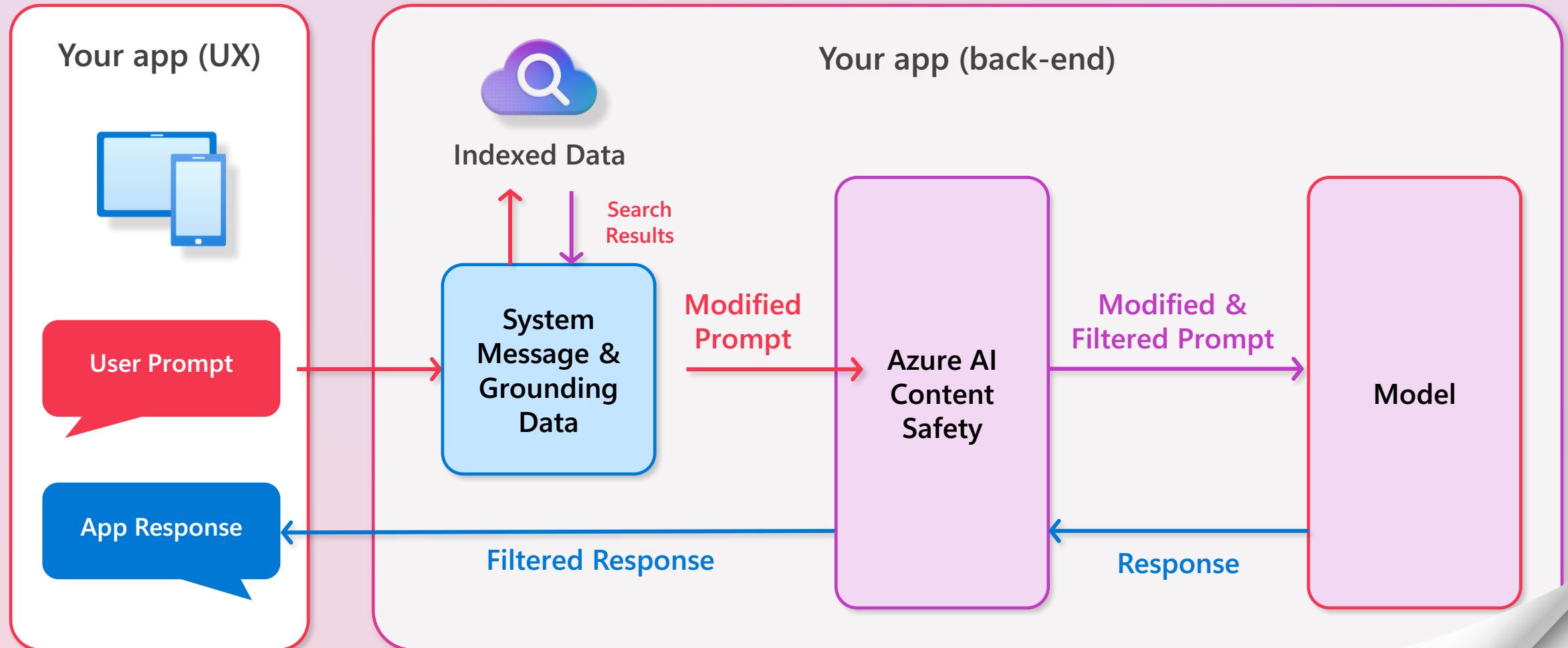
Built-in security and safety system



Risk mitigation layers



How these mitigations happen in real time



Responsible tooling for confidence



Multimodal filtering

Scan text, images and multi-media to identify, block and monitor **harmful content**



Customized systems

Create blocklists and **custom categories** to block entire topics, not just specific words



Prompt shielding

Identify and mitigate prompts that could expose you to **prompt injection attacks**



Protected materials

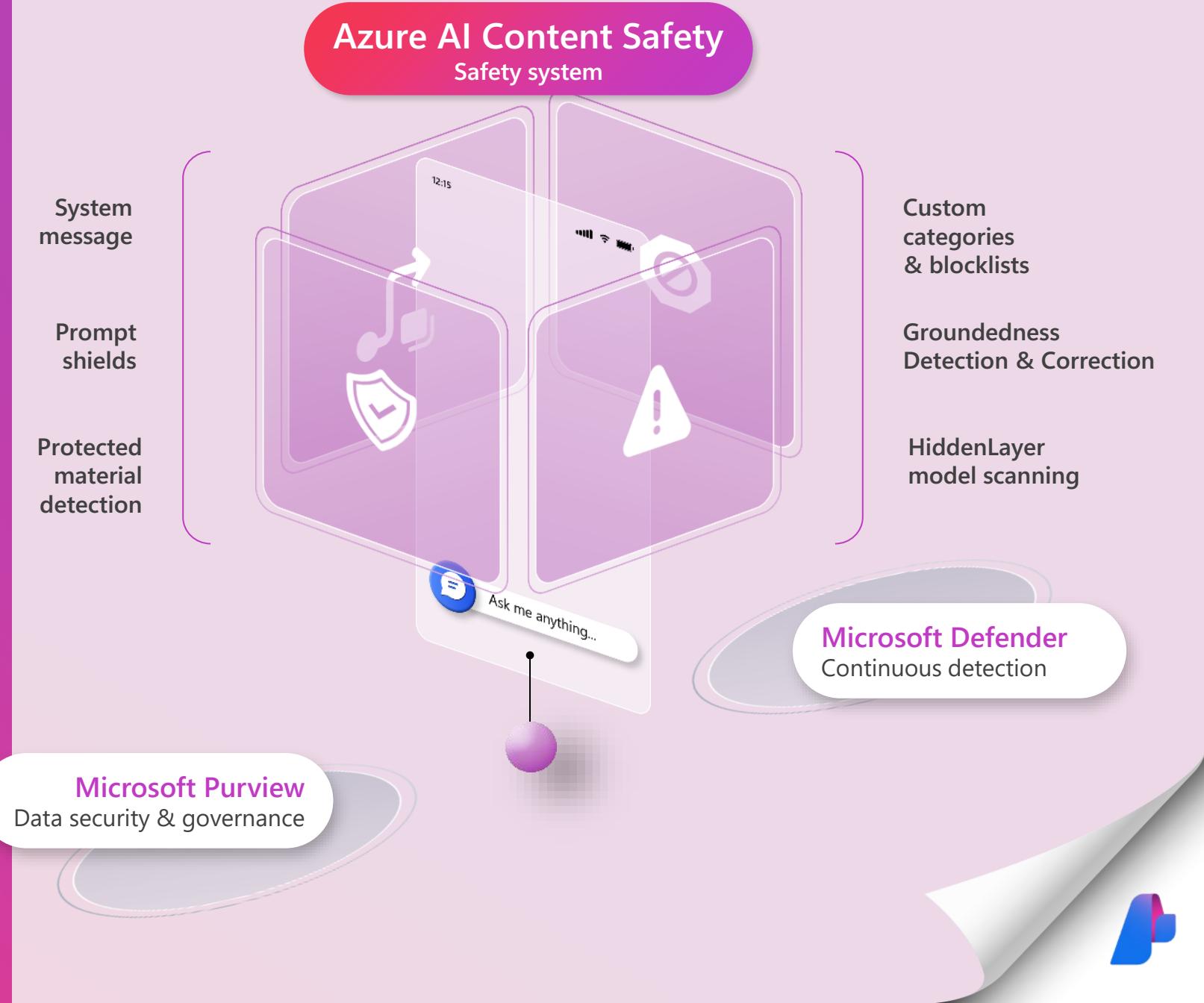
Avoid outputting **known or owned text** content with protected materials detection



Drive innovation with a built-in, comprehensive safety system

The Azure AI Foundry difference:

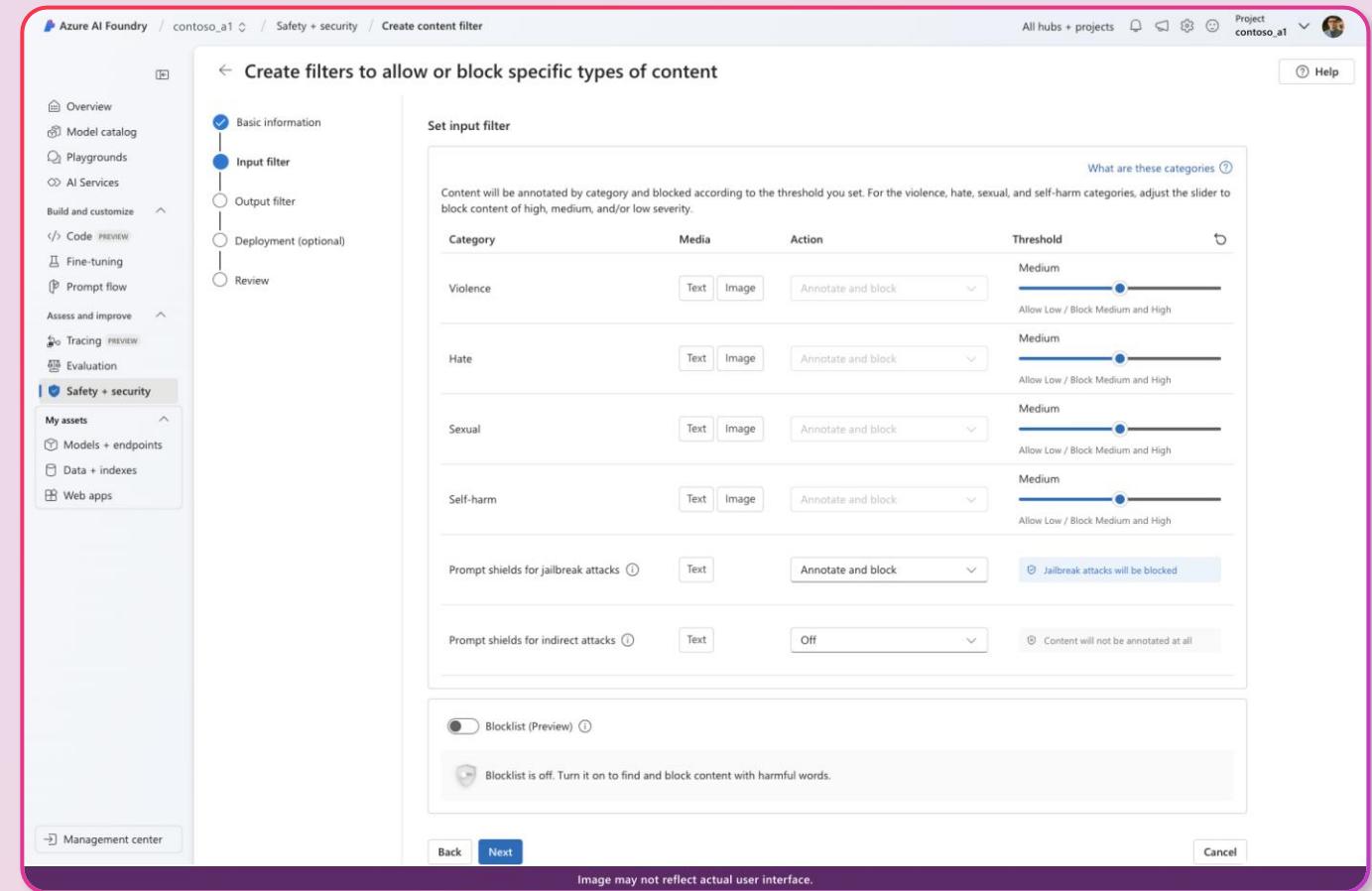
- Security and governance
- HiddenLayer model scanning
- Copyright commitment



Azure AI Content Safety

Azure AI Content Safety is a content moderation platform that uses AI to keep your content safe, creating better online experiences for everyone with powerful models that detect inappropriate content.

- Enable content harm filters to monitor inappropriate responses that may be hateful, sexual, violent, or lead to self-harm
- Customize content safety thresholds for each user type or individual
- Use Jailbreak Risk Detection to detect user prompts that provoke the GenAI model into breaking set rules
- Secure your data with Cross Prompt Injection Detection to defend against and identify potential Cross-Prompt Injection Attack (XPIA) attacks in input documents and large language model (LLM) conversations
- Identify text in language model output that matches known text context for protected materials
- Create user defined blocklists to manage content
- Insert built-in safety system messages in playground
- Customize your own categories for specific scenarios



Enhance user safety and data security with Azure AI Content Safety, providing tailored protection, real-time threat detection, and customizable controls.

Responsible by design

Because AI principles are not self-executing, we share our learnings and embed data-driven guardrails, guidance and best practices into Azure AI Foundry to help you operationalize trustworthy AI.

- Microsoft had early access to OpenAI models and gained valuable experience launching enterprise GenAI apps in the past two years—all built on Azure AI
- Microsoft has nearly 350 employees specializing in responsible AI at Microsoft, and we are investing to expand this number further
- Microsoft is a recognized leader in cloud platform services with highly secure, state-of-the-art Azure datacenters across 60+ announced regions
- Microsoft has committed to investing \$20 billion in cybersecurity over five years and we employ more than 8,500 security and threat intelligence experts
- Azure has one of the largest compliance certification portfolios in the industry and deep experience helping regulated industries and governments take advantage of AI technologies responsibly

Microsoft's Responsible AI principles



Fairness



Reliability & Safety



Privacy & Security



Inclusiveness



Transparency



Accountability



Microsoft first adopted our six AI principles in 2018, and they continue to drive our policy, research, and engineering investments.



Microsoft Copilot Copyright Customer Commitment

Copilot Copyright Customer Commitment ensures defense against IP infringement claims arising from Copilot services to enhance customer confidence.

- Microsoft defends customers against copyright claims, settling adverse judgements and settlements
- Microsoft protects output content generated by Copilot services, requiring adherence to built-in guardrails and content filters
- Microsoft address copyright, patent, trademark, and trade secrets claims, excluding claims unrelated to IP rights
- Microsoft requires use of content filters and other safety systems built into the product
- Microsoft helps facilitate innovation by defending against IP claims, fostering customer confidence in utilizing Copilot services for creative endeavors



The screenshot shows a Microsoft Licensing News article. The title is "Introducing the Microsoft Copilot Copyright". Below the title is the subtitle "Use Microsoft Copilot services with confidence." The date "September 7, 2023" is at the top left. The main text discusses the new benefit of intellectual property indemnity support for commercial Copilot services, starting October 1, 2023. It states that Microsoft will defend customers from IP infringement claims and pay for adverse judgements or settlements. The text also mentions that this commitment does not change Microsoft's position on intellectual property rights in its outputs. It highlights customer wants to harness generative AI technologies and work with confidence. Specifically, the commitment covers third-party IP claims based on copyright, patent, trademark, trade secrets, or right of publicity, but not trademark use in trade or commerce, defamation, false light, or other causes of action. It also covers the customer's use and distribution of output content generated by Copilot services, but not input data, modifications, or uses that infringe others' rights. Finally, it requires the customer to use content filters and safety systems.



Microsoft empowers businesses to leverage Copilot services confidently, mitigating legal risks and fostering innovation within established boundaries.



A professional office setting with four people. A woman in a black blazer and yellow skirt stands by a window, looking out. Two men are seated at a white conference table; one is writing in a notebook while the other looks on. A fourth person's back is to the camera on the left. Large windows provide a view of a city skyline.

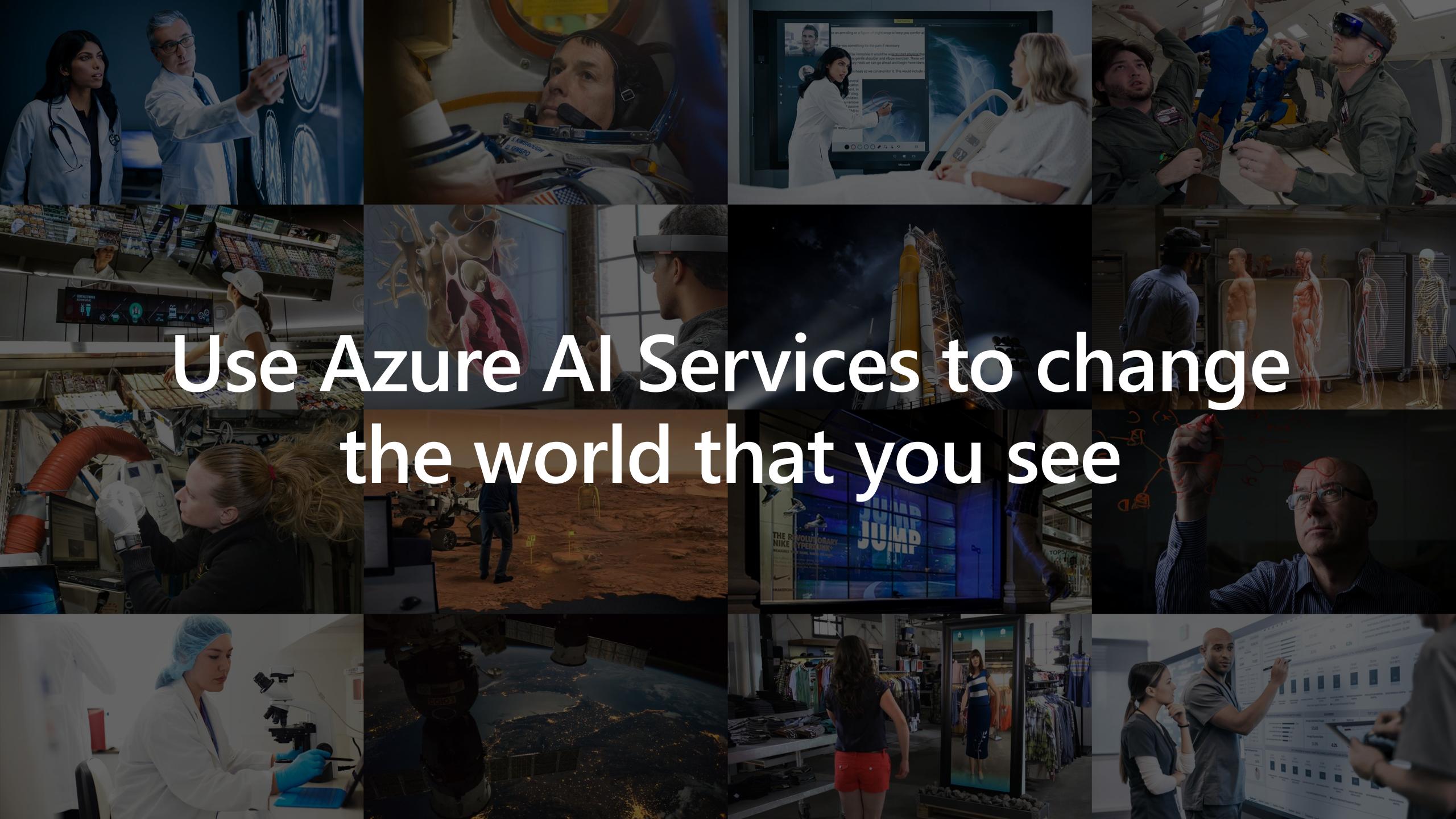
Making AI real for your Projects



You have more power at your
fingertips than ever before.

The collage of images includes:

- Two medical professionals examining a patient's heart on a chalkboard.
- An astronaut in a spacesuit inside a spacecraft module.
- A doctor and a patient in a hospital setting, with a Microsoft Surface tablet displaying medical data.
- A person in flight gear interacting with a Microsoft HoloLens device.
- A woman working in a food processing plant with a Microsoft Surface tablet.
- A man wearing a Microsoft HoloLens device, looking at a 3D anatomical model of a human skeleton.
- A large rocket launching from a launch pad at night.
- A man in a lab coat pointing at a whiteboard with a complex diagram.
- A female scientist in a lab coat using a microscope.
- A view from the International Space Station looking down at Earth at night.
- A woman in a clothing store interacting with a Microsoft Surface tablet displaying a virtual fitting room.
- A man and a woman in a business environment analyzing data on a large Microsoft Surface wall display.



The collage consists of 12 square images arranged in a grid, each depicting a different application of AI:

- Top-left: Two medical professionals in lab coats examining a chalkboard with anatomical drawings.
- Top-center: An astronaut in a spacesuit inside a spacecraft module.
- Top-right: A medical professional in a white coat interacting with a patient in a hospital bed via a video call on a tablet.
- Middle-left: A woman working at a food processing plant, with a screen showing a complex dashboard of data and video feeds.
- Middle-center: A man wearing a Microsoft HoloLens headset, looking at a 3D anatomical model of a human skeleton.
- Middle-right: A rocket launching from a launch pad at night.
- Bottom-left: A female scientist in a lab coat and hairnet using a microscope.
- Bottom-center: A view from the International Space Station looking down at Earth's city lights.
- Bottom-right: A man in a striped shirt pointing at a large wall-mounted screen displaying a complex network diagram.
- Second row, left: A woman in a white lab coat working in a laboratory.
- Second row, center: A woman in a white lab coat working in a clothing store, with a screen showing a virtual reality interface.
- Second row, right: Two medical professionals in scrubs reviewing a large wall of data on a screen.

Use Azure AI Services to change
the world that you see

