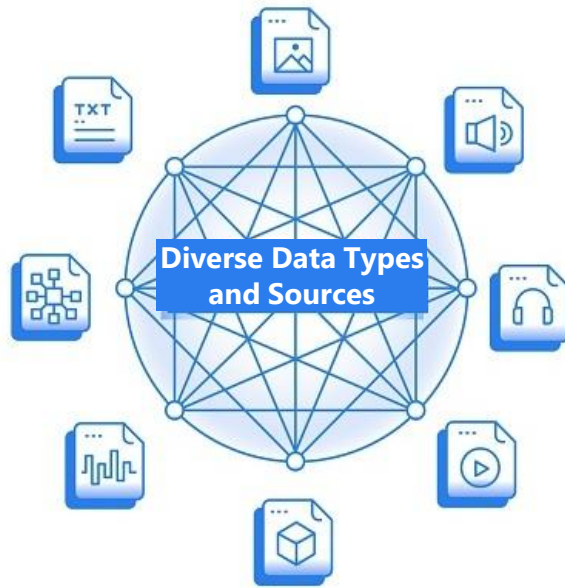


Content Understanding

Organizational challenges

Multimodal Data

ability to process and integrate data from multiple modalities such as text, images, videos, and audio



Comprehensive solution for handling diverse data types and converting them into structured, actionable insights

Unstructured Data

requires advanced analysis to extract meaningful information



Leverages AI to unlock insights from various data formats, making it especially suitable for complex scenarios where data is not neatly organized

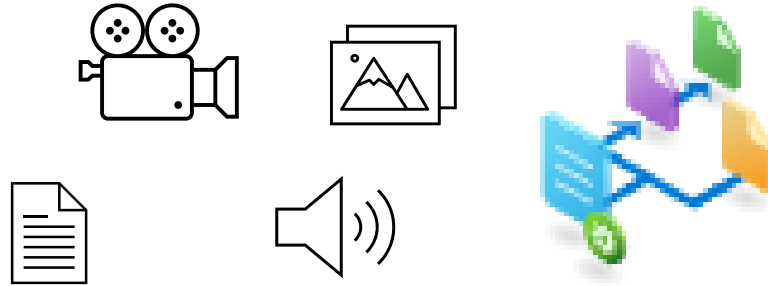
Comprehensive Analysis

to gain a holistic view across different channels



Content Understanding provides the tools to analyze and integrate insights from all available data sources

Azure AI Content Understanding



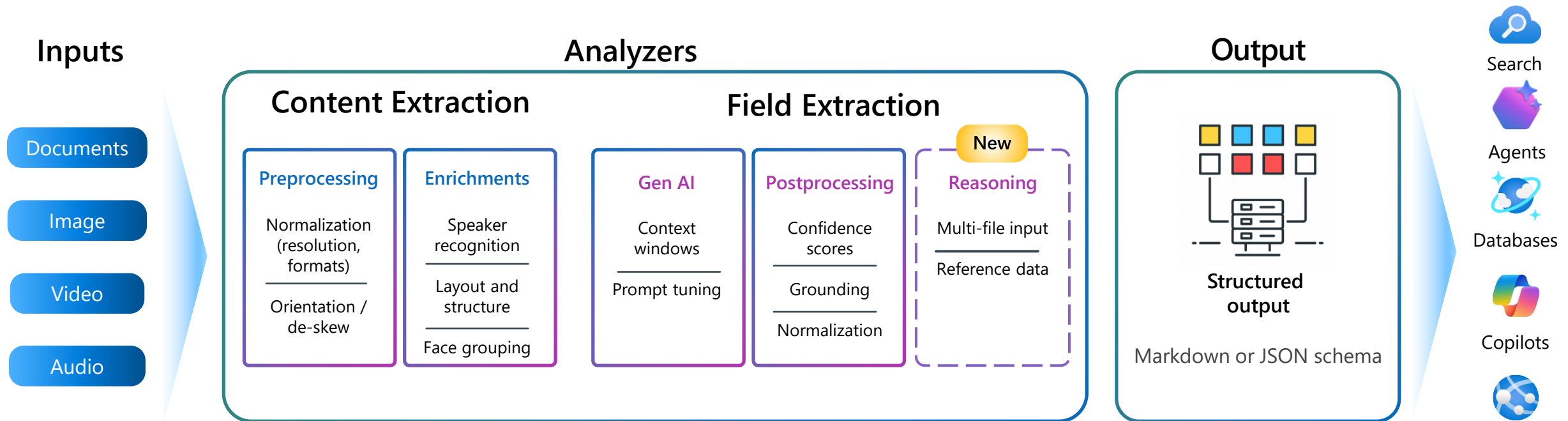
Unlock insights, enhance capabilities, improve efficiencies by leveraging unstructured data across diverse modalities

Build Comprehensive
Agentic Solutions

Feeds your Rag
Solutions

Powers Analytics and
Insights

Announcing a new Pro mode with reasoning



- ✓ Replaces multiple extraction calls followed by a human validation process with one simple operation
- ✓ Enable richer, more complex and adaptable task within agentic workflows

Benefits of Content Understanding



Multimodal Data Ingestion

- Ingest a range of modalities, including documents, images, audio, or video
- Use a range of Azure AI models to convert input data into a structured format
- Process and analyze through downstream services or applications



Information Extraction Schema

- Define schemas of extracted results
- Generate task-specific representations of the output, including insights, features, or summaries
- Build enterprise GenAI apps or agentic workflows to automate business processes through LLMs and RAG



Grounding and Confidence Scores

- Ensure schema-identified values are accurate and usable with GenAI tools
- Ground extracted information in underlying content
- Provide confidence scores to reduce human intervention
- Enable continuous improvement through user feedback

Improve data quality and accuracy

Streamline workflows and reduce costs

Customize output for any data type

Capabilities at a Glance



Documents

- OCR of printed or handwritten text
- Layout information
- Barcodes / Formulas
- Cross page tables
- Selection Marks
- Extract, Generate and classify field values
- **Content Classification**



Image

- Generate titles, descriptions, and summaries
- Classify elements



Audio

- Transcription
- **Multi-lingual Transcription**
- Diarization and language detection
- Identify speaker roles
- Generate conversation summaries and topics
- Determine sentiment



Video

- Transcription
- Shot detection
- Key frame extraction
- **Face Grouping (gated)**
- Face Detection (gated)
- Generate summaries
- Determine sentiment

Industry Scenarios



Enterprise

- **Contract Intelligence:** Extract clauses, obligations, and risks from legal documents and generate compliance checklists.
- **Invoice and purchasing:** Automate the extraction of key details from invoices and purchase orders, streamline approval workflows, and ensure compliance with financial regulations.
- **Supply Chain Optimization:** automating document-centric steps and providing real-time decision support across a web of information – orders, inventories, routes, and external factors like weather or port conditions.



Consumer Insights & Retail

- **Customer Feedback Analysis:** Aggregate and analyze reviews, customer calls, and social media posts to generate product improvement insights.
- **Visual Merchandising Optimization:** Analyze in-store images and videos to assess shelf placement, branding, and customer behavior.
- **Voice of the Customer:** Combine call center transcripts, emails, and survey data to generate customer satisfaction reports.



Government & Legal

- **E-Discovery Analysis:** Modern litigation and investigations must process multimedia evidence (e.g. police reports, witness audio, video evidence)
- **Administrative Process Automation (digitizing forms and records):** Automate information extraction from diverse forms and records, speeding up processes and improving efficiency.
- **Regulatory Compliance:** Analyze multimodal data for compliance violations and generate audit-ready reports.



Media & Entertainment

- **Media Asset Management and Content Indexing:** Studios, broadcasters, and content creators often need to organize and search through vast libraries of media. Tag and index audio-visual content to make it easily searchable.
- **Automated Highlights and Content Generation:** for long videos, such as sports games, AI can identify key moments, extract those segments and stitch them together to generate a draft highlight reel.

Agentic Solutions

RAG Solutions

Power Insights

Industry Scenarios



Health Care

- **Clinical Documentation Analysis:** Extract patient data from handwritten notes, scanned forms, and audio dictations; generate structured summaries.
- **Medical Imaging Insights:** Combine radiology images with patient records to generate diagnostic hypotheses or treatment suggestions.
- **Pharmacovigilance:** Monitor and summarize adverse event reports from text, audio calls, and scanned documents.



Marketing & Advertising

- **Personalized Content Generation and Dynamic Experiences:** agencies can feed in a brand's existing content (web pages, PDF brochures, product images, videos, audio testimonials, etc.) and then create new, tailored user experiences.
- **Advertising Analytics and Asset Optimization:** analyze ads and customer interactions to inform better marketing decisions, extracting successful elements from past campaigns to enhance future engagement and streamline asset organization.



Banking & Finance

- **Fraud Detection and Risk Management:** Multimodal AI improves fraud detection by integrating and identifying unusual patterns across diverse data sources, providing a holistic view for faster, more accurate fraud alerts and risk assessments.
- **Document Processing and Compliance Automation:** extract key information across many different document types regardless of format (loan applications, tax forms, insurance claims, regulatory filings, etc.)



Energy & Utilities

- **Integrated Asset Management and Resource Optimization:** unify diverse data streams for holistic decision-making, optimizing operations and safety in energy and utilities sectors by integrating sensor, image, and text data.
- **Predictive Maintenance and Grid Reliability:** enhances predictive maintenance by analyzing visual and sensor data, quickly identifying infrastructure risks to prevent outages and optimize grid reliability.

Agentic Solutions

RAG Solutions

Power Insights

Content Understanding Demos

Insurance Claims

Streamline insurance claims processing with automated content understanding.

Content Understanding

o1 Reasoning

Multi-modality

Select Insurance Demo

[Vimeo Live Demo](#)

Investigation Portal

Enhance investigation workflows with AI-powered content analysis.

Content Understanding

Embedding

Multi-modality

Select Investigation Demo

[Vimeo Live Demo](#)

Banking Loan Processing

Automate banking document processing and analysis with AI to accelerate loans.

Content Understanding

o1 Reasoning

Multi-modality

Select Banking Demo

[Vimeo Live Demo](#)

[Demo Landing Page: Azure AI in Action](#)

Reasoning on multimodal content for efficient agentic AI app building

Wednesday, May 21 | 2:00 PM - 3:00 PM Pacific Daylight Time

Duration 1 hour

Arch, 705 Pike, Level 6, Room 608 BRK169

Speakers:



[Ronak Chokshi](#) | [Microsoft](#)



[Tobias Fengler](#) | [ASC Technol...](#)



[Vinod Kurpad](#) | [Microsoft](#)



[Feras Siam](#) | [MediaKind](#)



Azure AI Content Understanding is designed for developers to streamline content processing tasks. It automates complex workflows and reduces manual intervention by delivering structured and reliable data tailored to specific goals. With advanced reasoning capabilities, it can now optimize processing performance, enabling efficient AI app creation and saving time, allowing developers to focus on building agentic AI apps quickly and effectively.

Content Understanding:



Try it in Azure AI Foundry:
<https://aka.ms/cu-landing>



Samples repository:
<https://aka.ms/cu-samples>



Documentation:
<https://aka.ms/cu-docs>



Pricing:
<https://aka.ms/cu-pricing>

Contacts:



Connect with the CU team
<https://aka.ms/cu-contact>



Connect with the DI team
<https://aka.ms/di-contact>

Solution Templates:



Conversation Knowledge Mining AI Solution:
<https://aka.ms/Conversational-Knowledge-Mining>



Azure Multimodal AI + LLM Processing
[Multimodal-ai-llm-processing](https://aka.ms/Multimodal-ai-llm-processing)