informatech

informatech CERTIFIED GLOBAL LEADERSHIP CONSULTANTS

DIGITAL INNOVATION AND TRANSFORMATION | DIT-020

# ChatGPT Advanced: Mastering Multimodal AI Integration


**UK**
+44 33 000 111 90
info@informatech.co.uk
https://informatech.uk
63-66 Hatton Garden Hatton Garden
EC1N 8LE , London

**NL**
+31 85 74 444 46
info@infomatech.nl
https://infomatech.nl
Waarderweg 50 - 2031PB
Haarlem - Netherlands

# Course content

## Why Attend

Artificial intelligence is rapidly evolving beyond text-based interactions into multimodal systems capable of processing and generating information across text, images, audio, video, and structured data. These developments are creating new opportunities for organisations to improve productivity, automate complex workflows, enhance customer experiences, and develop innovative digital services.

ChatGPT and related AI technologies enable professionals to interact with information through natural language while incorporating visual understanding, document analysis, voice interactions, data interpretation, and integrated business workflows. However, realising the full value of multimodal AI requires more than familiarity with conversational prompting.

Organisations must understand how different AI modalities work together, how to design effective multimodal prompts, how to integrate AI capabilities with enterprise systems, and how to evaluate the accuracy, reliability, security, and business value of AI-generated outputs.

Advanced applications also introduce challenges involving inconsistent model responses, hallucinations, sensitive information, system integration, computational costs, accessibility, and responsible AI governance. Successful implementation therefore requires a structured approach combining technical understanding, practical experimentation, workflow design, and organisational controls.

The ChatGPT Advanced: Mastering Multimodal AI Integration course provides participants with practical knowledge of advanced ChatGPT capabilities and multimodal AI integration techniques.

Participants explore multimodal prompting, image and document understanding, speech-based interactions, video analysis concepts, structured data processing, API integration, workflow automation, and the development of AI-enabled business applications.

Through practical demonstrations, guided exercises, and applied implementation workshops, participants develop the ability to design, evaluate, and deploy multimodal AI solutions that address real organisational challenges.

### Professional Problems This Course Is Designed to Solve

• Limited use of AI beyond text generation: Expanding ChatGPT applications to incorporate visual information, documents, audio, video, and structured business data.

• Inconsistent results from complex AI prompts: Developing structured multimodal prompting techniques that improve instruction clarity, contextual understanding, and output reliability.

# Course content

## Why Attend

• Disconnected AI tools and business systems: Designing integration workflows that connect AI capabilities with organisational applications, information sources, and operational processes.

• Time-consuming document and information processing: Applying multimodal AI to analyse reports, extract information, interpret visual materials, and summarise complex content.

• Difficulty developing practical AI applications: Understanding the architecture, implementation requirements, and evaluation methods for multimodal AI-powered solutions.

• Concerns about AI accuracy and information security: Establishing validation, human oversight, privacy safeguards, and responsible AI controls for multimodal workflows.

• Unclear return on AI investments: Identifying high-value use cases and measuring improvements in efficiency, quality, user experience, and operational performance.

## Course Methodology

The programme combines instructor-led discussions, live demonstrations, hands-on ChatGPT exercises, multimodal prompting workshops, integration design activities, and practical AI application development scenarios.

Participants work with realistic business materials, including documents, spreadsheets, images, diagrams, audio transcripts, and illustrative video content, to explore how multimodal AI can support analysis, communication, decision-making, and workflow automation.

Technical demonstrations introduce API-based integration concepts, structured outputs, data processing workflows, and application architecture. Exercises can be adapted to participants' technical experience, with conceptual design activities for business professionals and optional implementation tasks for technical participants.

The programme emphasises practical experimentation, responsible AI use, output verification, and the development of reusable multimodal AI workflows.

## Course Objectives

# Course content

## Course Objectives

By the end of this course, participants will be able to:

- Explain the principles and capabilities of multimodal generative AI systems
- Identify appropriate business applications for text, image, audio, video, and structured data processing
- Develop advanced prompts combining multiple information modalities
- Apply ChatGPT to complex document and visual information analysis
- Design effective speech-based and conversational AI interactions
- Understand video interpretation workflows and their technical limitations
- Integrate multimodal AI capabilities with business applications through APIs and automation tools
- Develop structured AI workflows using external information and business data
- Evaluate multimodal AI outputs for accuracy, consistency, relevance, and reliability
- Apply privacy, security, accessibility, and responsible AI principles
- Assess the operational and financial value of multimodal AI applications
- Design a practical multimodal AI integration solution for an organisational use case

## Target Audience

This course is suitable for:

- Digital transformation managers and specialists
- Artificial intelligence and innovation professionals
- IT managers and enterprise technology specialists
- Business analysts and process improvement professionals
- Software developers and application integration specialists
- Data analysts and business intelligence professionals
- Product managers and digital service designers

# Course content

## Target Audience

• Marketing, communication, and content development professionals

• Customer experience and service automation managers

• Business leaders responsible for AI adoption and digital innovation

## Target Competencies

Participants will develop competencies in:

• Multimodal AI capabilities and application design

• Advanced ChatGPT prompting techniques

• Visual and document intelligence

• Audio and conversational AI workflows

• Video and multimodal information analysis

• AI API integration and workflow automation

• Structured data processing and AI orchestration

• Multimodal output evaluation and validation

• Responsible AI, privacy, and security governance

• AI solution development and business value assessment

## Course outline

### Day 1: Advanced ChatGPT Capabilities and Multimodal AI Foundations

• Understanding the evolution of generative AI from text-based language models to multimodal systems that process text, images, audio, video, and structured information

• Exploring the capabilities, limitations, and practical differences between multimodal input processing, reasoning, and content generation

# Course content

## Course outline

• Examining advanced ChatGPT features, model selection considerations, context management, and available multimodal interaction methods

• Developing advanced prompting strategies using role definition, task decomposition, contextual information, examples, output constraints, and iterative refinement

• Understanding how to combine textual instructions with images, documents, tables, and other inputs to improve task execution and contextual interpretation

• Practical workshop: Designing and testing advanced multimodal prompts for business analysis, visual interpretation, document review, and structured information extraction

### Day 2: Image Intelligence, Document Understanding and Visual AI Applications

• Applying multimodal AI to interpret images, diagrams, charts, screenshots, and visual business information

• Analysing complex documents containing text, tables, figures, financial information, and mixed visual layouts

• Designing document intelligence workflows for classification, summarisation, information extraction, comparison, and question answering

• Exploring image generation and editing workflows, including prompt design, creative iteration, visual consistency, and quality evaluation

• Understanding visual AI limitations involving image resolution, ambiguous content, numerical accuracy, document structure, and unsupported visual assumptions

• Hands-on exercise: Developing a multimodal document analysis workflow that extracts information, evaluates visual content, generates structured findings, and applies human verification

### Day 3: Audio, Voice, Video and Real-Time Multimodal Interactions

• Understanding speech recognition, speech synthesis, audio understanding, and conversational AI capabilities within multimodal systems

• Designing voice-enabled workflows for customer interactions, meeting assistance, language support, accessibility, and internal knowledge retrieval

# Course content

## Course outline

• Exploring audio transcription, speaker-related considerations, summarisation, and extraction of action items from spoken information

• Understanding video analysis concepts, including frame interpretation, temporal context, scene understanding, and the challenges of processing extended video sequences

• Examining real-time multimodal interaction architectures, latency considerations, session management, user experience, and error handling

• Practical workshop: Designing a voice-enabled assistant and a video or audio analysis workflow with clear inputs, outputs, validation procedures, and operational safeguards

### Day 4: Multimodal AI Integration, APIs and Intelligent Workflow Automation

• Understanding multimodal AI integration architecture, including application interfaces, API requests, model inputs, structured responses, and external system connections

• Exploring the use of OpenAI APIs and relevant development tools to process supported combinations of text, images, audio, and other media

• Designing workflows that connect AI services with document repositories, business applications, databases, and automation platforms

• Applying structured outputs, function calling concepts, retrieval-augmented generation, and workflow orchestration to support reliable AI-enabled processes

• Evaluating integration challenges involving authentication, access permissions, data transfer, error handling, latency, scalability, and operational costs

• Integration workshop: Designing an end-to-end multimodal AI workflow that receives business information, processes multiple input types, produces structured outputs, and routes results for human review

### Day 5: Responsible Multimodal AI, Performance Evaluation and Enterprise Deployment

• Establishing evaluation methods for multimodal AI applications covering accuracy, relevance, completeness, consistency, latency, usability, and business effectiveness

# Course content

## Course outline

• Identifying multimodal AI risks involving hallucinations, prompt injection, misleading visual interpretation, sensitive information exposure, intellectual property, and inappropriate automation

• Applying responsible AI principles through human oversight, access controls, data minimisation, output validation, transparency, and documented governance procedures

• Developing deployment strategies covering pilot selection, technical readiness, stakeholder engagement, user training, operational monitoring, and continuous improvement

• Assessing multimodal AI business cases through implementation costs, productivity improvements, quality indicators, risk reduction, and return on investment

• Capstone exercise: Developing and presenting a Comprehensive Multimodal AI Integration Solution incorporating a defined business use case, advanced prompting strategy, multimodal processing workflow, integration architecture, evaluation criteria, security controls, and implementation roadmap

# Seminar dates

## Available seminar dates

Live dates and pricing for ChatGPT Advanced: Mastering Multimodal AI Integration generated from the course details page.

| Date | Location | Format | Fee |
|---|---|---|---|
| Dates on request | Venue on request | Classroom | **Contact us** |

| | |
|---|---|
| **Live online option** | Online delivery is available at €1,850.-. |