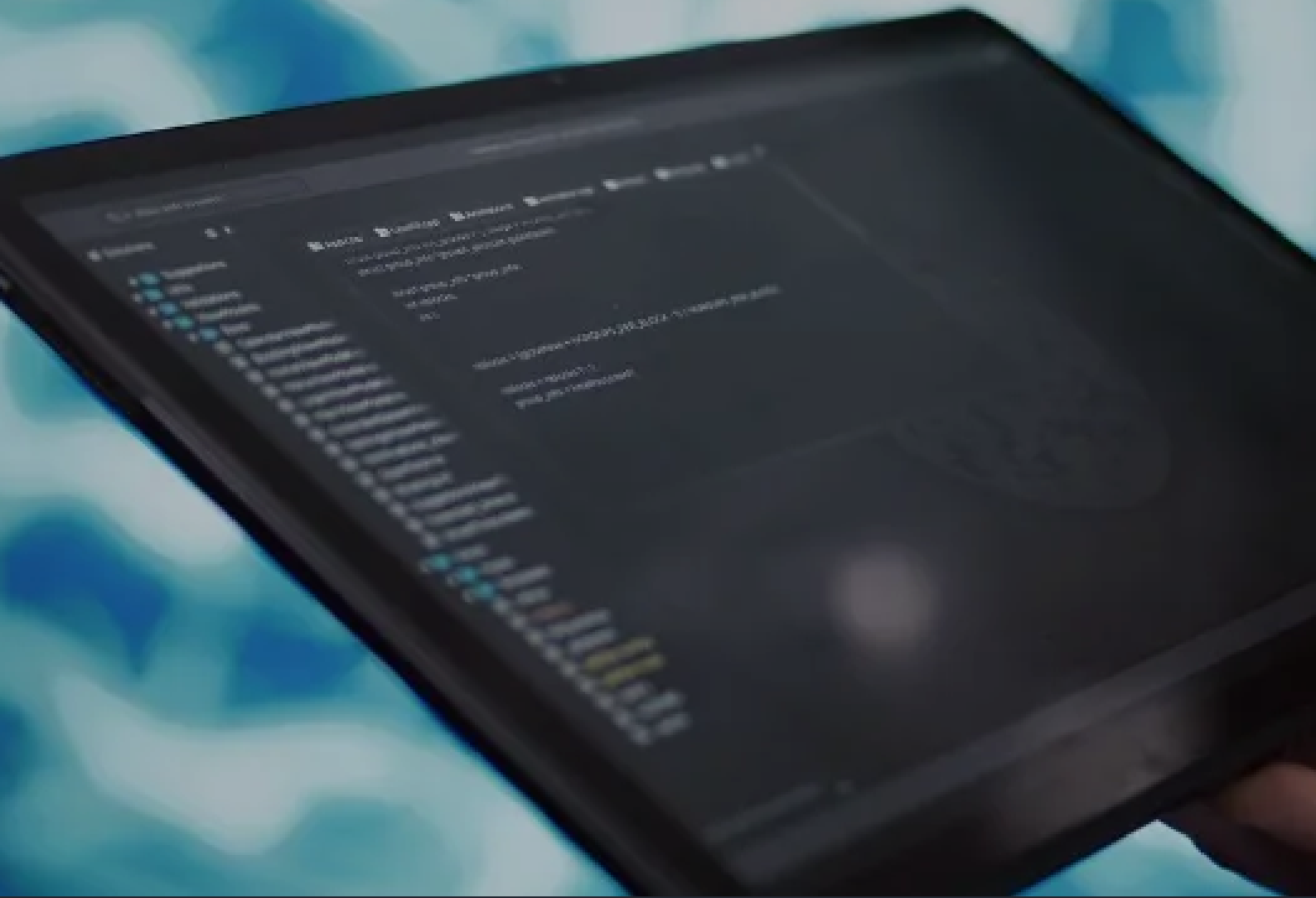




DIGITAL INNOVATION AND TRANSFORMATION | DIT-020

ChatGPT Advanced: Mastering Multimodal AI Integration

Contact

+31 85 7444446
info@aractech.com
<https://aractech.com>

Address

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.

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

- 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

Course content

- Design effective speech-based and conversational AI interactions
- Understand video interpretation workflows and their technical limitations

| Target Audience

- This course is suitable for:
- Digital transformation managers and specialists
- Artificial intelligence and innovation professionals
- IT managers and enterprise technology specialists

Course outline

Detailed course outline

Day-by-day outline for ChatGPT Advanced: Mastering Multimodal AI Integration.

Day 1 - Advanced ChatGPT Capabilities and Multimodal AI Foundations

- 09:00-09:30 | Understanding the evolution of generative AI from text-based language models to multimodal systems that process text, images, audio, video, and structured information
- 09:30-10:30 | Exploring the capabilities, limitations, and practical differences between multimodal input processing, reasoning, and content generation
- 10:30-11:00 | Examining advanced ChatGPT features, model selection considerations, context management, and available multimodal interaction methods
- 11:00-11:30 | Coffee break
- 11:30-12:30 | Developing advanced prompting strategies using role definition, task decomposition, contextual information, examples, output constraints, and iterative refinement
- 12:30-13:00 | Understanding how to combine textual instructions with images, documents, tables, and other inputs to improve task execution and contextual interpretation
- 13:00-14:00 | Practical workshop: Designing and testing advanced multimodal prompts for business analysis, visual interpretation, document review, and structured information extraction

Course outline

Detailed course outline

Day-by-day outline for ChatGPT Advanced: Mastering Multimodal AI Integration.

Day 2 - Image Intelligence, Document Understanding and Visual AI Applications

- 09:00-09:30 | Applying multimodal AI to interpret images, diagrams, charts, screenshots, and visual business information
- 09:30-10:00 | Analysing complex documents containing text, tables, figures, financial information, and mixed visual layouts
- 10:00-11:00 | Designing document intelligence workflows for classification, summarisation, information extraction, comparison, and question answering
- 11:00-11:30 | Coffee break
- 11:30-12:30 | Exploring image generation and editing workflows, including prompt design, creative iteration, visual consistency, and quality evaluation
- 12:30-13:00 | Understanding visual AI limitations involving image resolution, ambiguous content, numerical accuracy, document structure, and unsupported visual assumptions
- 13:00-14:00 | Hands-on exercise: Developing a multimodal document analysis workflow that extracts information, evaluates visual content, generates structured findings, and applies human verification

Course outline

Detailed course outline

Day-by-day outline for ChatGPT Advanced: Mastering Multimodal AI Integration.

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

- 09:00-09:34 | Understanding speech recognition, speech synthesis, audio understanding, and conversational AI capabilities within multimodal systems
- 09:34-10:41 | Designing voice-enabled workflows for customer interactions, meeting assistance, language support, accessibility, and internal knowledge retrieval
- 10:41-11:00 | Exploring audio transcription, speaker-related considerations, summarisation, and extraction of action items from spoken information
- 11:00-11:30 | Coffee break
- 11:30-11:45 | Exploring audio transcription, speaker-related considerations, summarisation, and extraction of action items from spoken information (continued)
- 11:45-12:19 | Understanding video analysis concepts, including frame interpretation, temporal context, scene understanding, and the challenges of processing extended video sequences
- 12:19-12:53 | Examining real-time multimodal interaction architectures, latency considerations, session management, user experience, and error handling
- 12:53-14:00 | Practical workshop: Designing a voice-enabled assistant and a video or audio analysis workflow with clear inputs, outputs, validation procedures, and operational safeguards

Course outline

Detailed course outline

Day-by-day outline for ChatGPT Advanced: Mastering Multimodal AI Integration.

Day 4 - Multimodal AI Integration, APIs and Intelligent Workflow Automation

- 09:00-10:00 | Understanding multimodal AI integration architecture, including application interfaces, API requests, model inputs, structured responses, and external system connections
- 10:00-10:30 | Exploring the use of OpenAI APIs and relevant development tools to process supported combinations of text, images, audio, and other media
- 10:30-11:00 | Designing workflows that connect AI services with document repositories, business applications, databases, and automation platforms
- 11:00-11:30 | Coffee break
- 11:30-12:00 | Designing workflows that connect AI services with document repositories, business applications, databases, and automation platforms (continued)
- 12:00-12:30 | Applying structured outputs, function calling concepts, retrieval-augmented generation, and workflow orchestration to support reliable AI-enabled processes
- 12:30-13:00 | Evaluating integration challenges involving authentication, access permissions, data transfer, error handling, latency, scalability, and operational costs
- 13:00-14:00 | 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

Course outline

Detailed course outline

Day-by-day outline for ChatGPT Advanced: Mastering Multimodal AI Integration.

Day 5 - Responsible Multimodal AI, Performance Evaluation and Enterprise Deployment

- 09:00-09:34 | Establishing evaluation methods for multimodal AI applications covering accuracy, relevance, completeness, consistency, latency, usability, and business effectiveness
- 09:34-10:08 | Identifying multimodal AI risks involving hallucinations, prompt injection, misleading visual interpretation, sensitive information exposure, intellectual property, and inappropriate automation
- 10:08-10:42 | Applying responsible AI principles through human oversight, access controls, data minimisation, output validation, transparency, and documented governance procedures
- 10:42-11:00 | Developing deployment strategies covering pilot selection, technical readiness, stakeholder engagement, user training, operational monitoring, and continuous improvement
- 11:00-11:30 | Coffee break
- 11:30-12:19 | Developing deployment strategies covering pilot selection, technical readiness, stakeholder engagement, user training, operational monitoring, and continuous improvement (continued)
- 12:19-12:53 | Assessing multimodal AI business cases through implementation costs, productivity improvements, quality indicators, risk reduction, and return on investment
- 12:53-14:00 | 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
------	----------	--------	-----

No seminar dates are currently available.