



# ChatGPT Document and Image Information Extraction Course

18 – 29 January 2027  
Tokyo

## ChatGPT Document and Image Information Extraction Course

**Ref:** 299\_125354 **Date:** 18 – 29 January 2027 **Location:** Tokyo **Fees:** 15000 Euro

### Overview

ChatGPT Document and Image Information Extraction Course is a ten-day practitioner course for records, operations, audit, and knowledge teams, who leave with a Reviewed Information Extraction Specification. Participants practice structured information extraction from business documents and images, using extraction schemas, document quality assessment, source provenance, confidence and exception flags, and human validation. The course addresses mixed layouts, unclear evidence, sensitive content, and inconsistent outputs. Agile Leaders Training Center provides training in document and image information extraction with ChatGPT.

### Who Should Attend

- Records teams responsible for classifying documents and retrieving required fields
- Operations teams responsible for converting files into usable business data
- Audit teams responsible for tracing extracted evidence to source material
- Knowledge teams responsible for organizing information from mixed file collections
- Business teams responsible for reviewing extracted tables, facts, and exceptions

The course assumes participants can inspect common business files and leaves out coding, API integration, workflow automation, model development, and general AI strategy.

### Departments and Industries

The course supports document-intensive work across operational, assurance, and information functions.

- Records management, administration, and shared services
- Operations, procurement, finance, and internal audit
- Legal operations, compliance, and knowledge management
- Healthcare, insurance, and financial services
- Manufacturing, logistics, retail, and professional services

### Learning Objectives

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



- Analyze document and image suitability for extraction
- Build field, table, and evidence extraction schemas
- Apply page, region, and layout instructions
- Use provenance and exception records for traceability
- Evaluate extracted information against source material
- Build a reviewed information extraction specification

## **Course Agenda**

### **Day 1: Extraction Scope and Evidence**

- Document-to-Decision Information Map
- Required Fact and Evidence Inventory
- Source Type Classification Matrix
- Extraction Boundary and Exclusion Statement
- Human Decision Ownership Record

### **Day 2: File and Image Quality Triage**

- Digital Text and Scanned Page Triage
- Image Resolution and Legibility Checklist
- Rotation, Cropping, and Obstruction Review
- Handwriting and Mixed-Language Risk Flag
- File Suitability Decision Gate

### **Day 3: Extraction Schema Design**

- Field Name and Data Type Dictionary
- Required and Optional Field Rules
- Allowed Value and Format Constraints
- Missing, Ambiguous, and Not-Applicable Codes
- Extraction Schema Version Record

### **Day 4: Page and Region Instructions**

- Page Range Selection Method
- Header, Footer, and Margin Exclusion Map
- Region-of-Interest Instruction Template
- Reading Order and Multi-Column Guide
- Repeated Section Identification Rule



## **Day 5: Fields and Key-Value Pairs**

- Label-to-Value Association Pattern
- Dates, Identifiers, and Amount Format Check
- Multi-Line Value Capture Rule
- Conflicting Field Resolution Log
- Field-Level Source Reference

## **Day 6: Tables and Repeating Records**

- Table Header and Column Mapping Grid
- Merged Cell and Split Row Handling Rule
- Repeating Line Item Extraction Template
- Subtotal and Total Reconciliation Check
- Table Cell Exception Register

## **Day 7: Images and Visual Evidence**

- Screenshot and Photograph Context Brief
- Chart Label and Value Capture Grid
- Diagram Entity and Relationship Record
- Visual Claim and Observation Separation
- Embedded Visual Availability Check

## **Day 8: Provenance and Confidence**

- Page, Region, and Source Provenance Pointer
- Field Confidence Rating Scale
- Low-Confidence Review Threshold
- Unsupported Inference Warning Flag
- Extraction Exception Queue

## **Day 9: Validation, Privacy, and Control**

- Source-to-Output Sampling Plan
- Accuracy and Completeness Validation Checklist
- Sensitive Information Handling Boundary
- Reviewer Correction and Approval Record
- Batch Consistency Comparison Sheet



## Day 10: Reviewed Extraction Practice

- Suggested Exercise: Triage a Mixed File Set
- Suggested Exercise: Build a Field and Table Schema
- Suggested Exercise: Extract with Provenance Pointers
- Suggested Exercise: Validate Exceptions and Corrections
- Capstone Exercise: Reviewed Information Extraction Specification

## Practical Exercises

The course uses suggested activities to convert mixed files into traceable and reviewed information.

- Suggested activity: classify digital documents, scans, screenshots, charts, and photographs by extraction suitability
- Suggested activity: design field and table schemas with formats, missing-value codes, and source pointers
- Suggested activity: extract records while logging confidence, ambiguity, unsupported inference, and exceptions
- Suggested activity: compare outputs with sources and document corrections, privacy decisions, and reviewer approval

## FAQs

### Who suits ChatGPT document and image information extraction training?

ChatGPT document and image information extraction training suits records, operations, audit, knowledge, and business teams that inspect common files and need structured, reviewed outputs without building software.

### How does document and image information extraction differ from general ChatGPT prompting?

Document and image information extraction uses defined schemas, source locations, exception codes, and validation rules. General prompting covers broader tasks and may not require field-level traceability or systematic comparison with source material.

### Can ChatGPT extract tables and fields from every document image?

Extraction depends on file support, legibility, layout, and available visual content. Participants use suitability checks, confidence flags, and human review instead of assuming every field or table is captured correctly.

### How should extracted information be checked against documents and images?

Extracted information should be sampled or fully compared with page and region references, required fields, format rules, totals, and visible evidence, with corrections and unresolved exceptions recorded.



## **How does source provenance improve document extraction review?**

Source provenance links each extracted fact to its file, page, region, or visual element, helping reviewers locate evidence, resolve ambiguity, and explain corrections.

## **Conclusion**

Participants take back a Reviewed Information Extraction Specification connecting source types, schemas, instructions, provenance pointers, confidence flags, validation checks, privacy boundaries, and review ownership. It makes document extraction decisions visible and repeatable. It supports more consistent handling of fields, tables, visual evidence, exceptions, and corrections across document-intensive work.