JUPITER PRIME
Consulting • Training • Advisory
جوبيتر برايم للاستشارات والتدريب

COURSE BROCHURE

ChatGPT Vision and Document Data Extraction Training Course

6 – 10 December 2026
Amman

ChatGPT Vision and Document Data Extraction Training Course

Ref: 1265_47558 **Date:** 6 – 10 December 2026 **Location:** Amman **Fees:** 4200 GBP

ChatGPT Vision and Document Data Extraction is a 1-week training course for operations, finance and records specialists who need reliable fields from invoices, forms, scans and photographs, and it produces an extraction specification and accuracy log. Multimodal models can now read layouts, tables and handwriting that defeated older template tools, yet unchecked output can quietly corrupt ledgers and registers. Jupiter Prime Training runs ChatGPT Vision and Document Data Extraction as a hands-on course.

Why ChatGPT Vision and Document Data Extraction Matters Now

Three pressures meet in back offices. Volumes of supplier invoices, delivery notes, inspection photos and application forms keep rising while headcount stays flat. Finance controllers and auditors expect every captured figure to be traceable to its source page. Records and privacy officers want personal and commercial details handled under clear rules. Teams that pair multimodal extraction with schemas and sampling cut keying effort and errors; teams that skip the checks import silent mistakes into systems of record. The course draws on ISO/IEC 42001, the AI management system standard, and the NIST AI Risk Management Framework AI RMF for oversight of extraction tasks.

Who Gains Most from ChatGPT Vision and Document Data Extraction

- Accounts payable specialists who capture invoice headers and line items
- Operations coordinators who process delivery notes, permits and site photos
- Records officers who index scanned archives and correspondence
- Procurement analysts who pull terms and quantities from supplier documents
- Process improvement staff who redesign manual keying steps

A shared-services unit sending finance and records specialists together gains most, as both sides then agree field definitions and acceptance rules.

ChatGPT Vision and Document Data Extraction Capabilities

Participants practise turning messy pages and pictures into validated fields that downstream systems can trust.

- Specify target fields and a JSON output schema for each document type — system owners
- Design extraction prompts that handle tables, stamps and handwriting — processing teams
- Test accuracy through field-level sampling and confusion logs — finance controllers
- Apply masking and retention rules to personal and commercial details — privacy officers
- Build exception routes for low-confidence or missing fields — operations supervisors
- Calculate keying time saved against error and rework rates — shared-services leads

Key ChatGPT Vision and Document Data Extraction Terms

- Multimodal model — A model that accepts images and text together, so it can read pages, photos and diagrams.
- Output schema — A fixed list of field names, types and formats the extraction must return.
- JSON — A plain text format of keys and values that spreadsheets and business systems can import directly.
- Field-level accuracy — The share of individual extracted values that match the source document exactly.
- Exception queue — A holding list where uncertain or failed extractions wait for human review.
- Source anchoring — Recording the page and position each value came from, so reviewers can check it quickly.

ChatGPT Vision and Document Data Extraction Day by Day

Day 1: Document Types and Extraction Scope

- Document Inventory — Grouping Invoices, Forms, Scans and Photos by Layout
- Multimodal Model Capabilities and Limits — Matching Tasks to Model Strengths
- Field Definition Sheets — Agreeing Names, Types and Mandatory Values
- ISO/IEC 42001 and NIST AI RMF — Framing Oversight of Extraction
- Masking and Retention Rules — Protecting Personal and Commercial Details

Day 2: Extraction Prompts and Schemas

- JSON Output Schemas — Forcing Consistent Fields and Formats
- Table and Line-Item Prompts — Capturing Quantities, Prices and Codes
- Handwriting, Stamps and Signatures — Handling Hard-to-Read Regions
- Multi-Page and Mixed-Language Files — Keeping Context Across Pages
- Few-Shot Page Examples — Teaching Unusual Layouts

Day 3: Images, Charts and Field Photos

- Meter, Label and Nameplate Photos — Reading Values From Site Images
- Chart and Diagram Interpretation — Converting Visuals to Figures
- Image Quality Checks — Rejecting Blurred or Cropped Inputs
- Source Anchoring — Linking Each Value to Its Page Location
- Confidence Flags — Marking Uncertain Fields for Review

Day 4: Accuracy and Workflow Integration

- Field-Level Sampling Plans — Measuring Extraction Accuracy
- Confusion and Error Logs — Finding Repeat Failure Patterns
- Exception Queues — Routing Low-Confidence Records to Reviewers
- Spreadsheet and System Import Templates — Feeding Ledgers and Registers
- Throughput and Rework Measures — Judging Benefit Against Manual Keying

Day 5: Practising the Extraction Specification and Accuracy Log

- Case Field Definitions — Applying Field Sheets to a Document Set
- Case Prompt and Schema Build — Applying JSON and Table Techniques
- Case Accuracy Trial — Applying Sampling and Error Logs
- Reviewer Walkthrough — Presenting Exceptions and Masking Steps
- Extraction Specification and Accuracy Log — Completing the Handover Pack

Applying ChatGPT Vision and Document Data Extraction After the Course

- Write a field definition sheet and JSON schema for one high-volume document type and trial it on a sample batch.
- Run a field-level accuracy sample on the trial output and share the error log with the finance controller.

ChatGPT Vision and Document Data Extraction Questions FAQ

What is ChatGPT Vision and Document Data Extraction and why does it matter now?

It is the use of multimodal models to pull structured fields from documents, scans and images. It matters now because document volumes keep rising, and unchecked extraction can push silent errors into ledgers and registers.

Which roles gain most from ChatGPT Vision and Document Data Extraction?

Operations, finance and records specialists gain most, especially accounts payable, procurement and archive staff who key data from documents every day.

How does ChatGPT Vision and Document Data Extraction differ from an OCR and capture platform course?

It works with general multimodal models rather than a dedicated capture platform. You design prompts, schemas and sampling checks yourself, which suits varied layouts and photos that template tools handle poorly.

What do you bring back from ChatGPT Vision and Document Data Extraction?

You bring back a specification for one document type, with field definitions, prompts, a JSON schema, masking rules and an accuracy log.

ChatGPT Vision and Document Data Extraction: The Way Forward

The extraction specification and accuracy log first inform the controller's decision on whether a document type can move from manual keying to assisted capture. Processing staff and reviewers then work from it. After the first month, revisit confidence thresholds and sampling size, since volumes in sectors such as construction or healthcare often expose new layouts.