

# RAG · AP-Automation · Fraud detection

## MODULE 4 — ACCOUNTING & TREASURY OPERATIONS

RAG on your documents. AP-Pilot. GoBD-compliant.

3-5 hours

Individual Coaching

Live-Demo RAG

**note:**



This training is conducted by an AI system as trainer (e.g. ChatGPT, Gemini, Claude) – upload both Markdown files (Systemprompt.md and content.md) and type “Start”. No human trainer required.



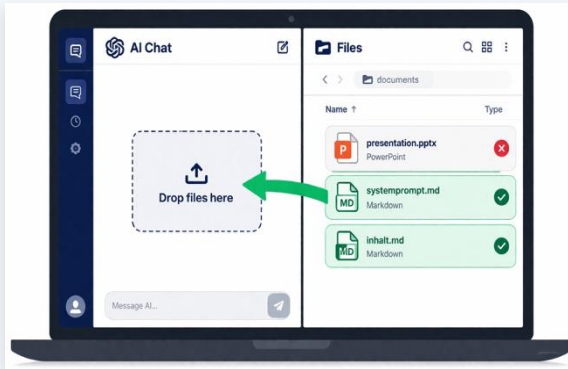
Download both files: <https://www.foundic.org//trainings/>



FOUNDIC.org



## 1 Upload Files to Chat

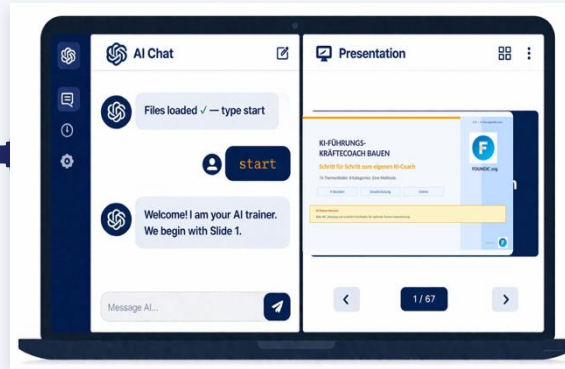


- ① **Open your LLM in the browser**  
ChatGPT, Claude or Gemini — Text mode (no audio!)
- ② **Upload both Markdown files**  
systemprompt.md + content.md via Drag & Drop
- ❌ **Do NOT upload the PowerPoint**  
Only the two .md files belong in the chat

⚠️ No audio yet — files can only be uploaded in text mode.

📁 Files missing? Download: [foundic.org/trainings](https://foundic.org/trainings)

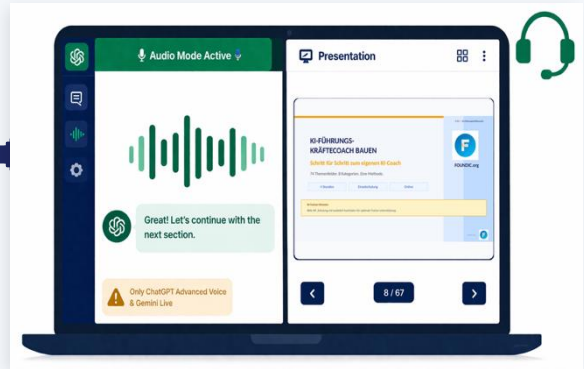
## 2 Begin training with "start"



- ① **Type the word: start**  
A single word is enough — the trainer starts automatically
- ② **Split Your Screen**  
Left: Chat window · Right: Open PowerPoint slides

✓ The slides are your visual anchor — the trainer tells you when to click forward.

## 3 Optional: Enable Audio Mode



- ① **Complete Steps 1 & 2 first**  
Upload files + type "start" — only then enable Audio
- ② **Activate audio mode**  
ChatGPT: headphone icon · Gemini: microphone icon
- ③ **Put on headset & speak**  
View slides, speak freely with the trainer

This feature is currently only available in ChatGPT (Advanced Voice) and Gemini Live — not in Claude.

# Transfer Check Module 3 — What Have You Implemented?

## Transfer Check Module 3

Which agent workflow have you built?  
 How much time does your workflow save per week?  
 → Brief Report: Make.com or Other Tools?  
 → HITL Gate Implemented? Who Approves?

## What went well?


Which Step in the Workflow Was Difficult?  
 → Guardrails: Does the HITL Gate Work?  
 → Time savings: note the concrete number  
 → What Would You Do Differently Next Time?

## Entry Check Module 4

Which Financial Documents Do You Search Most Frequently?  
 → loan agreements, guidelines, annual financial statements?  
 → I Note This for the Building Session

## Your Accounting Problem Today

Which Manual Accounting Process Costs You Most Time?  
 → Will Be Addressed in the Building Session  
 → No process too complex or too simple.

 files ready? northern\_light\_loan\_agreement.pdf + automation\_readiness\_scoring.xlsx

# Agenda — What to Expect Today

15 min	Transfer Check M3 + Entry Check	Introduction
10 min	Input: Accounting Ops — the hidden time drains	Input
15 min	RAG — Your Documents as Knowledge Base	Input
15 min	LIVE-DEMO: Q&A on Loan agreement (Northern Light Corporation)	Demo
15 min	AP-Automation: OCR → LLM → HITL	Input
10 min	Document verification prompt + duplicate check	Input
15 min	EXERCISE 1: Readiness-Scoring for your AP-process	exercise
10 min	Discussion: Should the Agent Post the Invoice?	analysis
10 min	LEARNING STATUS CHECK 1 (6 Multiple Choice)	Check
10 min	<i>BREAK</i>	break
15 min	Fraud- & Anomaly detection: Benford, IQR, Timing	Input
10 min	Account Reconciliation: Bank, Intercompany, Credit Card	Input
15 min	EXERCISE 2: Anomaly pattern detection (Excel)	exercise
10 min	LEARNING STATUS CHECK 2 (6 Multiple Choice)	Check
10 min	<i>BREAK</i>	break
35 min	BUILDING SESSION: RAG Q&A on Financial Documents	Hands-on
10 min	Mini-Case + Use Cases	analysis

2–4 hours. 2 learning checks. 1 building session. Result: your RAG Q&A on your own financial documents.



# Where Does Your Time Disappear in Accounting?

## Incoming Invoices

7 Manual Steps: Open → Check → Code → Approve → Post → File → Archive. Average 8 min./Invoice.

## Document Search


"Which Covenant Addresses a Leverage Ratio of 3.5x?" — 2 Hours Contract Reading vs. 30-Second RAG.

## Account Reconciliation

Intercompany Reconciliation: 3 Teams, 2 Systems, Endless Email Chains. Differences Remain Open Until Month-End Close.

## Fraud Suspicion

Anomalies in the Posting Journal: Who Looks at 15,000 Rows Manually? No One — Until the Auditor Asks.

 *Tanja, CFO Northern Light Corporation: "AI Doesn't Save Me 30 Seconds. It Gives Me 2 Hours Per Week Back — for Analysis Instead of Searching."*

# RAG — What Is It and Why Do You Need It?

## RAG — The Librarian Analogy

A Classical LLM Knows Much — but Not What's in Your Loan Agreement.

RAG = Retrieval-Augmented Generation: AI Answers Based on Your Documents.

Analogy: RAG Is Like a Very Fast Librarian:

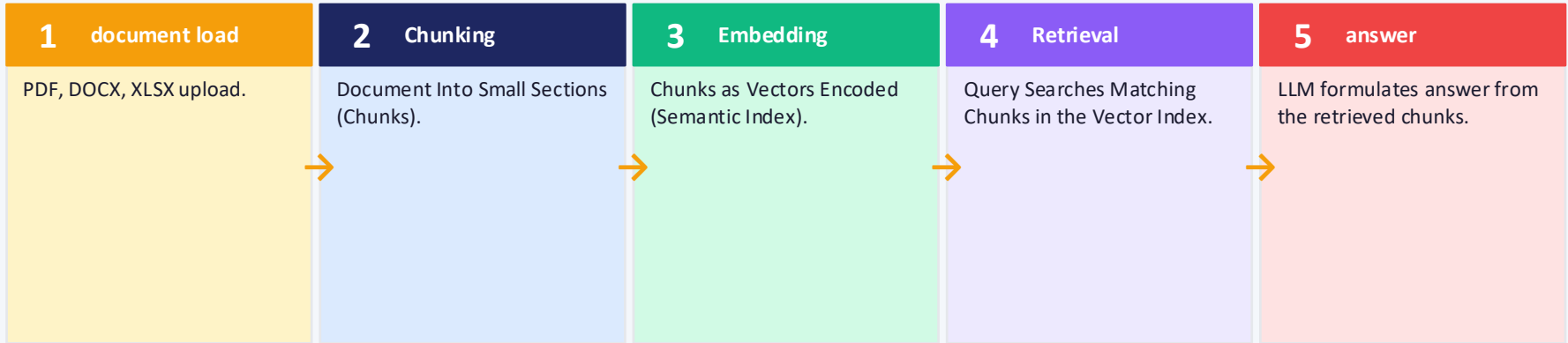
He Knows All Your Books (Documents), Finds in Seconds

The Relevant Page and Formulates a Precise Answer with Source Reference.

## RAG vs. Standard-LLM

Standard-LLM	RAG-system
training data (2024)	<b>your current documents</b>
No source reference	Always: document + page
Can hallucinate	<b>Only from your docs</b>
No access to internal data	your contracts, guidelines, reports

# How RAG Works — The Technical Process



No-code options: ChatGPT File Upload · Claude Projects · Microsoft Copilot (SharePoint) · Google NotebookLM — no server, no code, immediately usable.



# Which documents are ideal for RAG?

## ✓ Ideal RAG Documents

Loan Agreements — Covenants, Interest Rate Tiers, Events of Default  
Internal Guidelines — GoBD, Travel Expense Policy, Approval Procedures  
Annual Financial Statements — Balance Sheet Q&A, KPI Search  
Audit-reports — audit findings, recommendations  
Supplier Contracts — Payment Terms, Discount Rules, SLAs

## ⚠ Limitations to Consider


Scanned PDFs Without OCR — Not Readable or Searchable  
Highly Tabular Documents — Chunking Can Lose Table Context  
Personal Data — GDPR Review Required Before Upload  
Confidential Negotiation Documents — Observe Company Data Protection Policy

🔒 **GDPR Rule: No Personal Data (Names, Salaries, IBAN) in External AI Tools. Anonymize or Use Enterprise Plan with Data Protection Guarantee.**

# LIVE-DEMO — Loan agreement-Q&A in 10 minutes

 LIVE-DEMO — Trainer Leads Through. Participants Observe and Pose Questions.

- |          |                                |  |
|----------|--------------------------------|--|
| <b>1</b> | <b>document upload</b>         | northern_light_loan_agreement.pdf in Claude or ChatGPT Upload (File-Upload). No Setup, No Account Switch Required.         |
| <b>2</b> | <b>Ask Your First Question</b> | <i>"What financial covenants are included in this agreement? List all with threshold and measurement date."</i>            |
| <b>3</b> | <b>Second question</b>         | <i>"What does the contract stipulate for a covenant breach? What rights does the lender have in the event of default?"</i> |
| <b>4</b> | <b>Generate Table Output</b>   | <i>"Create a Table: Covenant   Definition   Threshold   Measurement Period   Reporting Obligation"</i>                     |

 Goal: Tasks requiring 2 hours of contract review, RAG completes in under 30 seconds — with source reference.

# Which RAG Tool for Which Purpose?

Tool	Cost	Document Scope	Strength	Limitation
<b>ChatGPT File-Upload</b>	Free (Plus: €20/mo.)	Single Document, Temporary	Quickest option, no set up	No persistent project
<b>Claude Projects</b>	Free / Pro €20/mo.	Up to 5 docs, persistent	Best Answer Quality	200k token context window
<b>Copilot (SharePoint)</b>	M365 license	SharePoint-documents	IT integration, GDPR-compliant	IT Coordination Required
<b>Google NotebookLM</b>	Free	Multiple Docs, Audio Summary	Good Visualization	Google account required

💡 Recommendation: Claude Projects or ChatGPT Plus — immediately usable, no IT required, best answer quality. Copilot for company SharePoint when M365 is available.

# How good is your RAG-answer? — 4 quality checks



## Check 1: source reference available?

Every RAG Answer Must Include Document Name + Section/Page Name.  
Without Source: Sharpen Prompt: "Always Provide Document Name and Page Number."  
Cross-check source against original PDF — takes 30 seconds.



## Check 2: Hallucinations-Test

Ask a Question That Is NOT Answered in the Document.  
Correct RAG Answer: "There is No Information About This in the Provided Documents."  
Incorrect Answer: RAG Invents Content — Adjust System Prompt.



## Check 3: precision test

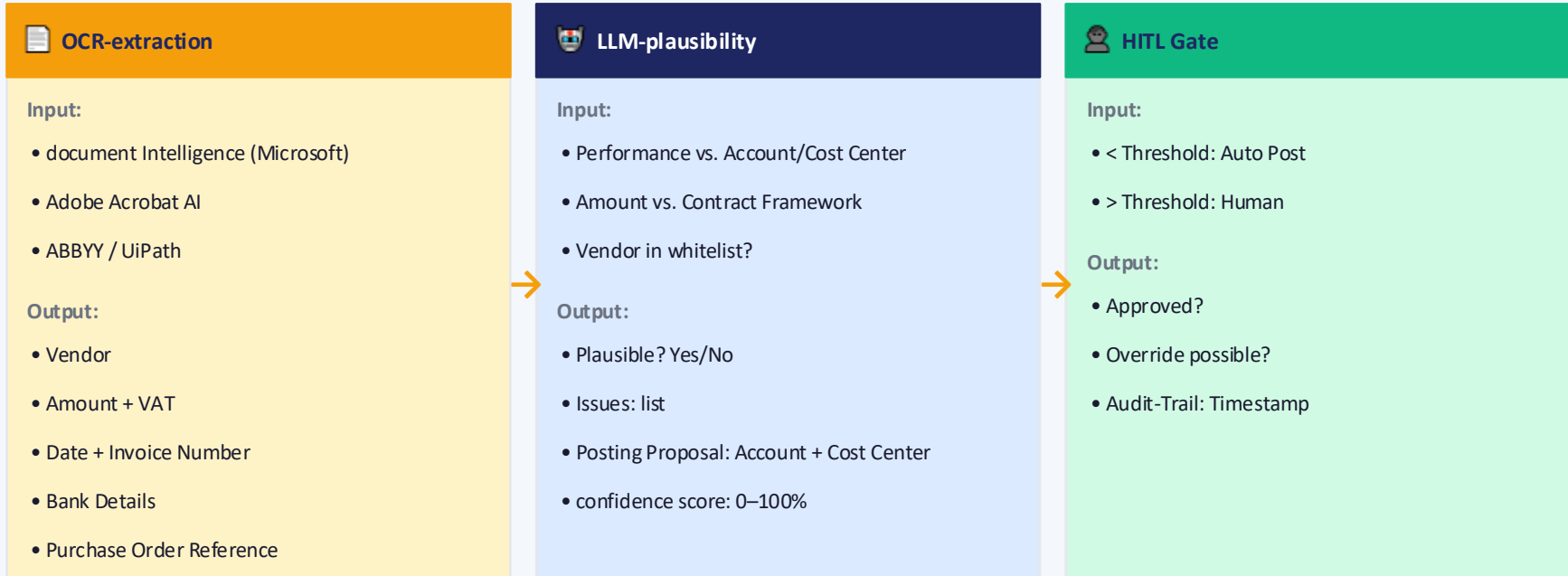
Ask a Question with a Known Answer (e.g., Interest Rate From Contract).  
Does the Number Match? If Not: Check Chunking, Shorten Document.  
Goal: <2% Error Rate for Structured Documents.



## Check 4: GDPR-compliance

Was the Document Reviewed for Personal Data Before Upload?  
Protocol Lead: Which Documents Were Loaded Into Which System When?  
When in Doubt: Involve Legal or Data Protection Officer.

# AP Automation — From 7 Manual to 3 Automated Steps



# The LLM Plausibility Prompt — Structure and Result

## Prompt Structure (Level 2)

You Are an Experienced Accountant.  
Check This Invoice Data for Plausibility:

### INVOICE DATA:

Vendor: [VENDOR]  
Amount: [AMOUNT] EUR  
Performance: [SERVICE]

### CHECK CRITERIA:

1. Vendor in Whitelist?
2. Performance fits account?
3. Amount in the Contract Framework?

### OUTPUT (JSON):

```
{ plausible: true/false,
  issues: [],
  account: "...",
  confidence_score: 0-100 }
```

## duplicate-Check (before posting)

Exact Match: Invoice Number + Amount + Vendor + Date Identical.

Fuzzy Match: Amount  $\pm 5\%$ , Similar Data — Mark as Suspect.

Duplicate rate: avg. 0.3% of invoices — but every one is a loss.

## exception-Management

LLM confidence < 70%: automatically routed to exception queue.

Unknown Vendor: Immediate HITL — No Auto-Posting.

Large amounts (> €5,000): always HITL regardless of confidence.

# EXERCISE 1 — How ready is your AP-process for Automation? (15 min.)

 **Task: Assess Your Most Important AP Process Against 5 Criteria (Scale 1–5).**

<b>Structure</b>	Are Incoming Invoices Always in the Same Format?	<input type="button" value="1"/>	<input type="button" value="2"/>	<input type="button" value="3"/>	<input type="button" value="4"/>	<input type="button" value="5"/>
<b>Volume</b>	More Than 20 Invoices/Month of This Type?	<input type="button" value="1"/>	<input type="button" value="2"/>	<input type="button" value="3"/>	<input type="button" value="4"/>	<input type="button" value="5"/>
<b>Rule Clarity</b>	Are There Clear Validation Rules (Whitelist, Amount Limits)?	<input type="button" value="1"/>	<input type="button" value="2"/>	<input type="button" value="3"/>	<input type="button" value="4"/>	<input type="button" value="5"/>
<b>System Integration</b>	Direct connection to accounting system possible?	<input type="button" value="1"/>	<input type="button" value="2"/>	<input type="button" value="3"/>	<input type="button" value="4"/>	<input type="button" value="5"/>
<b>HITL Readiness</b>	Is an approver defined for exceptions?	<input type="button" value="1"/>	<input type="button" value="2"/>	<input type="button" value="3"/>	<input type="button" value="4"/>	<input type="button" value="5"/>

 Evaluation: Score  $\geq 4.0$  = Quick Win (automate now) | Score 2.5–3.9 = Pilot | Score  $< 2.5$  = stabilize process first

# Discussion — Should the Agent Post the Invoice?



## The System Says: Plausible (Confidence 94%)

Vendor: TechServ GmbH (Whitelist: Yes)

Amount: 2,340 EUR (Framework: Up to 5,000 EUR: Yes)

Performance: IT Support February (Account 6800: Fits)

Duplicate Check: No Duplicate Found



## The Accountant Says: I Confirm

He Has 200 Suggestions This Week.

He Clicks Through Without Checking Each One Individually.

He Trusts the System — "It Was Always Correct Before."

He Has No Time for True HITL.



## Discussion Question: Who Bears Responsibility When a Posting Error Slips Through? The System? The Accountant? The CFO?

Which controls would have prevented these errors? · Sampling review · Reversal-rate KPI · HITL quality audit · Volume cap per day





Learning Status Check 1 — 6 Multiple Choice-questions | per 1 correct answer | trainer discusses every answer directly.

### 1. What Is the Main Advantage of RAG Compared to Standard LLM?

- A) RAG Is Faster
- B) RAG Answers Based on Your Documents with Source Reference
- C) RAG Is Cheaper
- D) RAG Requires No Internet Connection

### 2. What Should Always Be Reviewed Before Uploading a Document in an AI Tool?

- A) file size
- B) GDPR compliance (no personal data-related. data)
- C) document length
- D) Number of pages

### 3. Which level in AP automation decides whether a human is involved?

- A) level 1: OCR
- B) level 2: LLM
- C) level 3: HITL Gate
- D) Level 0: Preparation

### 4. What does a confidence score of 45% mean in the LLM plausibility check?

- A) Post automatically
- B) Reject
- C) In exception-Queue — human checks
- D) Recheck with a new prompt

### 5. Which check is mandatory in a RAG hallucination test?

- A) Ask a question that IS answered in the document
- B) Ask a question that is NOT in the document — RAG should state this clearly
- C) document new upload
- D) Clear the system prompt

### 6. What is "HITL-Theater"?

- A) A training program for accountants
- B) Approval without genuine review due to excessive volume
- C) An audit procedure
- D) A RAG quality method

# PAUSE

Screen Off — Brief Stand-Up — Back in 10 Minutes

After the Break: Fraud Detection — Benford, IQR, Timing Analysis

# AI-Powered Fraud Detection — What It Can Do and What It Cannot



## What AI Detection Achieves

Benford's Law: First Digit of Amounts Follows a Statistical Pattern. Variance = Suspicion.

IQR Method: Outlier Identification on Posting Amounts. Limitations:  $Q1 - 1.5 \times IQR$  to  $Q3 + 1.5 \times IQR$ .

Timing Analysis: Postings Outside Normal Hours (Night, Weekend, Holiday).

User Activity Analysis: Volume per User ID, Outliers Over Time.



## What AI Does NOT Achieve

AI Identifies Patterns — Does Not Interpret Intent.

Many false positives: anomaly  $\neq$  fraud. Frequent normal exceptions.

No final judgment: the human renders the final judgment — always.

Context knowledge missing: Bernd tinkers on Saturdays = normal. AI doesn't know that.



Correct use: AI provides a prioritized list of suspects. The reviewer decides. This keeps the human in the loop without reading 15,000 rows manually.



# Fraud patterns in Excel — 3 methods

## Method 1: Benford's Law

Extract First Digit: `=LEFT(TEXT(ABS([amount]),"0"),1)`

Target Distribution: Digit 1 = 30.1%, Digit 2 = 17.6%, Digit 3 = 12.5% ...

Actual Distribution: COUNTIF / Number of Postings, Then Calculate %

Variance > 5%: Suspicion Increased — Examine Closer

## Method 2: IQR Outlier

$Q1 = \text{QUARTILE}(\text{amounts}, 1)$ ,  $Q3 = \text{QUARTILE}(\text{amounts}, 3)$

$IQR = Q3 - Q1$

Lower bound =  $Q1 - 1.5 \times IQR$ , Upper bound =  $Q3 + 1.5 \times IQR$

Outlier:  $\text{amount} < \text{lower bound}$  OR  $> \text{upper bound}$

## Method 3: Timing Analysis

Posting hour: `=HOUR(Timestamp-column)`

Day of week: `=WEEKDAY(Date,2)` — 7 = Sunday

Pivot on hour/day of week — pattern immediately visible

AI prompt on flag list: "Prioritize anomalies and explain possible causes"

# Automate Account Reconciliation — 3 Types

## Bank Reconciliation

*Manual: 4 hrs./Month | Automated: 30 Min. Review*

XLOOKUP Match: Bank Entry vs. ERP Posting by Amount + Date  
(±1 day). Unmatched items → AI analysis.

## Intercompany Reconciliation

*1E2761*

Cross-Matrix Comparison: Receivable A for Company X = Liability X for Company A?  
Discrepancy email sent automatically via Make.com.

## Credit Card Reconciliation

*10B981*

Statement rows vs. submitted receipts. LLM match on amount + description.  
Missing receipts: automatic reminder.

# Automating Treasury Ops — DSO, DPO, Collections

## DSO & DPO-Tracking

$DSO = \text{SUM}(\text{Open Receivables}) / \text{Daily Revenue}$

$DPO = \text{SUM}(\text{Open Payables}) / \text{Daily Purchases}$

In Excel: TODAY() – Invoice date = aging. PIVOT on age groups.

AI trend analysis: "Has DSO changed over the last 3 months? What are the causes?"

## Collections Automation

Escalation logic: level 1 (14 days overdue) → level 2 (30 days) → level 3 (45 days).

LLM-generated dunning notices: customer name + amount + date + tone = personalized.

Make.com workflow: daily overdue check → level-dependent email sent automatically.

GDPR: archive dunning correspondence, document sending timestamps.

## Early Payment Discount Optimization

Algorithm: effective discount rate (discount% / payment terms in days × 365) vs. cost of capital. If effective rate > cost of capital: pay early and capture the discount.

Collections escalation logic in Excel + Make.com: 3 hours to set up, then runs automatically. First month: 15 minutes instead of 3 hours on collections.



# Exercise 2 — Find Anomaly Patterns in the Posting Journal (15 Min.)

 Task: Find 3 Anomaly Patterns in the `northern_light_corporation_dataset.xlsx` (Table "Posting Journal").

**1** Benford Check


Extract the First Digit of Every Posting Amount.  
Compare Your Distribution with the Target Distribution (Digit 1 = 30.1%).  
Which Digits Deviate Strongly?

**2** Timing-analysis


Create a pivot table: postings by day of week and hour.  
When do unusually many postings occur?

**3** IQR-outlier

Calculate Q1, Q3, IQR for the amount column.  
Flag all outliers. How many are there? Which line items stand out?

 File: `northern_light_corporation_dataset.xlsx` — open table "Posting Journal". Columns: Date, Timestamp, User ID, Amount, Posting Text, Account.

Note: The dataset contains 3 intentional anomalies. Whoever finds all 3 earns a star.

 Learning Status Check 2 — 6 Multiple Choice-questions | per 1 correct answer.

#### 1. Benford's Law: Which First Digit Appears Most Frequently in Natural Numbers?

- A) 5
- B) 1 (approx. 30%)
- C) 9
- D) All equally frequent

#### 2. What does the "below-threshold pattern" indicate in fraud detection?

- A) Postings just below the approval threshold
- B) Postings below the minimum amount
- C) Postings without a cost center
- D) Postings outside the budget

#### 3. What is the primary purpose of the IQR method in posting analysis?

- A) duplicate-detection
- B) Outlier identification on amounts
- C) Vendor validation
- D) currency conversion

#### 4. Which Excel formula returns the day of week of a posting?

- A) =HOUR()
- B) =MONTH()
- C) =WEEKDAY(Date,2)
- D) =DATE()

#### 5. What is the correct approach for an unmatched bank reconciliation item?

- A) Post manually immediately
- B) Ignore until month-end
- C) AI analysis of the unmatched item + human review
- D) Reverse automatically

#### 6. What is a realistic DSO improvement effect through automated collections?

- A) No effect
- B) 5–15 days reduction through earlier escalation
- C) DSO drops to zero
- D) DSO increases due to higher workload

# PAUSE

Screen off — briefly stand up — back in 10 minutes

After the break: SoD, Audit Trail, and Building Session — your RAG on your own documents

# SoD — Four-Eyes Principle Also for AI Workflows

## Core SoD Rule for AP Automation

Whoever creates the posting proposal (AI or clerk) may NOT also approve it.

Whoever configures the system (whitelist, thresholds) may NOT also post.

Whoever changes vendor master data may NOT also execute payments.


SoD applies to humans AND systems — enforce technically, do not merely document.

## Common SoD Violations When Introducing AI

"The accountant set up the system and now approves their own entries." → Violation.

Recommendation: AI deployment — configuration by IT/Finance management, NOT by the clerk.

Escalation approval: route automatically to a second person. Never self-approve.

 SoD-checklist for your AP-workflow:

1. Posting proposal creator ≠ approver defined? 2. Configuration rights documented? 3. Exception queue: escalation to a second person?

# Audit Trail — Every AI Posting Must Be Traceable

## GoBD Requirements for Automated Postings

Retention requirement: 10 years (AO §147). Applies equally to AI-generated posting documents.

Integrity: postings cannot be altered after entry without a protocol record.

Completeness: every posting must include timestamp, user ID (human or system), and original document.

Machine readability: audit trail must be importable into IDEA/ACL — not PDF only.

## Mandatory Audit Trail Fields per Posting

Timestamp (ISO 8601): When was the posting proposed? When was it approved?

Creator: user ID of the AI system or the clerk.

Approver: user ID + timestamp of approval.

LLM confidence score: value + model version.

Override: Was the suggestion overridden? By whom? Justification?

 KPI: Reversal rate of AI-generated postings (target: <2%). Measured weekly and reported to the CFO.

# Tool Stack Module 4 — What you can use today

Function	Tools	Cost	Onboarding
<b>RAG / Q&amp;A on documents</b>	Claude Projects ChatGPT File-Upload Microsoft Copilot (SharePoint) Google NotebookLM	Free / from €20/mo.	Immediately, no setup
<b>OCR &amp; Receipt Extraction</b>	Microsoft Document Intelligence Adobe Acrobat AI ABBYY FlexiCapture	from 0 (free tier) to enterprise	IT coordination recommended
<b>AP-Automation (workflow)</b>	Make.com n8n (Self-Host) Zapier	from 0 (free tier) approx. €9–20/mo.	No-code, immediate start
<b>Fraud detection</b>	Excel (Benford, IQR, HOUR()) Claude / ChatGPT on flag list	Included in Excel	Immediately usable
<b>Account Reconciliation</b>	Excel (XLOOKUP) Make.com for IC-E-Mails	Included in Excel	15 min. setup

Recommended today: Claude Projects or ChatGPT Plus for RAG. Make.com for AP Automation pilot. Excel for fraud detection.

# BUILDING SESSION — your RAG-Q&A in 35 minutes

 **Goal: A live RAG Q&A instance on your own financial document — ready to take with you.**

## your document

Use your own financial document (loan agreement, policy, annual financial statements).  
No own document: use northern\_light\_loan\_agreement.pdf (provided).

## Tool

Claude.ai Projects or ChatGPT (File-Upload). account already available?

## time

35 minutes. Steps on slides 28–31. Trainer supports.

## result

3 tested prompts + hallucination test + QC checklist completed.

Now prepare your document and open Claude.ai or ChatGPT in the browser. Let's go.

# 1 Prepare Document — GDPR Check + Format

## ✓ GDPR-Check (obligation before Upload)

No names of natural persons (CEO, contact persons).

No IBANs, account numbers, or social security numbers.

No salary or personal compensation data.

Trade secrets: thresholds, margins, and pricing — assess carefully before upload.

## 📄 Format Optimization

Prefer text-based PDFs (not scanned image PDFs).

Scanned: run OCR first (Adobe Acrobat: File → Save as Text).

Very long documents (>50 pages): extract the relevant chapters.

Filename without special characters — e.g.  
northern\_light\_loan\_agreement\_anon.pdf.

✓ Document reviewed and ready? → Proceed to Step 2.

## 2 Build RAG Instance — System Prompt + Upload

### 📍 Claude Projects

Create new project → Name: [document type], e.g. "Loan Agreement Q&A"  
Document upload: Project → Documents → select file.  
Set system prompt (next box).

### 🟢 ChatGPT

ChatGPT open → Paperclip-Icon → file upload.  
Or: GPT-4 Plus → My GPTs → new GPT with document.

### 📄 System Prompt (copy and use)

You are a precise document analyst.

Answer questions **EXCLUSIVELY**

based on the uploaded documents.

rules:

- Always cite document name + section name
- No external sources or assumptions
- If not in the document: state this clearly
- Answers in English, precise

## 3 Test Prompts — 3 Mandatory Questions + Hallucination Test

- 1 Prompt 1: content overview** *"Create a structured overview of the most important content in this document. Organize by sections with a brief summary."*
- 2 Prompt 2: Specific question** *"Which [covenant / rule / requirement] applies to [your topic]? State the exact wording and the section."*
- 3 Prompt 3: Table output** *"Create a table: [column 1] | [column 2] | [column 3] — populate from the document."*
- 4 Hallucinations-Test** *"What are the [FICTITIOUS term not found in the document]?" — Correct answer: "There is no information about this in the document."*

The hallucination test is mandatory: if RAG answers a fictitious question as if it were true, the system prompt is too weak.

## 4 Quality Control — Complete the QC Checklist

✓ **source reference** Does every answer cite document name + section / page?

✓ **Hallucinations-Test** Did RAG correctly respond to the fictitious question with "not in the document"?

✓ **precision test** Does an answer for a known fact match the original document?

✓ **GDPR Documentation** Documented: which document, when, in which system?

✓ **transfer-question** Which 3 questions will you actually ask your RAG next week?

All 5 checks passed? Congratulations — your RAG instance is production-ready for finance tasks.

# Use Cases — What works today, immediately

## UC 4.1 — controlling / FP&A

Automated Variance Analysis with RAG on Budget Guidelines.

Question: "Which Cost Centers Lie According to Guideline Outside the Plan Framework?"

**WOW: Answer in 30 Sec. Instead of 45 Min. Manual Guideline Search.**

## UC 4.3 — treasury

Automated Bank Reconciliation with Difference Escalation.

XLOOKUP Match + Make.com Difference Email.

**WOW: 4 hrs. Bank Recon → 30 Min. Review Queue.**

## UC 4.2 — Accounting

AI-powered incoming invoice review End-to-End.

OCR → LLM-plausibility → HITL. 200 invoices → 5 hrs. HITL.

**WOW: Error Rate <2%, Posting Proposal Already Formatted.**

## UC 4.4 — Strategy / IR

RAG-Q&A on own Annual financial statements & reports.

Answer Analyst Questions in Real-Time.

**WOW: IR Preparation: 2 hrs. → 30 Min. with Pre-Prepared RAG Answers.**

# Mini-Case — "AI Proposes, Accountant Approves Blindly"

## Max, AP Accountant

Approved blindly — under time pressure, without malicious intent.

200 suggestions/week. 90% had been correct so far. He trusts the system.

## CFO / Finance-manager

Must clarify responsibility and improve controls.

Deployed the system without defining SoD and a reversal rate KPI.

## Internal Auditor

Finds errors in the annual financial statements. Demands evidence and process correction.

"No supporting document, no posting audit trail — this is an audit finding."

## AI system (Moderator)

"The system proposed correctly. The approval rested with the humans."

The system bears no liability — the question is: who defined the governance?



Closing Check: Have you internalized the core concepts from Module 4? 6 questions | 1 point per correct answer.

### 1. What Distinguishes RAG from a Normal LLM Query?

- A) RAG is faster
- B) RAG is based exclusively on your uploaded documents
- C) RAG requires no internet connection
- D) RAG cannot output tables

### 2. Which step in AP automation decides whether a human is involved?

- A) OCR-extraction
- B) duplicate-Check
- C) HITL Gate with threshold
- D) Posting proposal generation

### 3. What does SoD mean in an automated AP process?

- A) system-on-Demand
- B) Posting proposal creator ≠ Approver
- C) Automated posting without human
- D) Software-on-Deployment

### 4. Which fraud detection pattern uses statistical digit distribution?

- A) IQR-method
- B) Timing-analysis
- C) Benford's Law
- D) duplicate-Check

### 5. What is the GoBD retention requirement for AI-generated posting documents?

- A) 3 years
- B) 6 years
- C) 10 years (AO §147)
- D) No legal obligation

### 6. What is a Hallucinations-Test for RAG?

- A) System speed test
- B) Ask a question that IS answered in the document
- C) Ask a question NOT in the document — RAG should respond "not in the document"
- D) Test the upload duration

All 6 correct? Then you are ready for Module 5: Strategy & Governance.

# Key Takeaways — What You Take With You Today

**1****RAG = Your Knowledge Base**

No hallucination when correctly configured. Always source reference. System prompt is the foundation.

**2****Document verification in 3 levels**

OCR → LLM plausibility → HITL. Thresholds are governance decisions — not technical ones.

**3****Fraud detection: AI + human**

Benford, IQR, timing provide a prioritized suspect list. The human renders judgment — always.

**4****SoD applies also for AI**

Posting proposal creator ≠ approver. Enforce technically, do not merely document.

**5****Readiness before Automation**

Poor data × AI = scaled errors. Stabilize the process first, then automate.



# Transfer Task — Your Assignment Before Module 5

 **Assignment (20–30 min., to be completed before Module 5): 3 mandatory deliverables.**

- 1 RAG-Q&A in practice use**  
Ask your RAG instance at least 5 questions from your real work.  
Document: answer quality (1–5), source reference present, hallucination observed?
- 2 Readiness-Scoring conduct**  
Assess a concrete AP or treasury process using the scoring model from slide 13.  
Result: score + class (Quick Win / Pilot / stabilize process first).
- 3 Sketch SoD Concept**  
For your planned automation pilot: who creates suggestions? Who approves?  
A simple flowchart (hand-drawn is fine) is sufficient.

 Acceptance criteria: RAG source references reviewed · Readiness score documented · SoD clearly defined · No real personal data used.

Transfer task = your entry ticket. Those who bring all 3 deliverables will start Module 5 with a running pilot.

# Digital Finance Masterclass — Your Progress

## The Masterclass-roadmap

1

### Introduction & Quick Wins

*AI-fundamentals, CRAFT, data protection, Custom GPT*

 Completed

2

### Data Analytics

*PowerQuery, KPI-Dictionary, Forecasting, IBCS*

 Completed

3

### Virtual Finance Team

*agents, Make.com, guardrails, HITL*

 Completed

4

### Accounting & Operations

*RAG, AP-Automation, Fraud detection, SoD*

 Completed

5

### Strategy & Governance

*EU AI Act, ROI calculator, roadmap, change management*

→ Next session

## My Quick Win Plan — Starting Tomorrow

### 1. RAG-protocol:

*Load your own financial document and ask 5 test questions*

### 2. AP-Readiness:

*Complete the 5-criteria score for your AP process*

### 3. Fraud-Check:

*Apply Benford analysis to your own posting journal*

### 4. SoD sketch:

*Creator / reviewer / approver for AI workflow*



# Module 4 Complete!

## ACCOUNTING & TREASURY OPS

### Module 5: Prepare your transfer task

*Apply RAG, assess readiness, sketch your SoD.*

# Glossary — Key Terms Module 4

**RAG**

Retrieval-Augmented Generation: AI Answers Questions Based on Your Documents.

**Chunking**

Dividing a Document Into Small Sections for Better Retrieval.

**Embedding**

Vector Encoding of Text Meaning for Semantic Search.

**OCR**

Optical Character Recognition: Text extraction from image-based PDFs.

**HITL Gate**

Human-in-the-Loop: Human Checkpoint in Automated Workflows.

**SoD**

Segregation of Duties: Four-Eyes Principle — Creator ≠ Approver.

**GoBD**

Principles of Proper Bookkeeping and Data Processing — 10-Year Retention.

**AP-Automation**

Accounts Payable Automation: Automated Incoming Invoice Processing.

**Benford's Law**

Statistical Law: First Digit of Natural Numbers Follows Logarithmic Distribution.

**IQR-method**

Interquartile Range: robust outlier detector for posting amounts.





**DSO**

Days Sales Outstanding: receivables aging in days.

**Audit-Trail**

Audit-proof record of every posting change.




# AI in Posting-Adjacent Processes — Decision Matrix

 <b>Permissible</b>	 <b>Only with HITL</b>	 <b>Prohibited</b>	 <b>Always Human</b>
Variance comment draft	Posting proposal < 500 EUR + Review	Auto-posting without human-Review	Period accrual (Accrual-decision)
Loan agreement-Covenants read via RAG	IC-Posting proposal with cross-check	Salary / compensation posting automatically	Journal Entry at the Month-end close
Anomaly detection on posting journal	Invoice approval < approval threshold	M&A / Insider-data in cloud LLM	Payment approval > limit
Forecast narrative and commentary	AP Automation Pilot with SoD-control	Closing entry without Controller-Review	Reversal on published closing

# You've done it!

Thank you for your time and your trust.

---

-  I hope you took something away — and put it into practice tomorrow.
-  Suggestions for improvement? Leave a comment at [foundic.org](https://foundic.org) — we read every note and continuously improve our training.
-  Our goal: free training for everyone — because continuing education should never be a question of budget.

 If you enjoyed the training — buy us a coffee. It helps us develop more free training.

 [Treat us to a coffee](#)

→ Feedback & comments: [foundic.org/trainings](https://foundic.org/trainings)