

WHERE TO START WITH AI · NO. 1

Employee questions on HR policy

An internal assistant for HR policy, rolled out in two stages: first it helps the HR team answer, then it answers approved, standard questions on its own.

Stage 1 · HR answers, with AI support

Stage 2 · AI handles the routine, HR handles the exceptions



WHY START HERE

It meets all four tests of a good first workflow



High volume

Hundreds of questions a month across about twenty topics.



Clear owner

HR leadership, already accountable for these policies.



Clear output

A correct answer that cites its source.



Known pain point

For HR and employees alike.

A closed body of content. The assistant answers from documents the company already owns. The risk does not go away. It becomes more visible, measurable, and manageable once the sources, the scope, and the escalation rules are defined.



ON ONE CONDITION

Low risk, but only within a strict scope

- 1 General policies**
written, approved, and current.
- 2 No individual employee files**
used as a source.
- 3 No HR decisions**
or judgment calls handed to AI.
- 4 Individual, sensitive, or ambiguous situations**
routed to HR.

Outside this scope, it becomes a different use case with a different risk profile.



THE WORKFLOW AS IT RUNS TODAY

Six steps, all done by hand

Example Services Inc. · 450 employees · fictitious company

1	Question sent to HR by email, by phone, or in the hallway	
2	Triage and assignment in the team's shared inbox	● Overloaded shared inbox
3	Finding the policy handbook, policies, internal memos	● Multiple versions, scattered documents
4	Drafting the answer case by case	● The same answer rewritten
5	HR manager review when the case is out of the ordinary	● Waiting for sign-off
6	Reply and filing answer sent to the employee	● No usable record

≈ 200

questions a month

≈ 70%

concern a written policy

≈ 2 days

average response time

● friction point · illustrative figures



BEFORE YOU START

Get the policies in order

The work starts with the sources: identify the handbook and policies that apply, resolve contradictions, assign an owner to each document, and date each version. The tool comes after.

1**Identify the documents that apply****2****Resolve contradictions****3****One owner per document****4****Dated versions****The tool comes after.**

STAGE 1 · HUMAN IN THE LOOP

The assistant drafts, the HR team decides

The question comes in
to the same HR inbox as today.



The assistant drafts a reply
citing the section and its version.



An HR team member decides
reviews, edits if needed, then sends.



Every correction is logged
topic by topic.

What this stage produces: a topic-by-topic measure of how good the drafted answers are.

Estimated duration: 8 to 12 weeks.



STAGE 2 · THE ASSISTANT HANDLES THE STANDARD CASES

It answers on approved topics; HR handles the exceptions

THE ASSISTANT ANSWERS WHEN

- The question falls under an approved topic
- A general, written, current policy applies
- The answer can cite the section and its version

IT ROUTES TO HR WHEN

- The situation is individual, sensitive, or ambiguous
- The source is missing, contradictory, or outdated
- The question involves topics such as health, conflict, pay, extended leave, or termination

Every week, HR reviews a sample of answers and keeps the sources up to date.



UNDER THE HOOD · RAG AND LEARNING LOOP

Reference architecture: how the system stays under control



UNDER THE HOOD · CONTROLS AND SECURITY

The controls required before any autonomous answer

1 Evaluation

A reference set of about 100 real questions. Metrics: accuracy, faithfulness to sources, correct citation, and the rate at which exceptions are detected and correctly routed. Rerun whenever the model, prompt, or corpus changes.

2 Observability

Every request is logged: question, retrieved passages, answer, confidence signals, routing decision. Because questions may contain personal data, the log is minimized, accessible only to authorized staff, and kept for a defined period.

3 Identity and access

SSO (SAML or OIDC). Role-based access control enforced at retrieval, not after generation.

4 Data

Hosting and data transfers verified for the chosen solution; contractual terms barring use of the data for model training; encryption; log retention and access defined; Swiss FADP compliance verified.

5 Application security

Protection against prompt injection. The assistant can only answer or route to HR.

Three options: the assistant built into your existing office suite, a managed RAG platform, or a custom build (orchestration, vector database, model API). The choice comes after the source audit, not before.



MOVING FROM STAGE 1 TO STAGE 2

A decision based on evidence, topic by topic

The move to stage 2 should not be decided in advance. It depends on what the HR team observes during stage 1: the quality of the drafted answers, the corrections needed, and the system's ability to recognize situations that fall outside its scope.

WHAT THE HR TEAM OBSERVES

- Accuracy of the answer and the citation
- Corrections needed, by topic
- Exceptions correctly routed
- No material incident within the tested scope
- Current sources, each with an owner

Vacation and leave could move to stage 2 while expenses stay in stage 1.



THE GOAL, AND HOW TO MEASURE IT

The goal is not to replace an HR decision. It is to cut response times on general-policy questions, stop the same answers from being rewritten, and give HR more time for situations that call for context and judgment.

Response time on general-policy questions

≈ 2 days —→ **in seconds**
today (illustrative) stage 2 target

Also tracked: the share of questions handled without HR involvement,
50–60% (illustrative target)

Data protection

Swiss FADP: employees informed; data processed and hosting location clarified before launch.

Governance

HR leadership owns the outcome and reviews a sample of answers every week.



In this example, AI is not meant to replace an HR decision.

It steps in first where a written policy already provides the answer.

So the useful question is not "Which tool should we buy?"

It is "Which repetitive work can we have AI assist with, measure, and then delegate under clear controls?"

What repetitive work in your company already meets these conditions?



Sergio Castagna Consulting.

sergiocastagna.com · info@sergiocastagna.com · Geneva