

Enterprise RAG Knowledge Base MVP

Comparable to: Enterprise RAG

A web application for mid-market companies that creates a shared internal knowledge portal across disconnected business systems.

Estimated MVP cost

\$16,450

Estimated effort

290 hours

To get your own customised cost for an app like this, visit www.chatstack.app and start a ChatStack interview describing how you want it customised.

Key Features

- Conversational AI assistant with cited, source-grounded answers and a searchable results list.
- Five API integrations spanning HR, SharePoint-style knowledge bases, project management, and Slack.
- Ingestion for PDFs, spreadsheets, and wiki content.
- Hybrid vector and BM25 retrieval with re-ranking and GraphRAG-style entity linking.
- OpenAI-based RAG answer generation.
- Query-time enforcement of existing source-system user and group permissions.
- Self-serve connection setup, sync-status visibility, and ingestion troubleshooting.

Out of Scope for This MVP

- Target US mid-market customers and host customer data in the United States.
- Support a SOC 2 and HIPAA-ready security and compliance posture.
- Preserve source-system permissions so users only retrieve content they are authorized to access.
- The MVP prioritizes a practical shared knowledge portal rather than a broad enterprise governance suite.
- Detailed audit analytics and advanced content-control dashboards are deferred unless needed for the initial launch.

Full Cost Breakdown

Item	Hours	Cost
The portal shall allow employees to search the shared knowledge base using keywords and phrases.	7	\$390
The portal shall allow employees to ask questions in natural language and receive answers grounded in authorized internal content.	17	\$940
The portal shall display the internal sources supporting each answer.	5	\$280
The portal shall enforce source-system user and group permissions when processing searches and generating answers.	15	\$830
The portal shall ingest and index PDFs, spreadsheets, and wiki content for search and question answering.	19	\$1,050
The portal shall allow administrators to connect Google Drive as a knowledge source.	10	\$550
The portal shall allow administrators to configure, authenticate, manage, and disconnect the Google Drive connection.	12	\$660
The portal shall synchronize approved Google Drive content and make synchronization status visible to administrators.	14	\$770
The portal shall allow administrators to troubleshoot Google Drive ingestion failures by retrying, re-running, correcting configuration, and resuming synchronization.	15	\$830
The portal shall record administrative, integration, synchronization, access, and security-relevant events in an audit trail.	12	\$660
Planning	32	\$1,760
Project Management	25	\$1,380
Testing	15	\$830

DevOps	12	\$660
Design & Frontend Build	38	\$2,090
AI Tokens	0	\$430
Security	15	\$830
Risk (10% Contingency)	27	\$1,490
Total	290	\$16,450

Estimate Assumptions

- Estimates cover application development and integration work, excluding NFR overhead, in frastructure operations, compliance certification, and vendor fees.
- Supabase Auth, PostgreSQL, object storage, a US-hosted vector store, and OpenAI Responses API access are available and configured.
- Google Drive OAuth consent, API quotas, permission metadata, and customer a

administrator access are available.

- Google Drive remains authoritative for source permissions, with sufficient metadata to filter retrieval at query time.
- PDF, spreadsheet, and wiki parsing uses existing libraries or services; advanced OCR and unsupported formats are excluded.
- Background workers or scheduled jobs are available for asynchronous synchronization and document processing.
- The portal uses the specified Next.js, Supabase, REST API, and Vercel stack with shared tenant and role

primitives.

- Audit retention, export, deletion, and detailed compliance evidence requirements are not yet specified.

Risks That Could Change This

- Google Drive API quotas, permission propagation, and shared-drive behavior may increase synchronization and authorization effort.

- Permission-aware retrieval can leak content if metadata is stale or filtering is not consistently applied across search and answer paths.

- Document parsing quality, spre

adsheet
structure,
and wiki
export
formats
may
require
format-
specific rem
ediation.

- Answer
grounding
and citation
quality may
require
prompt,
chunking,
model, and
evaluation
iterations
beyond the
base
estimate.

- Retry
and resuma
bility
behavior
may be
difficult to
validate
without
defined
failure
scenarios
and
recovery
objectives.

User Stories

US-1 - Employee

As an
employee, I
want to
search the
shared
knowledge
portal so
that I can
quickly find
relevant

company
information.

**US-2 -
Employee**

As an
employee, I
want to ask
questions
in natural
language
and receive
source-
grounded
answers so
that I can
understand
internal
information
without
searching
multiple
systems.

**US-3 -
Employee**

As an
employee, I
want to
view the
sources
supporting
an answer
so that I
can verify
the
information
and
investigate
further.

**US-4 -
Employee**

As an
employee, I
want
search and
answer
results to
respect my
existing
source-
system per
missions
so that I do

not retrieve
unauthorize
d company
information.

**US-5 -
Employee**

As an
employee, I
want to find
information
from
documents
and wiki
content so
that
knowledge
stored in
different
formats is
available in
one place.

**US-6 - Adm
inistrator**

As an admi
nistrator, I
want to
connect
Google
Drive so
that
company
knowledge
can be
included in
the shared
portal.

**US-7 - Adm
inistrator**

As an admi
nistrator, I
want to
configure
and
manage
the Google
Drive
connection
so that the
portal can
access the
intended
company

knowledge
source.

US-8 - Administrator

As an administrator, I want to view Google Drive synchronization status so that I know whether the connected knowledge source is current.

US-9 - Administrator

As an administrator, I want to troubleshoot Google Drive ingestion problems so that failed or incomplete knowledge updates can be resolved.

US-10 - Administrator

As an administrator, I want customer knowledge data handled with audit-ready security controls so that the organization can

support
SOC 2 and
HIPAA-
ready deplo
yments.

Require ments

FR-1 - Search

The portal
shall allow
employees
to search
the shared
knowledge
base using
keywords
and
phrases.

FR-2 - Question Answering

The portal
shall allow
employees
to ask
questions
in natural
language
and receive
answers
grounded
in
authorized
internal
content.

FR-3 - Source Attribution

The portal
shall
display the
internal
sources
supporting
each
answer.

FR-4 - Auth orization

The portal
shall
enforce

source-
system
user and
group permissions
when
processing
searches
and
generating
answers.

**FR-5 -
Content
Ingestion**
The portal
shall ingest
and index
PDFs, spreadsheets,
and wiki
content for
search and
question
answering.

**FR-6 -
Google
Drive
Integration**
The portal
shall allow administrators
to
connect
Google
Drive as a
knowledge
source.

**FR-7 -
Integration
Administration**
The portal
shall allow administrators
to
configure, authenticate,
manage,
and
disconnect
the Google
Drive
connection.

FR-8 - Synchronization

The portal shall synchronize approved Google Drive content and make synchronization status visible to administrators.

FR-9 - Ingestion Operations

The portal shall allow administrators to troubleshoot Google Drive ingestion failures by retrying, re-running, correcting configuration, and resuming synchronization.

FR-10 - Audit Administration

The portal shall record administrative, integration, synchronization, access, and security-relevant events in an audit trail.

NFR-1 - Security

The system shall protect customer knowledge and enforce least-privilege access across users, administrators, integrations, and stored data.

NFR-2 - Compliance

The system shall provide audit-ready security controls and evidence suitable for SOC 2-oriented deployments.

NFR-3 - Healthcare Security

The system shall support HIPAA-ready safeguards for customers handling HIPAA-sensitive data.

NFR-4 - Data Residency

Customer data, indexed content, generated answers, and audit records shall be hosted in the United States.

NFR-5 - Performance

The system shall provide responsive search and answer interactions, with standard API responses targeted below 200 milliseconds where applicable.

NFR-6 - Reliability

The system shall preserve synchronization integrity and support recovery from interrupted or failed Google Drive ingestion jobs.

NFR-7 - Observability

The system shall provide monitoring, logging, and performance visibility for portal requests, answer generation, integrations, and synchronization jobs.

NFR-8 - Scalability

The system shall support growth in employees, indexed documents, search traffic, and synchronization workload for mid-market customers.

Requirements Assumptions

- The initial release is a web portal, not a mobile or app-store application.
- Google Drive is the only external integration in scope for the initial release.
- HR systems, SharePoint-style knowledge bases, project management systems, and Slack are deferred.
- Standard authentication and customer identity-management capabilities are available or provided separately.
- Google Drive permissions remain the authoritative source for content access decisions.
- The initial user experience is English-only.
- Customer-specific retention, deletion, backup, disaster-recovery, and service-level targets are not yet defined.
- HIPAA-ready deployment may require customer configuration, contractual terms, and operational procedures beyond the product features.
- No document authoring, workflow approval, payments, or knowledge-content editing features are in scope.
- Answer quality, supported file-size limits, synchronization frequency, throughput targets, and concurrency limits require estimation assumptions or later confirmation.

Technical Specification

TS-1 - Web Application

Implement the employee portal and administrator console as a responsive Next.js application using React, TypeScript, Tailwind CSS, and shadcn/ui.

TS-2 - Authentication

Use Supabase Auth with JWT-based sessions and OAuth 2.0-compatible integration over HTTPS/TLS 1.2 or

later.

TS-3 - Authorization and RBAC

Enforce tenant isolation, employee and administrator roles, least-privilege access, and source-system user/group permissions during retrieval and answer generation.

TS-4 - Backend and API

Use REST APIs with optional Supabase Edge Functions for portal operations, administration, Google Drive integration, synchronization workflows, and answer generation.

TS-5 - Data and Search Storage

Use managed PostgreSQL through Supabase for tenants, users, configuration, metadata, synchronization state, audit records, and application data; use object storage for source files and a RAG vector store for indexed content.

TS-6 - Retrieval-Augmented Answering

Use the OpenAI Responses API with retrieval over authorized document and wiki chunks; return source-grounded answers with references to permitted internal sources.

TS-7 - Google Drive Integration

Provide administrator-managed Google Drive OAuth connection, configuration, content discovery, permission-aware ingestion, synchronization status, retries, resumability, and failure remediation.

TS-8 - Administration and Audit

Provide an administrator UI for connection management, synchronization monitoring, error remediation, tenant configuration, and audit-trail review.

TS-9 - Hosting and Environments

Host the web application and admin console on Vercel and backend data services on Supabase, configured for United States data residency; maintain development, staging, and production environments.

TS-10 - CI/CD and Testing

Use GitHub Actions for build, validation, security checks, migrations, and deployment workflows; provide unit, integration, and end-to-end coverage for core portal and ingestion flows.

TS-11 - Security and Compliance

Apply HTTPS/TLS, encryption in transit and at rest, secret management, tenant isolation, least privilege, secure token handling, vulnerability management, and audit-ready controls supporting SOC 2 and HIPAA-ready deployments.

TS-12 - Observability and Reliability

Provide centralized application and security logging, metrics, tracing, alerts, synchronization job visibility, error reporting, and operational dashboards for portal, APIs, retrieval, integrations, and storage.

Technical Assumptions

- The reference stack uses Next.js, React, TypeScript, Tailwind CSS, shadcn/ui, Supabase PostgreSQL/Auth, REST APIs, OpenAI Responses API, GitHub Actions, and Vercel.
- Supabase and Vercel US-region deployment options are available and can satisfy customer data-residency requirements.
- A customer identity provider or standard authentication capability is available if enterprise SSO is required.
- Google Drive OAuth credentials, API quotas, permission metadata, and customer administrative consent are available.
- Google Drive remains the authoritative source for document access permissions.
- Supabase Storage or an equivalent US-hosted object store is used for source files and ingestion artifacts.
- The RAG vector store supports tenant isolation, metadata filtering, permission-aware retrieval, and US hosting.
- Background workers or scheduled jobs are available for asynchronous Google Drive synchronization and document processing.
- Initial ingestion supports PDFs, spreadsheets, and wiki content; OCR, unsupported formats, and advanced parsing require separate estimates.
- Customer-specific retention, deletion, backup, disaster-recovery, service-level, throughput, concurrency, and synchronization-frequency targets are not yet defined.

- SOC 2-oriented and HIPAA-ready controls do not imply certification and may require organizational and contractual measures beyond product implementation.
- No document authoring, workflow approval, payments, mobile application, or deferred external integrations are included in the initial release.

To get your own customised cost for an app like this, visit <https://www.chatstack.app> and start a ChatStack interview.

This is a starting estimate for a custom-built MVP, designed and built from scratch. You own all the code and IP with no platform lock-in. Enterprise RAG is a trademark of its owner; ChatStack / App Developer Studio is not affiliated with, endorsed by, or sponsored by Enterprise RAG. © 2026 ChatStack.

Created with ChatStack (www.chatstack.app)