Anar Labs

Anar Docs

Read the paperwork once.

Anar Docs turns PDFs, scans and Word files into text, document types and key fields, with the model endpoint you configure.

For
Records and digital service teams

Date
November 2026

The problem

Arabic paperwork waits for someone to type it.

01

Text is locked in images

Scans and photos hold the text. Staff type it again before a system can use it.

02

Every file needs a label

Someone reads each file to decide what type of document it is.

03

Answers stay in files

An assistant cannot answer from a document that nobody has turned into text.

What it is

Upload a file. Get text, a type and fields.

01

Your file

A PDF, a Word file or an image, up to 50 MB. Seven file formats.

→

02

Anar Docs

Reads the text, sets the document type, extracts fields and checks them.

→

03

Your systems

Text, fields and chunks return as JSON. Chunks can go to Anar Chat.

Anar supplies no models. Scanned-page reading and field extraction use the model endpoint you configure.

Reading

One call returns text, tables, metadata and chunks.

Digital files

Digital PDF and Word files give exact text and tables. No model is needed.

Shipped

Scanned pages

Scans and images go to your vision model. A local OCR engine is the fallback. Accuracy is not yet measured.

Partial

Many files

A batch request processes many files in one call.

Shipped

Shipped Partial Roadmap

Understanding

Docs sets the type, finds GCC data and pulls out fields.

Document type

Ten document types. Docs asks your model first, then uses Arabic and English keywords.

Shipped

GCC metadata

Finds Hijri and Gregorian dates, GCC national IDs, phone numbers and amounts. No model is needed.

Shipped

Field extraction

Your model reads the first 4,000 characters. Long documents lose content. A pattern fallback runs without a model.

Partial

Validation

Docs checks for required fields. Today the check reads English field names only.

Partial

Shipped Partial Roadmap

Knowledge base

Processed documents can feed Anar Chat.

Docs splits the text into chunks and sends them to a Chat knowledge base. One call can process and send.

Shipped. Turn it on per install.

Tested against a mock of Anar Chat. Chat accepts the chunks from an admin or manager account.

What Chat receives

- Text chunks ready for retrieval
- The document type
- Dates, IDs and other metadata

Status today

What works now, and what does not yet.

Capability	Status	Note
Upload, digital text and tables, batch	Shipped	Seven formats. 50 MB per file.
Document type, GCC metadata	Shipped	Ten types. Arabic and English keywords.
Scanned-page OCR	Partial	Works. Accuracy is not yet measured.
Field extraction, validation	Partial	First 4,000 characters. Validation reads English field names.
Search	Partial	Keyword match on text. No semantic search.
Handwriting, redaction, review queue	Roadmap	Not built yet.

Roadmap

The order follows what a pilot needs first.

Now

- Upload, OCR and extraction
- Type and GCC metadata
- Export to Anar Chat

Next

Roadmap

- Measured Arabic OCR accuracy
- Each field linked to its place on the page
- Personal data redaction

Later

Roadmap

- Human verification queue
- Handwritten Arabic, stamps and seals
- Arabic interface

The platform

Docs feeds other Anar products.

Anar Chat

Docs can send chunks to a Chat knowledge base. Turn it on per install.

Anar Flow

A Flow step can send a document to Docs for processing.

Anar Procure

Procure can send tender text to Docs to list requirements.

Anar Gateway

Type and field calls can go through the gateway. Scanned-page reads do not yet.

See it read your documents.

A one-month pilot: we deploy in your environment and run it with one team.

anarlabs.ai/contact