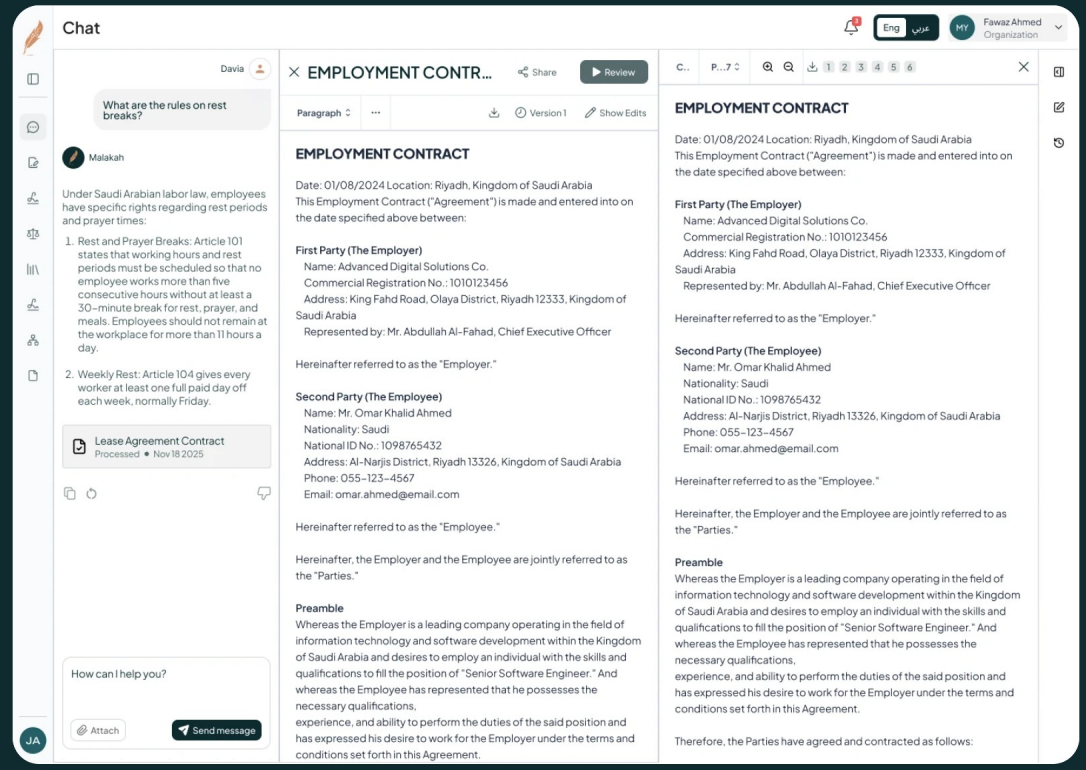
CASE STUDY · AI TRUST · SOURCE ATTRIBUTION

Making a legal AI show its sources, *down to the article*

Every claim in Malakah's answers opens the exact article of Saudi law it came from.

Abdullah Omer

Product Designer · Stixor, for Malakah · 2024–2026



AT A GLANCE

Make every claim cheap to check

Problem

10 of 12 lawyers would re-check any AI answer without a source. Unverifiable answers moved time, not saved it.

Reframe

Not “make the answer more accurate”, but design the path from a sentence to the law behind it.

What I designed

Four citation patterns compared, a legislation CMS that splits laws into articles, and a retrieval evaluation tool.

Outcome

Article-level citations shipped to 1,000+ legal professionals. One lawyer: “It saves me 90% of my time.”

01 · THE PROBLEM

An answer you can’t check is a task

10 of 12

would re-check any AI answer without a source

8 of 12

drafted in Arabic first

1,000+

legal professionals on the platform

DESIGN GOAL

From any claim, reach the exact article it rests on, in its current version, in one click, without leaving the conversation.

02 · APPROACH

One system, three layers, three users

1

Structure the law

Legislation CMS: chapters, parts, articles, annexes, amendments.

Users: content team

2

Measure retrieval

Evaluation tool: do the right articles come back, in the right order?

Users: AI team

3

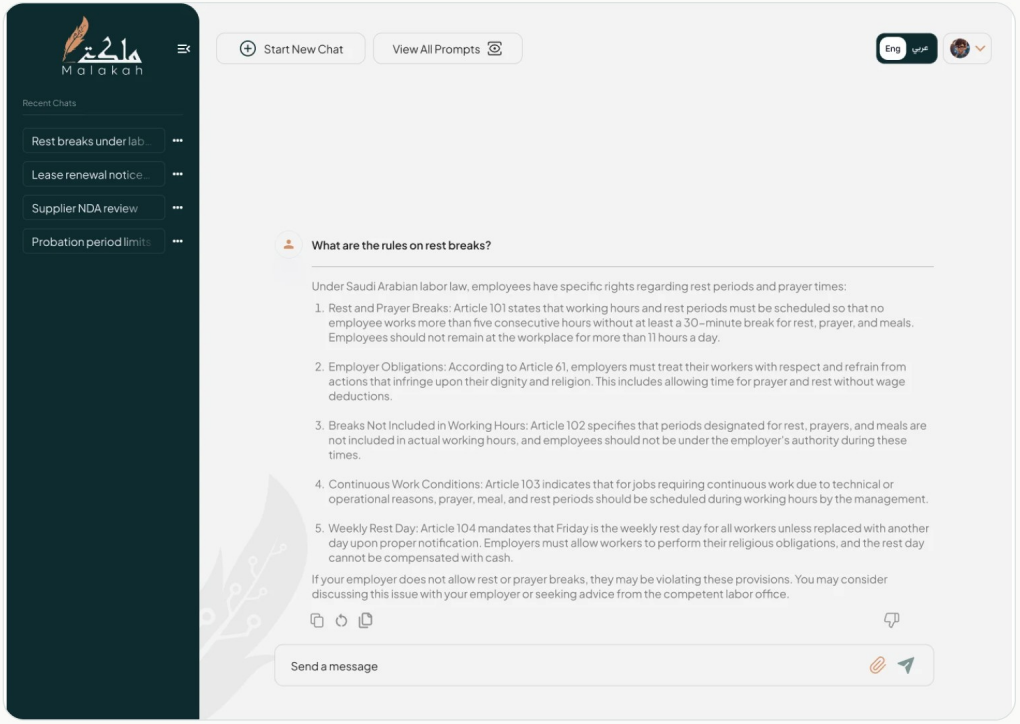
Show the source

Legal chat: each claim links to its article, opened beside the answer.

Users: lawyers

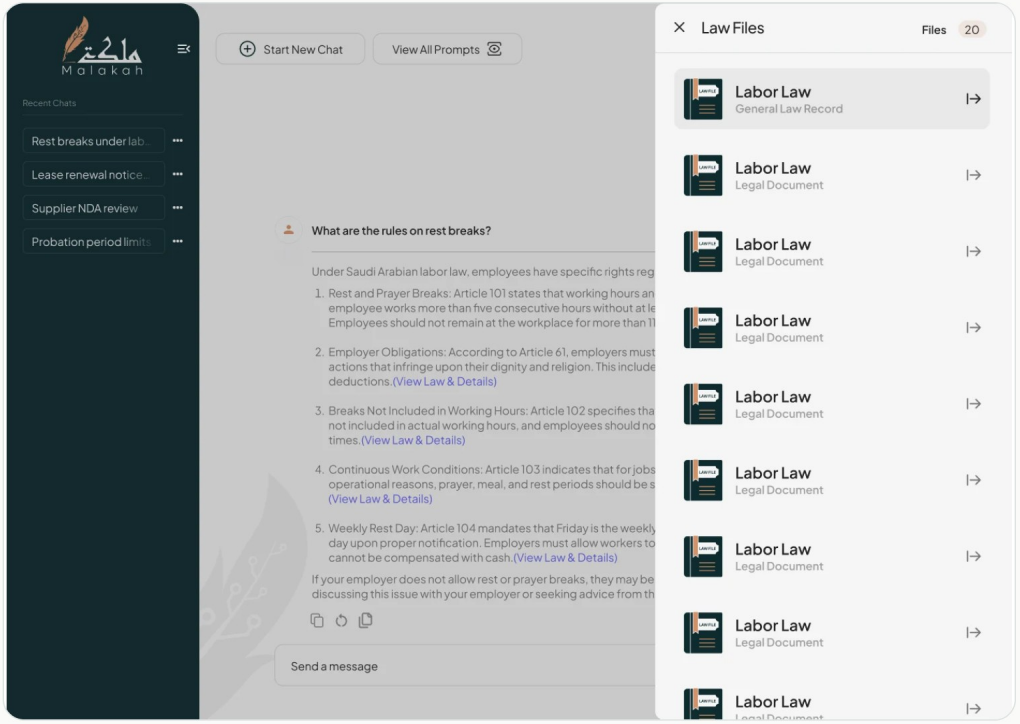
03 · EXPLORATION · OPTIONS A AND B

How much checking is left to the lawyer?



A · No sources

Every article number is a lookup elsewhere.

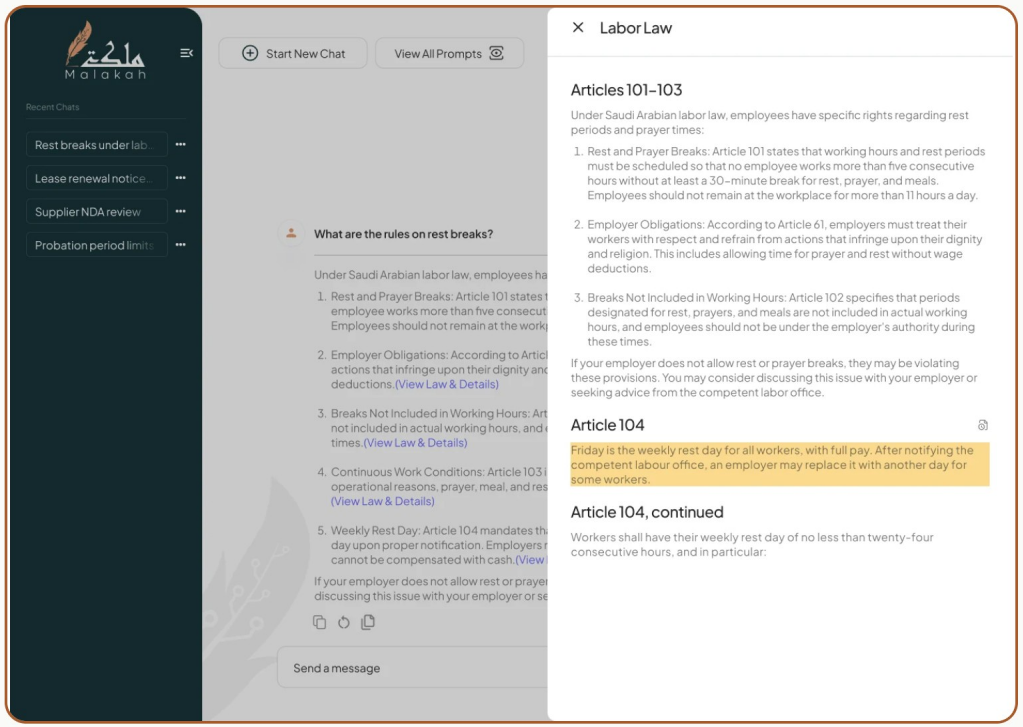


B · A list of law files

20 files, all “Labor Law”. Which law, not which article.

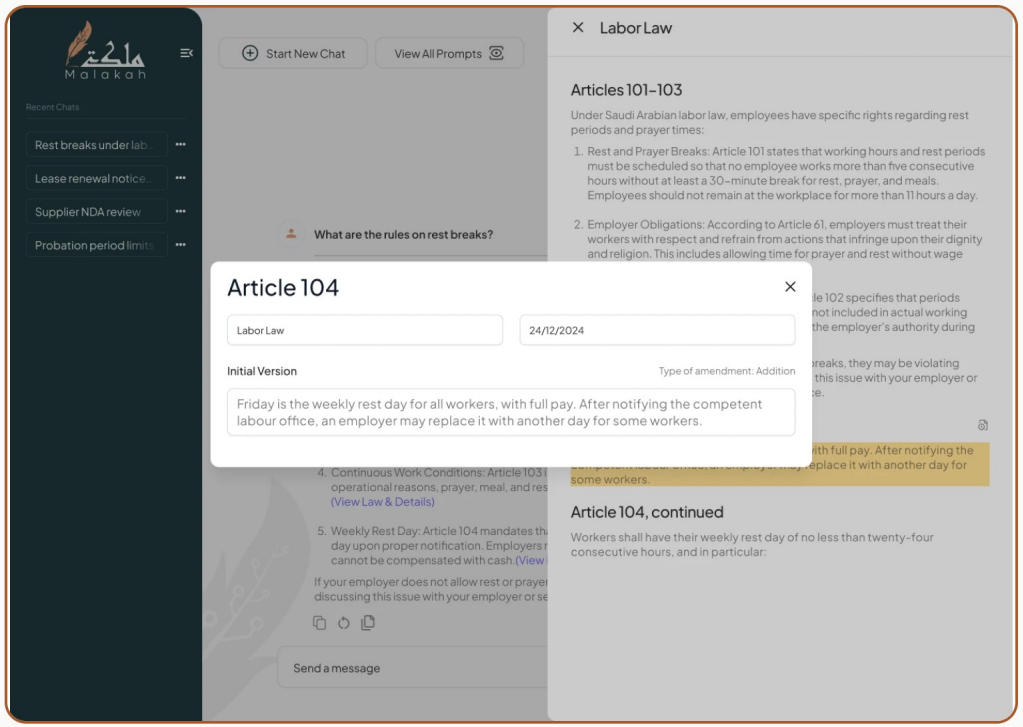
03 · EXPLORATION · OPTIONS C AND D

A link on every claim, with its version



C · Claim links to the exact article

Opens beside the answer, cited text highlighted.



D · Plus version history

Date, amendment type and earlier text.

03 · EXPLORATION

How the four options compare

Option	What the lawyer gets	Work left to verify
A · No sources	Article numbers in prose	Find and read each article elsewhere
B · File list	The right laws, not the right passages	Open a file, search, match it to a sentence
C · Claim links	The exact article, highlighted, beside the answer	Read the highlighted text
D · + Version	The article and when it last changed	Confirm the version is current

C and D shipped. The source opens as a third panel next to the chat and the contract, so lawyers check claims **while drafting**.

04 · KEY DECISION

Cite the article, not the law

SIMPLER · NOT CHOSEN

Cite whole laws

Easy to build and always “correct”, but it pushes the real work, finding the passage, back onto the lawyer.

CHOSEN

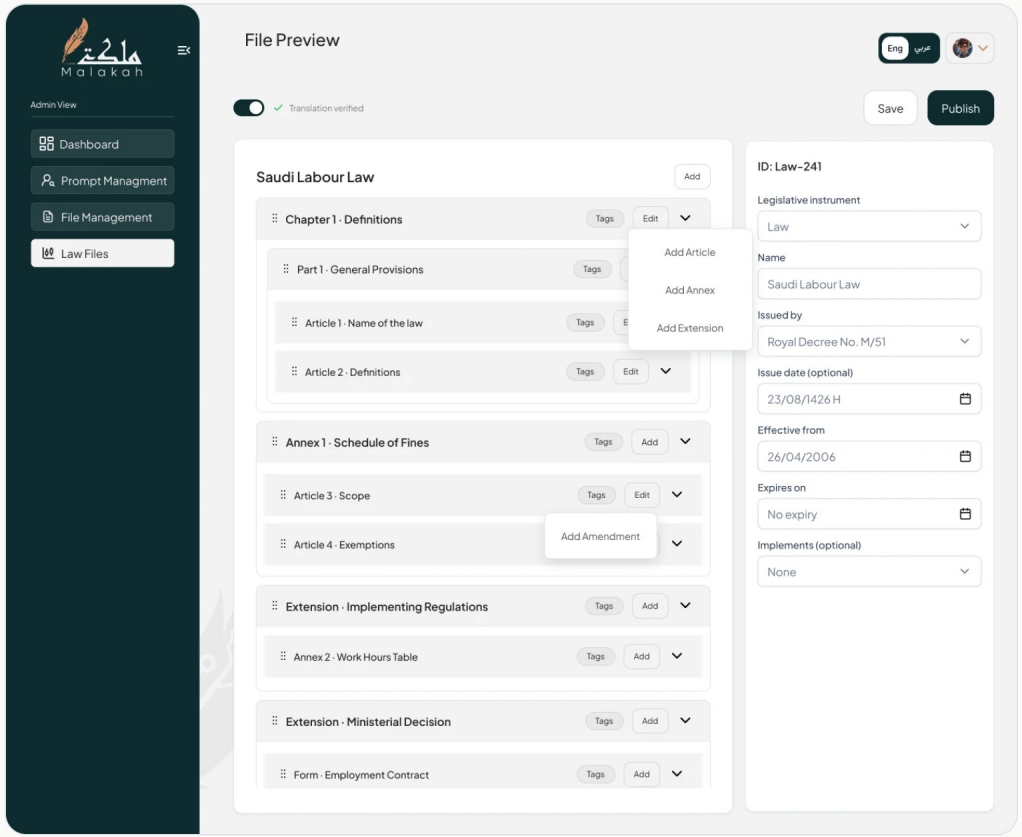
Cite articles

Far more work behind the scenes: every law stored article by article, and retrieval that finds the right one.

That cost is what the next two layers are about.

05 · LAYER 1 · CONTENT

Making the law addressable



A tree that matches how lawyers cite

Chapter, part, article; annexes and amendments as their own blocks.

“How binding is this?”

Instrument, issuing decree, effective and expiry dates.

Source priority

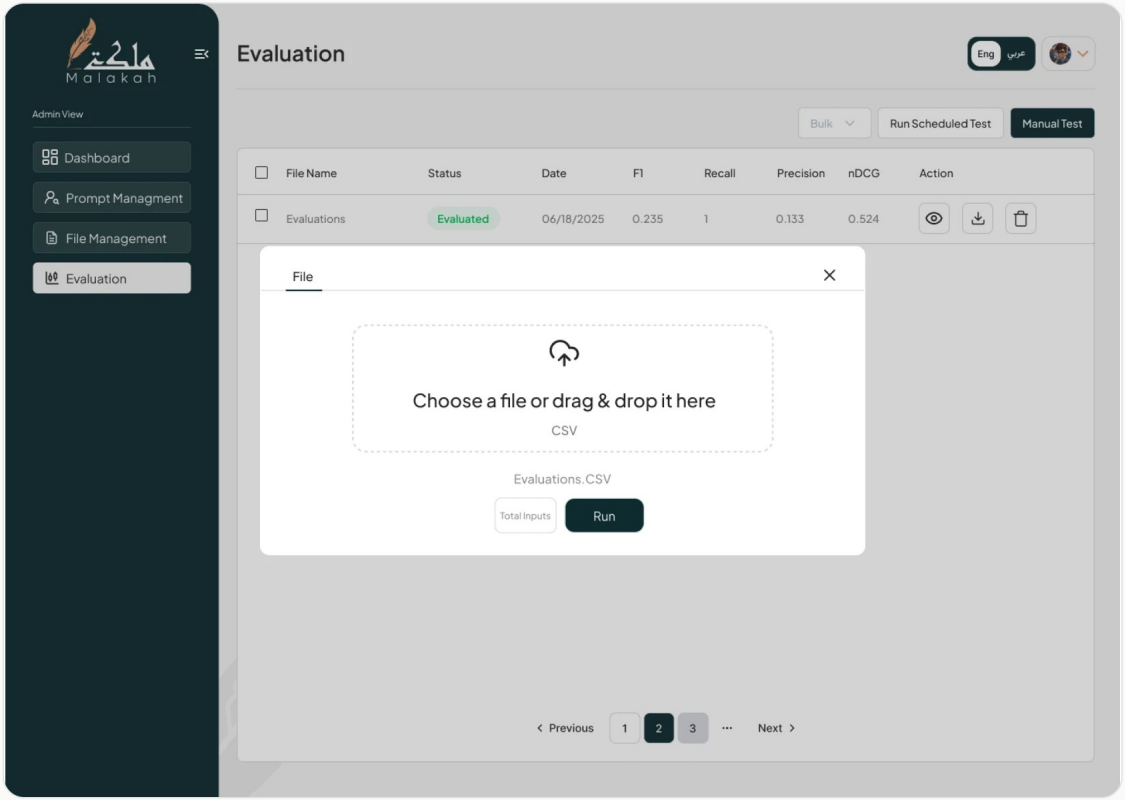
Definitional, authoritative or informative, so “the law says” differs from “a guideline suggests”.

Translation check by Publish

An unverified translation is a citation to the wrong text.

05 · LAYER 2 · QUALITY

Does the right article come back?



Test sets as files

Upload a CSV, run it manually or on a schedule after every change.

The right metrics, side by side

Recall: found at all? Precision: how much noise? nDCG: ranked first?

Runs as a history

Spot a regression and trace it to a change.

06 · MEASURING SUCCESS

Trust you can measure

Task success

Time to verified

Answer appearing to lawyer moving on, with vs without citations.

Engagement

Citation open rate

Falling over time with steady quality means earned trust.

Happiness

Thumbs-down rate

Split by whether the cited article was right.

Quality · tracked

nDCG on the test set

Does the correct article rank first, run by run.

Coverage

Article-level claims

Points linking to an article, not a whole law.

Adoption

Weekly active lawyers

Do teams come back for real work?

“It saves me 90% of my time.”

— Lawyer using Malakah, feedback after launch

One lawyer’s own estimate, not a measured average.

07 · REFLECTION

What I took from this

01

Trust is a path, not a badge

“Sources” didn’t help until a sentence was one click from its article.

02

Visible patterns need invisible structure

The link was easy to draw. The CMS and evaluation tool made it true.

03

Design for the people behind the AI

The content and AI teams were users as much as lawyers were.

NEXT Test citation density on long answers: grouped per section vs inline, by time to verify.

Thank you.

Questions are welcome, on AI trust, the trade-off or the internal tools.

NAME

Abdullah Omer

EMAIL

abdullah542903@gmail.com