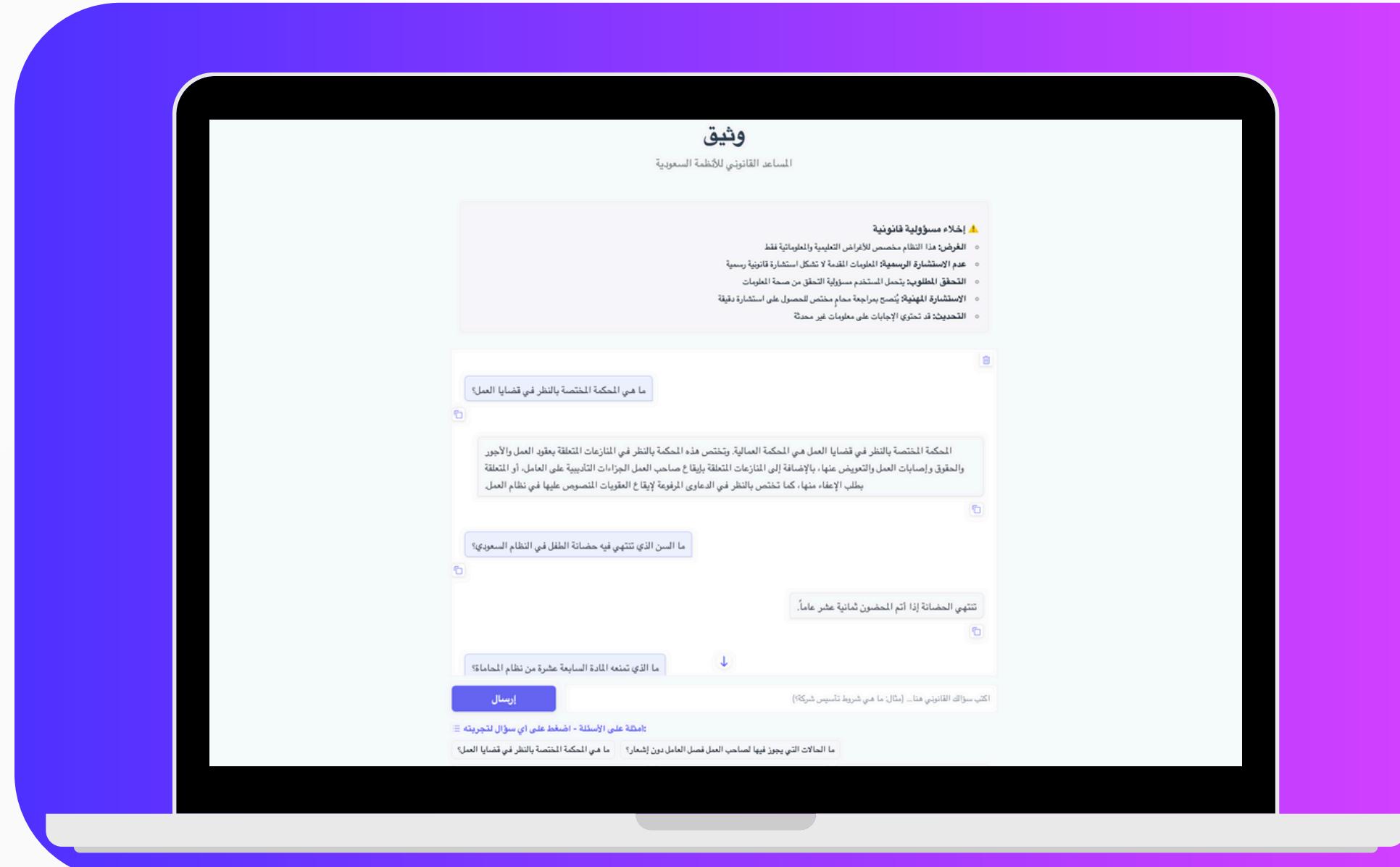


# وَثِيق Watheeq

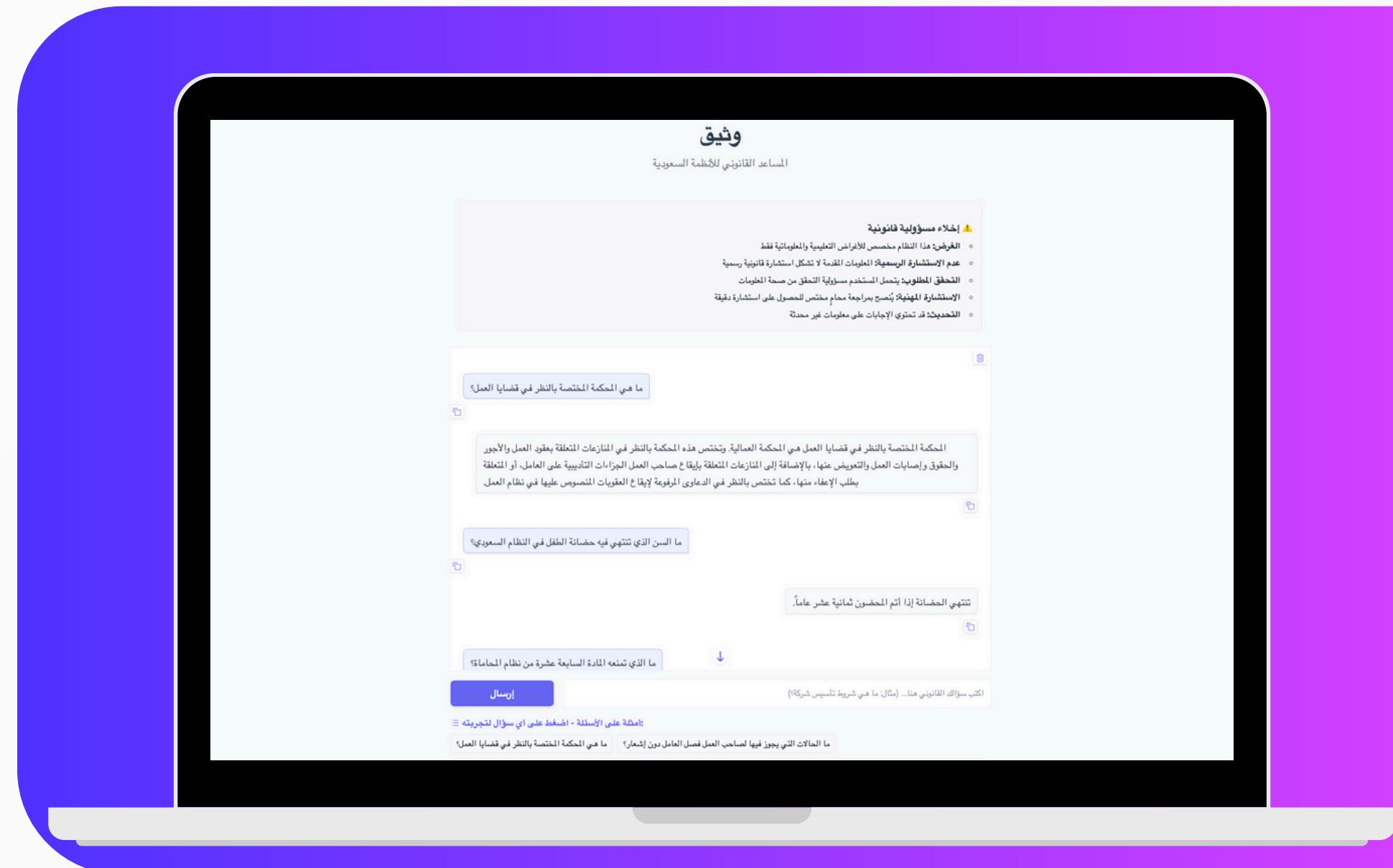
Shahad Almutairi  
Reem Alsahli  
Taif Alharbi  
Yasminah Alali



# Project Overview

**Watheeq** is a legal assistant chatbot designed to simplify access to Saudi Arabian laws and regulations. Legal documents are often complex and difficult for the general public to interpret.

Watheeq helps users understand their rights and obligations by answering legal questions in clear, natural Arabic supporting Saudi Arabia's digital transformation efforts.



# Motivation Behind Watheeq

## Complex Legal Language

Saudi laws are often written in formal legal Arabic, making them hard to understand for non-specialists.

## Time-Consuming Legal Research

Finding relevant legal information requires navigating long, unstructured documents.

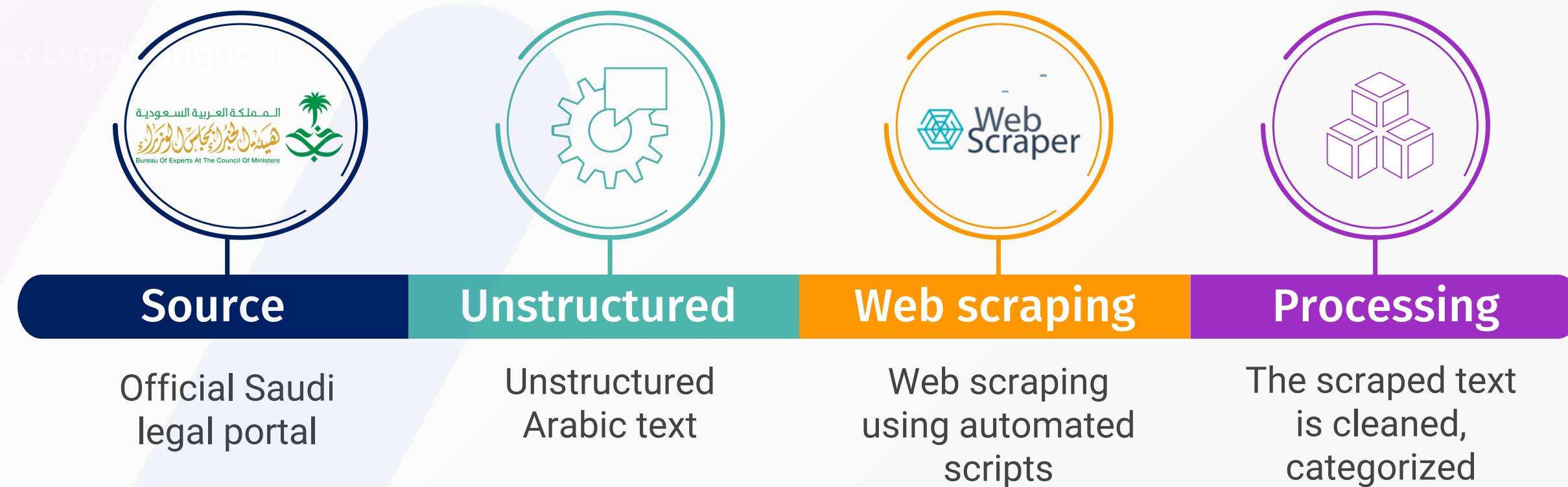
## Lack of Centralized Access

There's no unified tool that lets user ask legal questions and get reliable answers instantly.



# Data Collection

We use unstructured text data sourced from the official website of هيئة الخبراء ب مجلس الوزراء (Bureau of Experts), which publishes Saudi laws and regulations. The data includes the full legal texts of various systems (labor, personal status, business laws, etc.).



# Data Preprocessing

## 1. Load Legal Documents

We loaded our scraped legal data, which includes systems like Labor Law, Personal Status Law, and Law of Lawyers.

## 2. Split into Chunks

The documents were split into smaller text chunks for better context during search.

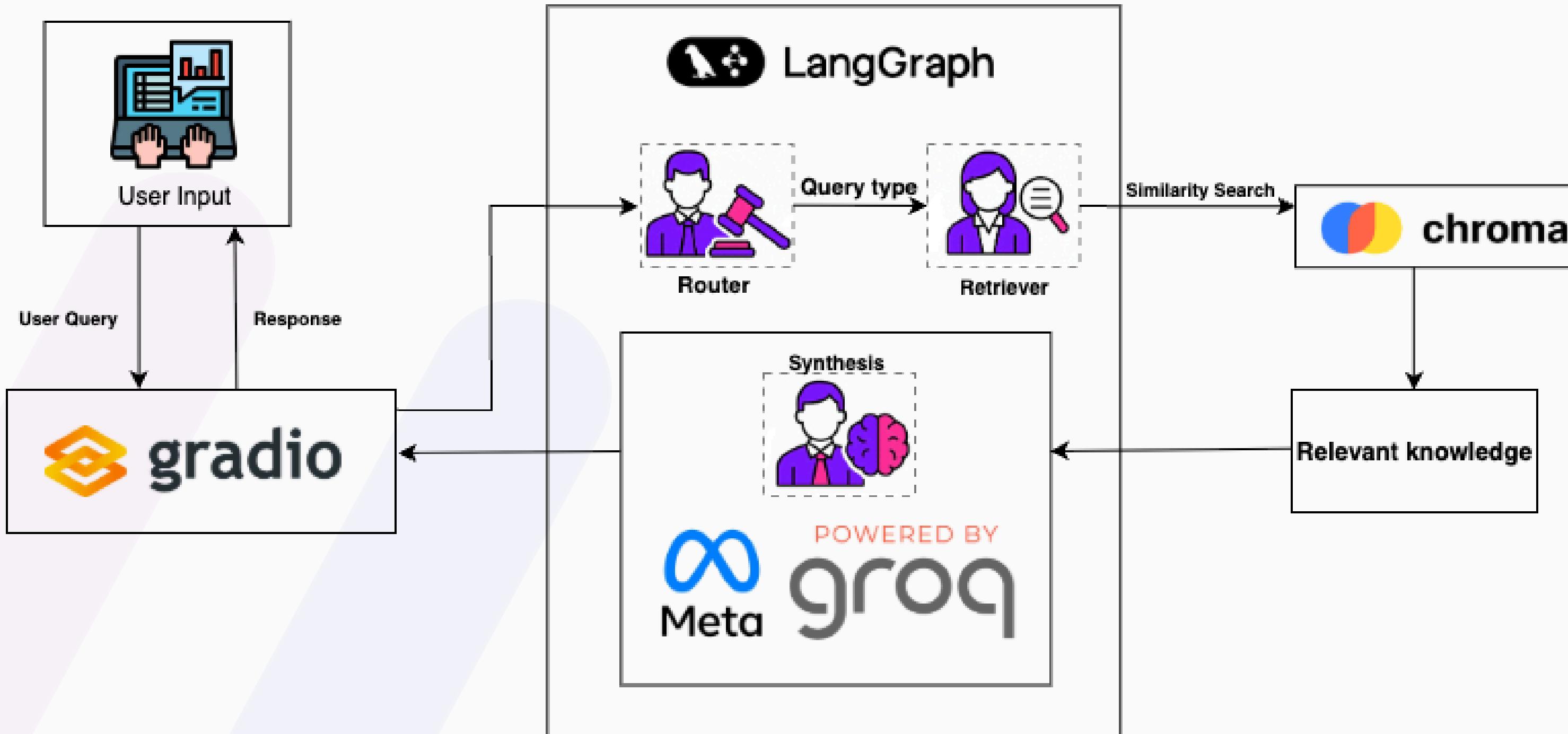
## 3. Create Embeddings

We used Cohere's multilingual embedding model to convert each chunk into vector form.

## 4. Store in ChromaDB

All embeddings were saved into a Chroma vector database, ready for fast and accurate retrieval.

# Architecture



# LLM Selection

## Primary Generation Model:

- **Groq LLaMA 3.3 (70B Versatile)**

- Delivers accurate, context-aware responses to Arabic legal queries
- Optimized for fast, low-latency inference via Groq API
- Temperature tuned to 0.3 for concise and factual answers



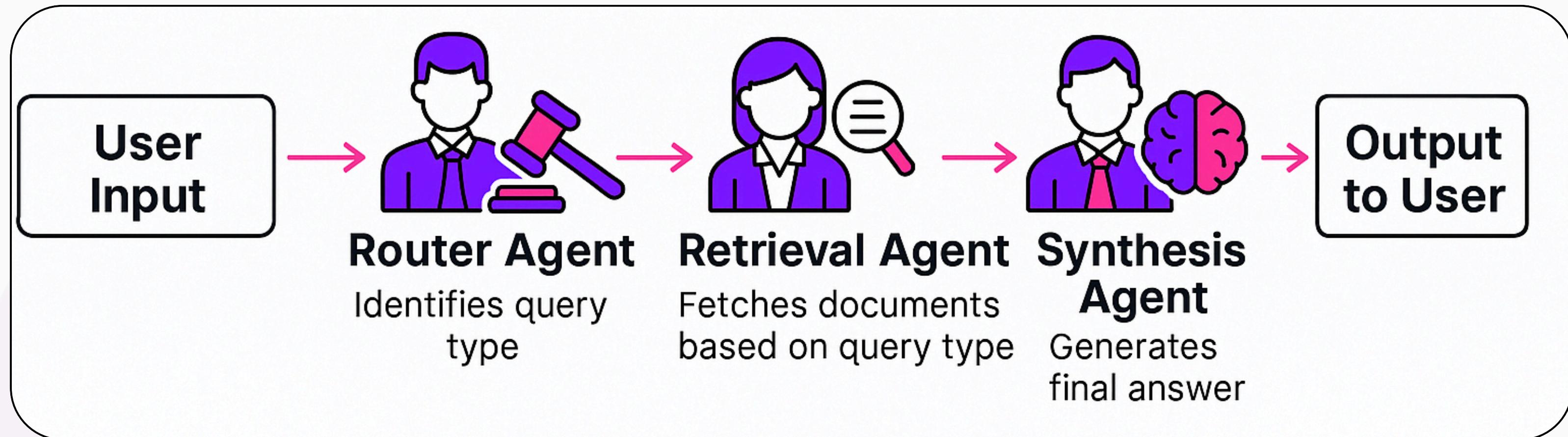
## Router Model (for Query Classification):

- **Model: XLM-RoBERTa Large (XNLI)**

- **Purpose:**

- Classifies user questions into categories (Labor Law, Commercial Law, Personal Status, Legal Procedures, etc.)
- Ensures the system retrieves the most relevant documents before the LLM generates the answer
- Provides a fallback keyword-based classifier for resilience

# Multi-Agent Workflow



# User Interface

## Gradio (Custom-Styled)

- RTL (Right-to-Left) Arabic support
- Clean, mobile-friendly layout
- Custom disclaimer and branding

## Key UI Features:

- Interactive Chatbot
  - Users can ask legal questions in both Arabic and English
  - Answers generated using retrieved laws
- Preloaded Example Queries
  - Quick start for common legal topics (e.g., labor, commercial)
- Disclaimer Box
  - Clarifies non-official advisory role

# Demo

## وثيق

المساعد القانوني لأنظمة السعودية

إخلاء مسؤولية قانونية

الفرض: هذا النظام مخصص للأغراض التعليمية والمعلوماتية فقط

عدم الاستشارة الرسمية: المعلومات المقدمة لا تشكل استشارة قانونية رسمية

التحقق المطلوب: يتحمل المستخدم مسؤولية التحقق من صحة المعلومات

الاستشارة المهنية: ينصح بمراجعة محام مختص للحصول على استشارة دقيقة

التحذير: قد تحتوي الإجابات على معلومات غير محدثة

مرحباً! كيف يمكنني مساعدتك في الأسئلة القانونية؟

ما هي المحكمة المختصة بالنظر في قضايا العمل؟

أمثلة على الأسئلة - اضغط على أي سؤال لتجريمه:

ما الحالات التي يجوز فيها لصاحب العمل فصل العامل دون إشعار؟

ما هي المحكمة المختصة بالنظر في قضايا العمل؟

ما شرط الحصول على ترخيص مزاولة مهنة المحاماة؟

ما السن الذي تنتهي فيه حضانة الطفل في النظام السعودي؟

ما هي العقوبات المقررة لجرائم الاحتيال المالي؟

إرسال

# Evaluation

## Evaluation Metric: LLM-as-Judge

Evaluate the accuracy of an LLM response by comparing it against the most relevant top documents retrieved for the user query.

Scoring Scale (0 to 1):

- **1.0** → Fully correct answer based on top documents
- **0.5 to 0.9** → Partially correct or mostly relevant
- **0.1 to 0.49** → Low relevance or mostly incorrect
- **0.0** → Completely unanswered or irrelevant

Overall accuracy :0.91

# Future Work

## 1. Comprehensive Extraction from هيئة الخبراء ب مجلس الوزراء

Expand document coverage by extracting and structuring all available legal content and systems issued by the Bureau of Experts.

## 2. Improved Chunking Strategy

Refine document chunking techniques to preserve semantic coherence and improve retrieval accuracy especially for long legal texts and nested articles.

## 3. Support for More Legal Systems

Scale the pipeline to include additional systems and regulations beyond the current dataset

## 4. Advanced File Upload Handling

Improve parsing, validation, and pre-processing for uploaded legal PDFs, ensuring clean and accurate ingestion across various formats and layouts.

# Conclusion

**Watheeq** has successfully demonstrated accurate legal question answering based on six Saudi legal documents, delivering responses in clear and natural Arabic. It maintains high precision by avoiding hallucinations and refusing to answer questions beyond its supported scope, ensuring trustworthy interactions aligned with national legal standards. By simplifying complex legal language for the public, **Watheeq** contributes to broader legal accessibility and understanding.

