

WebVM

Manuale d'uso e documentazione tecnica aggiornata

Il progetto WebVM è un servizio web di intelligenza artificiale, ossia è un'Application Pool del [Portale degli studi notarili](#) costruita su ASP.NET Core in Visual Studio 2026 – Codice C#, che consente agli iscritti del Portale di interrogare in linguaggio naturale i documenti PDF archiviati e indicizzati nei diversi Settori del Portale.

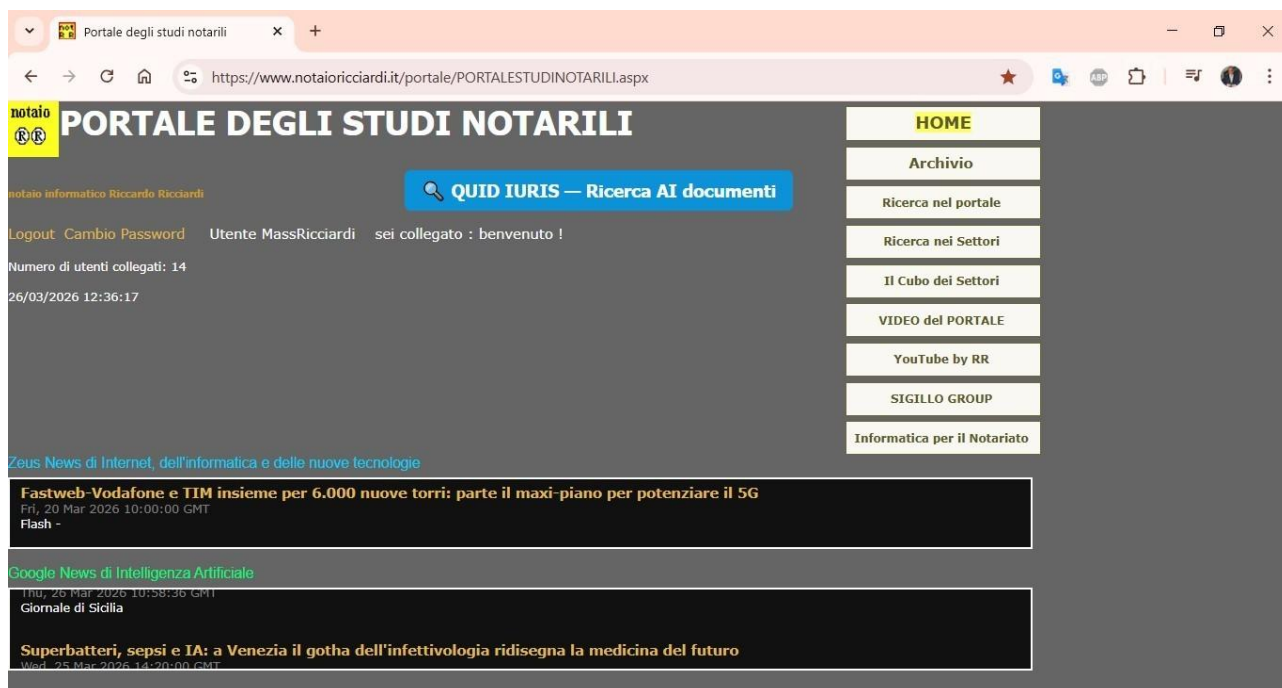
Il servizio legge esclusivamente i documenti selezionati e genera risposte basate sul loro contenuto, indicando sempre il documento, la pagina e il punto da cui è tratta l'informazione.

Rispetto alla versione precedente sono state inserite nuove funzionalità:

- Gemini 3.6 Flash per la generazione delle risposte.
- Ricerca full-text in più settori, nuova pagina di ricerca che si apre tramite il collegamento presente nella pagina Ask.
- Ricerca full-text nel contenuto dei PDF di un settore selezionato tramite il database fts-search.db.
- Filtro dei PDF per parole contenute nel testo.
- Aggiornamento automatico dell'elenco dei PDF ai soli file selezionati per generare la risposta
- Pulsante Visualizza PDF selezionati.
- Azzeramento immediato della risposta quando si cambia stile.
- Tre modalità di Domanda singola: Domanda singola: risposta unica - Domanda singola: risposte separate - Domanda singola: sintesi unica

1. Accesso alla pagina di ricerca

Dal Portale degli studi notarili cliccare su “Quid Iuris – Ricerca AI documenti”.



2. Settore e modello

Cliccando su : Quid Iuris – Ricerca AI documenti si apre la seguente pagina di ricerca, che consente di scegliere un Settore indicizzato, filtrare i PDF con ricerca dei file nel contenuto, selezionare lo stile, formulare una domanda e ottenere una risposta generata dal nuovo modello Gemini 3.6 Flash.

“L'estrazione, attraverso il Portale degli studi notarili, dell'acqua nel deserto (metafora potente che indica qualcosa di raro, prezioso e disperatamente necessario) era solo il primo passo di un processo complesso realizzato dal magico ErreErre, mirante a trasformare aridi Pdf in Oasi di formazione e aggiornamento professionale.”

[Istruzioni d'uso \(WebVM.pdf\)](#) [Ricerca full-text nei settori](#) [Nuova ricerca](#)

SETTORE AGGIORNAMENTI
Modello Gemini 3.6 Flash
Stile NotebookLM (discorsivo)

NotebookLM: per risultati ottimali seleziona al massimo 5 file.

Fonti (file)
Seleziona i file da interrogare. Tutti Nessuno (agisce solo sui file visibili)

Cerca file Visibili: 32/32

Apri PDF selezionati Seleziona uno o più file per aprirli.

- ☐ 2026 30 Rassegna novità giurisprudenziali n. 30_2026.pdf

Istruzioni d'uso

Apri il presente file

Ricerca full text nei settori

Permette di cercare parole nel testo dei PDF indicizzati in WebVM, anche quando quelle parole non compaiono nel nome del file. Questa nuova pagina di ricerca si apre tramite il collegamento presente nella pagina Ask. Nella pagina di ricerca sono mostrati contemporaneamente i settori disponibili. L'utente può selezionarne uno, più di uno o tutti. I risultati sono presentati in una pagina separata. Per ogni PDF trovato vengono indicati ****il settore, il nome del file e un collegamento per aprirlo****. Se la ricerca non può essere completata in un settore selezionato, la pagina lo segnala separatamente: un errore non viene presentato come «nessun PDF trovato».

[← Torna ad Ask](#)

Ricerca nel testo dei PDF

Parole da cercare

Scegli uno o più settori. AGGIORNAMENTI non è incluso.

☐ ANTIRICICLAGGIO

☐ COOPERATIVE

☐ DIRITTO INTERNAZIONALE PRIVATO

☐ MUTUI

☐ PRASSI IMPOSTE

☐ RR_II

☐ SUCCESSIONI DONAZIONI/DICHIARAZIONE SUCCESSIONE/

☐ URBANISTICA

☐ ATTO INFORMATICO

☐ CRISI IMPRESA

☐ ESECUZIONI IMMOBILIARI

☐ NOTARTEL

☐ PRIMA CASA

☐ SOCIETA CAPITALI

☐ TASSAZIONE

☐ VOLONTARIA GIURISDIZIONE

☐ CATASTO

☐ DEONTOLOGIA ISPEZIONI

☐ MEDIAZIONE

☐ PERSONE FAMIGLIA

☐ REDDITI

☐ SOCIETA PERSONALI

☐ TERZO SETTORE

☐ CERTIFICAZIONE ENERGETICA

☐ DICHIARAZIONE SUCCESSIONE

☐ MODELLO UNICO

☐ PLUSVALENZE

☐ REGISTRO IMPRESE

☐ SUCCESSIONI DONAZIONI

☐ TRUST

Seleziona tutti

Deseleziona tutti

Cerca nei settori selezionati

Nuova ricerca

Azzerare la ricerca corrente e riaprire la pagina senza la risposta precedente, le fonti e le selezioni della sessione.

Settore

Dopo aver scelto un Settore, vengono elencati i PDF disponibili nel Settore selezionato. Ogni Settore dispone di due database: **rag.db**, usato dalla ricerca RAG e **fts-search.db**, usato per la ricerca testuale nel contenuto dei PDF.

Modello

Gemini 3.6 Flash è il nuovo più performante modello utilizzato da WebVM per generare le risposte AI, al posto del precedente Gemini 2.5 Flash-Lite. Google lo descrive come un modello generalista ad alta efficienza, progettato per conciliare velocità, capacità di ragionamento e contenimento dei costi. Rispetto alla precedente generazione, Gemini 3.6 Flash offre prestazioni migliorate nel lavoro basato sulla conoscenza, nella generazione di codice e nell'elaborazione multimodale, con maggiore efficienza nell'uso dei token di output e minore tendenza a risposte inutilmente verbose. Il modello supporta output strutturati, caratteristica utilizzata da WebVM per ottenere risposte, riepiloghi e FAQ in formato controllato.

3. Ricerca dei file nel contenuto

Dopo aver selezionato un Settore vengono elencati in ordine alfabetico tutti i PDF contenuti nel Settore e il contatore mostra il rapporto tra PDF selezionati/PDF visibili.

Il campo “Cerca file” è una ricerca full-text nel contenuto dei PDF. Non cerca più soltanto nel nome, come nella versione precedente, ma ora invia le parole al database **fts-search.db** e rende visibili solo i PDF che contengono nel testo tutti i termini indicati.

È possibile inserire una parola oppure più termini caratteristici.

La **ricerca full-text** comporta i seguenti miglioramenti:

- seleziona rapidamente i documenti pertinenti;
- riduce il contesto inviato al modello di intelligenza artificiale;
- migliora la pertinenza e riduce i tempi della risposta;
- consente di cercare espressioni tecniche o termini giuridici presenti nei documenti.

Suggerimento: la ricerca di una frase molto lunga può essere meno efficace se termini rilevanti sono presenti in parti diverse del documento. Per ottenere un filtro preciso è consigliabile usare due-cinque parole tecniche e distintive.

4. Selezione e visualizzazione PDF

I pulsanti “**Tutti**” e “**Nessuno**” agiscono soltanto sui file attualmente visibili. Dopo una risposta, restano visibili i PDF selezionati e interrogati.

Selezionando un PDF si attiva il pulsante “**Visualizza PDF selezionati**”, che apre il documento in una nuova scheda. Se si selezionano più PDF, cliccando su “**Visualizza PDF selezionati**”, viene aperta una pagina HTML con i link ai PDF selezionati.

5. Modalità di risposta

Il menu “**Stile**” controlla il formato della risposta. Il numero massimo di file e il numero di chiamate AI cambiano secondo la modalità.

Modalità	Max file	Chiamate AI	Output	Uso
NotebookLM (discorsivo)	5	1 per file	Un riepilogo narrativo per ciascun documento	Comprensione generale
Domanda singola: risposta unica	8	1 complessiva	Una risposta unica sui file selezionati	Risposta mirata a una domanda
Domanda singola: risposte separate	6	1 per file	Da 3 a 8 punti separati per ciascun documento	Confronto tra documenti e requisiti
Domanda singola: sintesi unica	6	1 complessiva	Un unico elenco che integra i file selezionati	Sintesi trasversale dello stesso tema

Modalità	Max file	Chiamate AI	Output	Uso
FAQ automatico	3	1 + una chiamata per ogni domanda generata	Domande generate e relative risposte	Esplorazione automatica

NotebookLM (discorsivo)

Genera un riepilogo discorsivo del documento o dei documenti selezionati senza richiedere una domanda sostanziale. Il campo viene impostato automaticamente su “fammi un riepilogo”. La risposta viene redatta in italiano corretto, con paragrafi leggibili e contenuto ancorato ai PDF. Ideale per ottenere spiegazioni approfondite su temi complessi con tono espositivo e collegamenti tra concetti. Modalità RAG classica. È la scelta più adatta quando occorre una risposta mirata a una questione specifica.

Domanda singola: risposta unica

Formula una domanda chiara ed esplicita: Gemini produce un’unica risposta fondata sui file selezionati, indicando i riferimenti al documento, alla pagina e al PDF di origine.

Domanda singola: risposte separate

Formula una domanda chiara ed esplicita: per ciascun documento selezionato, Gemini genera un elenco ordinato di punti, con informazioni rilevanti e riferimenti alle fonti.

È utile per confrontare documenti diversi sullo stesso tema o per individuare, documento per documento, requisiti, condizioni, eccezioni, conseguenze, termini e criteri applicabili.

Domanda singola: sintesi unica

Formula una domanda chiara ed esplicita: Gemini esamina congiuntamente i file selezionati e produce un solo elenco ordinato che integra le informazioni rilevanti provenienti dai vari documenti, con riferimenti alle relative fonti.

È utile quando vuoi una visione complessiva di uno stesso tema, senza ricevere una risposta distinta per ogni PDF: ad esempio, per ricostruire regole comuni, differenze, requisiti ricorrenti, eccezioni e conseguenze emerse nell’insieme dei documenti.

FAQ automatico

Gemini genera domande pertinenti sulla documentazione selezionata e produce le relative risposte con i riferimenti alle fonti.

È utile per esplorare rapidamente il contenuto dei documenti quando non hai ancora una domanda precisa o vuoi individuare i temi principali da approfondire.

6. Domanda e generazione

NotebookLM e FAQ automatico compilano automaticamente il campo della domanda.

Per Domanda singola è necessario inserire una domanda; in caso contrario compare un avviso.

Premendo “Genera”, se non è stato selezionato alcun PDF compare un avviso. Durante l’elaborazione vengono mostrati “Generazione in corso...” e il pulsante rosso “Annulla”. Il cambio di modalità azzerà immediatamente la risposta precedente, le fonti e gli eventuali avvisi, senza dover ricorrere ad una nuova ricerca.

Dopo la risposta, la scelta di un nuovo stile azzerà la risposta, ma conserva Settore e PDF selezionati.

Le risposte riportano i riferimenti ai documenti selezionati.

7. Fonti e riferimenti

Le fonti includono:

- nome del PDF;
- pagina;

- indice del passaggio utilizzato;
- link diretto al documento, che consente di aprire il PDF alla pagina indicata.

Se le fonti non consentono una risposta univoca, il modello dichiara che l'informazione non è presente nel contesto senza aggiungere deduzioni esterne.

8. Documentazione tecnica aggiornata

Il sistema usa un'architettura RAG (*Retrieval-Augmented Generation*):

1. i PDF vengono indicizzati localmente tramite Ollama, app desktop;
2. il testo estratto viene salvato in un database SQLite rag.db per ogni Settore del Portale;
3. viene costruito un secondo database SQLite, fts-search.db, per la ricerca full-text nel contenuto dei PDF di ogni Settore del Portale;
4. l'utente filtra e seleziona i PDF pertinenti;
5. WebVM recupera i chunk testuali rilevanti;
6. Gemini 3.6 Flash genera una risposta basata esclusivamente sul contesto recuperato.

9. Database per Settore

Ogni Settore ha una propria cartella e utilizza due database SQLite distinti:

1) **rag.db** contiene i chunk estratti dai PDF e, per ogni chunk, informazioni quali:

- percorso del file;
- numero della pagina;
- indice del chunk;
- testo estratto;
- vettore embedding;
- dati necessari al controllo incrementale.

È il database utilizzato dalla ricerca RAG e dalla risposta AI.

2) **fts-search.db** contiene un indice SQLite FTS5 ricostruito a partire dai chunk presenti in rag.db.

È utilizzato dal campo **Cerca file** della pagina Ask per trovare PDF in base alle parole contenute nel loro testo, non soltanto nel nome del file.

La ricerca restituisce, per ciascun documento trovato:

- nome del PDF;
- numero di chunk trovati;
- pagina o pagine;
- snippet con evidenziazione del termine;

8.1. Indicizzazione locale

Estrazione e chunking

L'indicizzazione utilizza PdfPig per leggere il testo pagina per pagina. Il testo viene normalizzato e suddiviso in chunk che rispettano, per quanto possibile, confini di frasi e paragrafi. L'obiettivo è conservare porzioni di testo sufficientemente ampie da mantenere il significato giuridico, ma non così grandi da rendere imprecisa la ricerca.

Embedding

Durante l'indicizzazione, gli embedding sono prodotti localmente da Ollama e salvati in rag.db.

Ollama è quindi necessario localmente per l'indicizzazione, ma non può essere tecnicamente utilizzato sul server Aruba per rispondere alle domande degli utenti. A tal fine viene utilizzato Gemini 3.6 Flash.

Comportamento nell'interfaccia Ask

Il campo Cerca file:

- richiede almeno tre caratteri;
- cerca il testo nel contenuto dei PDF;
- mostra solo i file trovati;
- indica il numero di PDF trovati rispetto al totale del Settore;
- deseleziona automaticamente file già selezionati ma non più corrispondenti al filtro;
- aggiorna i contatori e il pulsante di visualizzazione PDF.

La ricerca di una frase molto lunga può essere meno efficace se termini rilevanti sono presenti in chunk diversi. Per ottenere un filtro preciso è consigliabile usare due-cinque parole tecniche e distintive.

Ask Recupero RAG e TextScore

Le query non generano embedding online. Il recupero dei chunk rilevanti è basato su TextScore.

TextScore considera principalmente:

- corrispondenza esatta della frase cercata;
- corrispondenza di token significativi;
- occorrenza dei termini nel testo;
- densità e lunghezza del contenuto;
- ordinamento dei documenti e dei chunk più rilevanti.

I chunk con punteggio più alto vengono passati a Gemini insieme a file, pagina e indice della fonte.

Generazione AI con Gemini 3.6 Flash

Secondo Google, Gemini 3.6 Flash è un modello generalista ad alta efficienza, progettato per combinare capacità di ragionamento, velocità e costo contenuto. Google lo presenta come un modello adatto a coding, knowledge work, compiti multimodali e workflow in più passaggi.

Nel contesto WebVM sono particolarmente utili:

- capacità di organizzare contenuti provenienti da più chunk;
- output strutturati JSON;
- buona velocità per domande singole e riepiloghi;
- gestione di risposte discorsive, elenchi e FAQ;
- maggiore efficienza dei token rispetto alla generazione Flash precedente.

Le caratteristiche del modello non sostituiscono la qualità del recupero RAG: la pertinenza finale dipende anche dai PDF selezionati, dall'estrazione del testo, dai chunk e dal prompt.

Uso consigliato

- **Per una domanda precisa:** cerca parole tecniche, seleziona uno o pochi PDF e usa Domanda singola: risposta unica.
- **Per capire rapidamente un documento:** seleziona il PDF e usa NotebookLM (discorsivo).
- **Per confrontare condizioni o requisiti:** usa Domanda singola: risposte separate.
- **Per esplorare automaticamente un documento:** usa FAQ automatico, con l'avvertenza che richiede più tempo per la risposta.

Gestione errori temporanei

In caso di indisponibilità temporanea del servizio di intelligenza artificiale, comparirà un messaggio che invita a riprovare più tardi. Questo non comporta la perdita dei dati: è sufficiente attendere qualche minuto e ripetere la richiesta.