# Roadmap: Legal PDF Summarizer – Version 1 to Version 2 Enhancements

## Version 1 (Current Capabilities)

* Upload a single legal PDF and extract text
* Summarization into structured sections:
* - Parties
* - Effective Date
* - Term
* - Confidential Info
* - Obligations
* - Jurisdiction
* - Risk Flags (general, flat text)

## Planned Version 2 Enhancements

### 1. Advanced Risk Flags

Classify risk flags into High, Medium, Low. Highlight clauses or omissions that lead to each category.

### 2. Key Clauses Extraction

Automatically extract and display critical clauses: Termination, Liability, Indemnity, Confidentiality, Payment obligations, etc.

### 3. Expanded Date Recognition

In addition to Effective Date, detect Expiry/End Dates and Renewal/Auto-renewal terms.

### 4. Named Entity Highlighting

Highlight parties (organizations/individuals), dates, and monetary amounts for quick scanning.

### 5. Downloadable Summary Report

Add a button to export the generated summary as a PDF or Word document.

### 6. Batch Processing

Enable uploading multiple documents at once to generate side-by-side summaries.

### 7. UI Improvements

Enhance layout with collapsible sections, cleaner interface, and a model selection dropdown (fast vs. detailed) to control cost and accuracy.

Testing and Performance

### Stress-Test Checklist

1. **Vary Document Lengths**
   * Upload 1 short (1-page), 1 medium (5–10 pages), and 3 long PDFs.
   * Check performance + LLM token limits (we’re slicing to 10,000 characters per file now).
2. **Filename Edge Cases**
   * Try filenames with spaces, dashes, non-ASCII characters.
3. **Empty or Corrupt PDFs**
   * Upload a blank PDF or a corrupt one to see if any errors are thrown.
4. **Download Button**
   * Confirm the .docx file downloads as expected for each document.
5. **API Rate Limits**
   * If using GPT-4o free-tier or capped key, observe whether repeated uploads trigger throttling.

The next enhancement steps

* Add a **comparison table**
* Merge multiple summaries into a single .docx
* Add Excel export of key findings
* Toggle between **Single vs Multi modes** inside one clean layout