# 📋 PRD – Product Requirements Document

## 1. Visão Geral

O sistema analisa arquivos enviados (PDF ou imagem) e identifica se o conteúdo é um artigo científico. Caso seja, alerta que o documento não é válido. Se o documento for válido, realiza análise textual e gera resumo com base em regras de conformidade.

## 2. Objetivos

- Evitar upload de artigos científicos indevidos.  
- Extrair e analisar textos de diferentes formatos.  
- Automatizar a validação textual e gerar resumo com IA.

## 3. Requisitos Funcionais

|  |  |  |
| --- | --- | --- |
| ID | Requisito | Descrição |
| RF01 | Upload de arquivo | Permitir envio de PDF ou imagem. |
| RF02 | Detecção de artigo | Identificar se o texto é um artigo científico. |
| RF03 | Extração de texto | Converter conteúdo para texto (OCR + PDF). |
| RF04 | Segmentação | Separar o texto em parágrafos. |
| RF05 | Contagem | Contar número de parágrafos e palavras. |
| RF06 | Análise de frequência | Identificar palavras mais frequentes. |
| RF07 | Resumo com LLM | Gerar resumo e verificar conformidade (>2000 palavras e >4 parágrafos). |

## 4. Requisitos Não Funcionais

|  |  |  |
| --- | --- | --- |
| ID | Categoria | Descrição |
| RNF01 | Desempenho | Processar em até 10s documentos até 5 MB. |
| RNF02 | Segurança | Excluir arquivos temporários após análise. |
| RNF03 | Precisão | Acurácia ≥ 90% na detecção de artigos. |
| RNF04 | Modularidade | Etapas independentes: OCR, segmentação, análise, resumo. |

## 5. Critérios de Aceitação

- Detecta artigo científico corretamente.  
- Extrai e analisa texto de PDFs e imagens.  
- Gera resumo apenas com uso do LLM.  
- Exibe conformidade com as regras (>2000 palavras e >4 parágrafos).