**什么是WebBaseLoader 大标题**

* + WebBaseLoader 是 LangChain 中用于抓取 **静态网页内容** 的文档加载器。
  + 通过 HTTP 请求直接获取网页 HTML，并提取其中的文本内容（自动清理标签、脚本等非文本元素）
  + 生成包含网页文本和元数据的 Document 对象
  + 适用于新闻文章、博客、文档页面等静态内容的快速提取。
  + 场景
    - 知识库构建（知识问答、企业知识库）、舆情监控（新闻/社交媒体分析）
    - 竞品分析（产品功能/价格监控）、SEO 内容聚合