

What is Docling ?

Docling simplifies document processing, parsing diverse formats — including advanced PDF understanding — and providing seamless integrations with the gen AI ecosystem.

Features

- 📁 Parsing of multiple document formats incl. PDF, DOCX, PPTX, XLSX, HTML, WAV, MP3, WebVTT, images (PNG, TIFF, JPEG, ...), LaTeX, plain text, and more
- 📄 Advanced PDF understanding incl. page layout, reading order, table structure, code, formulas, image classification, and more
- 🧬 Unified, expressive DoclingDocument representation format
- 🔄 Various export formats and options, including Markdown, HTML, WebVTT, DocTags and lossless JSON
- 📋 Support of several application-specific XML schemas incl. USPTO patents, JATS articles, and XBRL financial reports.
- 🔒 Local execution capabilities for sensitive data and air-gapped environments
- 🧩 Plug-and-play integrations incl. LangChain, LlamaIndex, Crew AI & Haystack for agentic AI
- 🔍 Extensive OCR support for scanned PDFs and images
- 🕶️ Support of several Visual Language Models (GraniteDocling)
- 🎙️ Audio support with Automatic Speech Recognition (ASR) models
- 🔌 Connect to any agent using the MCP server
- 🖥️ Simple and convenient CLI

What's new

- 📄 Structured information extraction [ beta]
- 📄 New layout model (Heron) by default, for faster PDF parsing
- 🔌 MCP server for agentic applications
- 📋 Parsing of XBRL (eXtensible Business Reporting Language) documents for financial reports
- 💬 Parsing of WebVTT (Web Video Text Tracks) files and export to WebVTT format
- 💬 Parsing of LaTeX files
- 📄 Parsing of plain-text files (.txt, .text) and Markdown supersets (.qmd, .Rmd)
- 📄 Chart understanding (Barchart, Piechart, LinePlot): converting them into tables, code or adding detailed descriptions Coming soon
- 📄 Metadata extraction, including title, authors, references & language
- 📄 Complex chemistry understanding (Molecular structures)