

Text extractor

- **Easy-to-use web service** - REST-like API with JSON exchange format.
- **Pattern and rotation recognition** – finds paper outlines.
- **Optimised** - can be trained specifically for your document set increasing recognition accuracy.
- **Advanced pdf extraction** – using specific pdf extractor with internal language analysis, specially supporting accented characters.
- **Linguistic analysis** - includes language recognition and additional linguistic analysers with optional spellchecker.
- **Full access** - can be deployed at your company site and/or integrated into your application and services.
- **Advanced usage** - can return individual word positions with position templating subsystem.

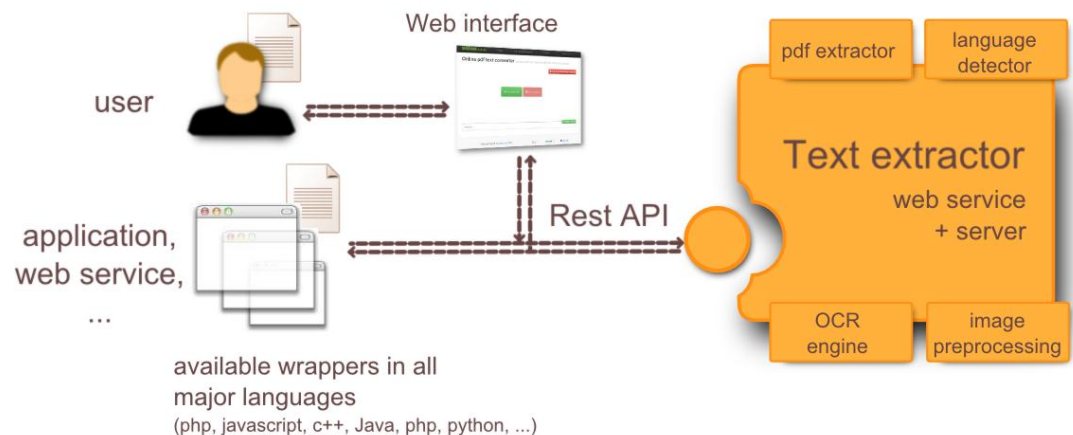
Intelligent Text Extractor Server with OCR¹

Mazoea s.r.o.

Our service helps you extract searchable text from PDF files, scanned documents and digital images. Text extractor supports many languages. We can train and adjust the server specifically for your document set.

Text extractor server

Easy-to-integrate and easy-to-use text extractor service converts your electronic documents into machine readable and searchable text. Advanced pdf extractor, optical character recognition together with linguistic tools improves the result text quality and accuracy.



¹ **Optical character recognition (OCR)** is the electronic conversion of text in images into machine-readable text.

Innovative
Effective
Reliable



Mazoea s.r.o.

<http://www.mazoea.eu>
textextractor@mazoea.eu