

Document Concept Extraction

The Document Lens allows fully automated concept recognition from Document data, stored as fully W3C compatible semantic web data.

KEY FEATURES

- Supports:
 - PDF
 - Microsoft Word™
- Customisable Taxonomies.
- Supports all common Knowledge Graphs
- Automated or Manual
- Provenance as standard
- Fully W3C Semantic Web compliant
- Fast
- Flexible configuration
- Platform Agnostic
- Highly Scalable
- Lightweight
- Robust

Knowledge Graph Support

Inbuilt support for all the common Knowledge Graphs, and generic SPARQL support for the less common.

Speed

xxx

Configurable Taxonomies

xxx

W3C Compliance

The outputted data is fully W3C semantic web compliant and can be used with any other compliant tooling in any way that you wish

Provenance as standard

All concepts extracted by the Document File Lens come with full provenance as standard, your data is fully traceable to source.

Triggering

Triggering may be automated via a message queue, or manually triggered via its RESTful endpoint.

Platform Agnostic

The Document Lens will run on any container infrastructure that supports Docker. This can be in the Cloud or on Bare Metal

Highly Scalable

The Document File Lens can support any size of source file. The Lens itself may also be scaled horizontally to fit your particular requirements

Lightweight

xxx

Robust

Build from the ground up to be enterprise capable

