Convert your unstructured PDF content into a valuable data source

From PDF content to a valuable data source – with our Semantic Analyze Workflow

SCAS
Smart Content
Automation Services

Most of the data in our digital world is not structured enough – if at all – for digital transformation processes, e.g. automated text generation in ecommerce.

Our Semantic Analyze Workflow offers you a powerful AI supported tool to extract, analyze and structure PDF content into any data format required.

Our solution operates beyond simple OCR. The Semantic Analyze Workflow scans even complex structured PDF content, identifies the visual layout and classifies single modules.

With an embedded grammar parser you can align, unify and correct your data on the basis of multiple PDF documents. The analyzed data can then be written into any database via API or can be exported in any format required (PDF, xml, HTML, XHTML, xlsx).

Save time, resources and money while getting not only structured data, but, for the first time, corrected and semantically enriched data.

The Semantic Analyze Workflow is part of our Smart Content Automation Services.

Key Facts

- PDF content is **automatically converted into structured data**
- The self-learning system lets you **benefit from every further analyze workflow**
- **Highly scalable** for large numbers of PDFs
- All export formats possible: **PDF, xml, HTML, XHTML, xlsx** etc.
- Use of native Windows application relies on **standard software technology**

Please call us for a free trial

text2net GmbH

Dr. Horst Wigger
CEO
Kaiserstraße 16

53113 Bonn

Phone: +49 (0)228 928 900 14
Mail: horst.wigger@text2net.de