

ABBY FlexiCapture Engine 12

Das umfangreiche SDK für Data Capture

Was ist FlexiCapture Engine?

ABBY FlexiCapture Engine ist ein umfassendes Software Development Kit für Entwickler, die automatisierte Datenerfassung, Formularverarbeitung und intelligente Dokumentenklassifizierung in ihre Applikationen integrieren möchten. Durch die Integration des SDK über eine API können Daten aus eingescannten Dokumenten extrahiert werden, um von Programmen und IT-Systemen genutzt zu werden. Es lassen sich sowohl Standardformulare wie Fragebögen, als auch Dokumente mit variierender Struktur wie z.B. Lieferscheine, Rechnungen oder Verträge automatisch verarbeiten. Eingehende Dokumente können auch automatisch verschiedenen Dokumententypen zugeordnet werden, um Prozesse zu beschleunigen. Somit erhalten eigene IT-Anwendungen einen zusätzlichen Mehrwert.

Erweiterung Ihrer bestehenden Applikation oder Entwicklung automatisierter Data Capture Lösungen

FlexiCapture Engine liefert eine große Bandbreite neuester Technologien für Datenextrahierung, Datenklassifizierung und Indizierung. Bereits entwickelte Lösungen können schnell um den Datenexport zu Backend-Systemen oder Dokumentenkonvertierung erweitert werden. Mit FlexiCapture Engine erhalten Entwickler alles was sie brauchen aus einer Hand: Tools zur Entwicklung, Training aber auch Support und Wartung.

Intelligente Dokumenten-Verarbeitung

Automatisierte Datenextrahierung aus Formularen, Verarbeitung von semi- und unstrukturierten Dokumenten und Dokumentenklassifikation in einem einzigen SDK.

Verbesserte Verarbeitung von mehrseitigen Dokumenten

Einzigartige Technologie zur vereinfachten Verarbeitung von mehrseitigen Dokumenten sowie Tabellen, die sich über mehr als eine Seite erstrecken.

Zuverlässige Verarbeitungsplattform

Das leistungsstarke SDK entspricht Data Capture Industriestandards zur Datenverarbeitung in Unternehmensprojekten.

Automatisierte Datenextraktion in Backend-Applikationen

Einfache Integration von Daten-Export-Funktionalitäten in Backend-Applikationen anhand von Standardschnittstellen.

Verarbeitung mehrsprachiger Dokumente

Unterstützung verschiedenster Sprachen zur Entwicklung von lokalisierten Anwendungen im internationalen Bereich.

Auto-Learning Funktionen

Die Auto-Learning API und ein Beispiel im Quelltext erlauben es, einfache Dokumente schnell über eine eigene Point-&-Click Oberfläche zu trainieren, um dann aus ähnlichen Dokumenten die Daten auszulesen.

Kosteneinsparungen

Reduktion der Kosten bei der Datenextraktion durch Integration einer automatisierten Verarbeitung.

Effiziente Entwicklung

Standard-Komponenten ermöglichen eine schnelle Integration des SDK in existierende Applikationen und verkürzen dadurch die Entwicklungszeit.

Qualitativ hochwertige Datenextraktion

Erstklassige OCR-Erkennungstechnologien in Kombination mit regelbasierten und automatisierten Verifizierungsmethoden garantieren eine optimierte und valide Datenausgabe.

Nahtlose Integration in Applikationen

FlexiCapture Engine garantiert Entwicklern eine nahtlose Integration von Data Capture Technologien in ihre eigenen Applikationen.

FUNKTIONEN

- Integration intelligenter Data Capture Funktionalitäten zur Verarbeitung von Dokumenten
- Dokumentenklassifizierung und automatische Dokumententrennung
- Leistungsstarke Datenextrahierung für semi- und unstrukturierte Dokumente
- Verifizierungs- und Überwachungsoptionen für eine optimale Datenausgabe
- Qualitativ hochwertige Erkennungstechnologien für OCR, ICR, OMR, 1D und 2D Barcodes
- Sprachunterstützung für bis zu 205 OCR und 126 ICR Sprachen
- Automatischer Datenexport in Backend-Applikationen
- PDF, MRC PDF & PDF/A Dokumentenexport zur Archivierung

VORTEILE FÜR ENTWICKLER

- Erweitern Sie Ihre Applikationen um Data Capture Komponenten
- Einzigartige und leistungsstarke Features sowie API für Datenextraktion
- Gebrauchsfertige SDK-Komponenten zur Beschleunigung des Entwicklungsprozesses
- Vollständige Kontrolle über Dokumentenverarbeitung, extrahierte Daten und das Routing von Dokumenten
- Kompatibel mit ABBYY FineReader Engine 12 API, um weitere Funktionen zur Dokumentenkonvertierung zu integrieren

Übersicht über den Entwicklungsprozess

Integration, Entwicklung von Dokumentbeschreibungen und Verarbeitung

Die Integration von Data Capture Funktionalitäten mit FlexiCapture Engine ist ein zweistufiger Prozess. Zuerst wird das SDK in die eigene Applikation integriert. Dazu können Entwickler die bereitgestellten FlexiCapture SDK Komponenten verwenden (Stufe 1). Für die spätere Verarbeitung unterschiedlicher Dokumententypen in einem Projekt müssen Dokumentenvorlagen - sogenannte Templates – entwickelt werden. Für Dokumente mit variierendem Layout müssen mit Hilfe von speziellen Tools Dokumentbeschreibungen definiert werden (Stufe 2). Das Entwicklungskit enthält alle nötigen Design-Tools: einen Document-Template-Editor für strukturierte Formulare, FlexiLayout Studio für die Beschreibung unstrukturierter oder semi-strukturierter Dokumente sowie eine API die es erlaubt einfache, einseitige FlexiLayouts über eine eigene Benutzeroberfläche zu erzeugen.

Applikationsintegration und Dokumentenverarbeitung

FlexiCapture Engine hat eine einfach strukturierte und leicht zu verwendende API. Die Entwicklung kann in C#, C++, Visual Basic .Net, Delphi, Java und anderen COM-kompatiblen Tools erfolgen. Die folgenden Schritte zeigen, wie Sie Ihre Applikationen um Data Capture Funktionalitäten erweitern können:

Schritt 1 – Integration:

Integrieren Sie die FlexiCapture Engine DLLs in Ihre Applikationen, Workflows oder Geschäftsanwendungen.

Schritt 2 – Projektauswahl:

Laden Sie ein vordefiniertes Set von Dokumentvorlagen, die mit der neuen Auto-Learning API erstellt oder mit FlexiLayout Studio entwickelt und getestet wurden.

Schritt 3 – Dokumentverarbeitung:

Laden Sie die Dokumente und PDFs, die verarbeitet werden sollen. Dokumententrennung, Klassifizierung und Erkennung werden automatisch ausgeführt, basierend auf der Logik der FlexiCapture Dokumentendefinitionen.

Schritt 4 – Datenverifizierung:

Die API für die Datenverifizierung von FlexiCapture Engine erlaubt Ihnen den Zugriff auf die internen Erkennungsergebnisse. Die Visual Components, erlauben eine schnelle Implementierung der manuellen Verifizierung.

Schritt 5 – Verarbeitung von Dokumenten und Daten:

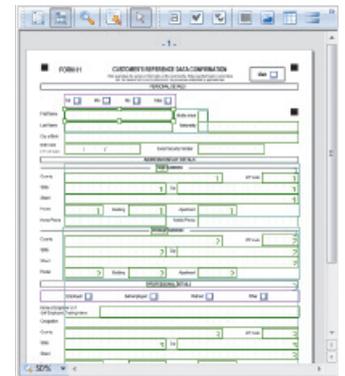
Alle extrahierten Daten können nun in Ihrer Applikation verarbeitet, an andere Applikationen geschickt oder in Form von XML-, XLS- und CSV-Dateien exportiert werden. Die verschiedenen Dokumente können sowohl als Bild-dateien als auch als durchsuchbare PDFs oder PDF/A-Dateien exportiert werden.

Entwicklung von Dokumenten-Templates für Formulare

Mit dem Document-Template-Editor können schnell und intuitiv Dokumenten-Templates zur Verarbeitung von Formularen erstellt werden.

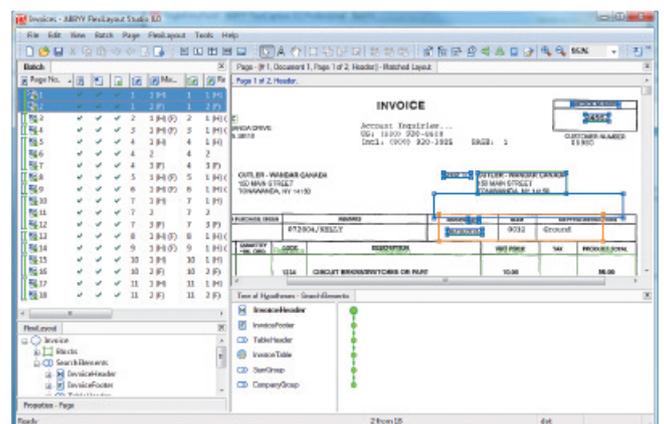
- 1) Verschiedene Segmente eines mehrseitigen, strukturierten Formulars werden in den Editor geladen.
- 2) Dann erfolgt die Definition der Elemente, die zum „Template Matching“ verwendet werden, z. B. Anker (Quadrate, Kreuze), statischer Text oder Linien.

- 3) Anschließend werden die Erkennungsbereiche mit Hilfe eines graphischen Editors festgelegt, z. B. Textblöcke, Tabellen, Kontrollkästchen, Barcodes oder Bildbereiche.
- 4) Zuletzt werden die Erkennungsoptionen festgelegt, also die Definitionen ob der Inhalt mit OCR oder ICR gelesen wird und welche Datentypen, Wörterbücher und Verifizierungsregeln angewendet werden sollen.



Entwicklung von mehrseitigen FlexiLayout-Templates für Dokumente mit unterschiedlichem Layout

- 1) Zunächst wird eine Auswahl repräsentativer Dokumente mit unterschiedlichem Layout ins FlexiLayout Studio geladen.
- 2) Spezielle Elemente zur Identifikation und Orientierung innerhalb eines Dokumentes werden bestimmt, z. B. Texte, Barcodes, Linien oder auch Abstände zwischen Elementen.
- 3) Im letzten Schritt erfolgt die Definition der Elemente, die die gesuchten Daten aufnehmen sollen. Dies können Texte unterschiedlicher Länge, Zahlen, ein Datum oder ganze Tabellen sein. Zusätzlich werden diese Elemente in geometrischen Bezug zu den unter 2.) definierten Ankern gesetzt, z. B. „Lies den Text rechts neben „Rechnungsnummer“, der zwischen 8-16 Zeichen lang ist und Zahlen enthält“.



Funktionen im Überblick

ABBYY FlexiCapture Technologie

Einzigartige Technologien für OCR, ICR, OMR und Barcode-Erkennung

Die Kerntechnologien von FlexiCapture Engine verarbeiten bis zu 205 Sprachen bei OCR und 126 Sprachen bei ICR, und bieten zudem die Erkennung von Ankreuzfeldern und Barcodes.

Verschiedene Dokumentarten auf einmal verarbeiten

FlexiCapture Engine bietet fertige Komponenten, die es Entwicklern ermöglichen, die automatische Extrahierung von Daten aus einem Stapel von Formularen und Dokumenten in eigene Anwendungen zu integrieren. Die automatische Erkennung der unterschiedlichen Dokumententypen ermöglicht die Verarbeitung von verschiedenen Dokumenten (strukturiert, semi-strukturiert, unstrukturiert) mit Hilfe von Formularvorlagen und FlexiLayouts.

Einfache und erweiterte Dokumententrennung

FlexiCapture Engine bietet zwei Arten von automatischen Dokumententrennung. Die einfache Trennung basiert auf Leerseiten und Barcodes, die erweiterte auf intelligenten Dokumentendefinitionen. Die Anwendung der Templates auf einen Dokumentenstapel kann ohne manuelles Eingreifen ausgeführt werden.

Dokumentenklassifizierung

Sobald ein Dokument zu einem Formulartemplate oder einem FlexiLayout passt, wird automatisch ein Dokumenttyp klassifiziert, z. B. Kundenumfrage, Rechnung, Vertrag.

Verarbeitung von Formularen mit fester Struktur

FlexiCapture Engine ist perfekt dafür geeignet, Stapel von Standardformularen zu verarbeiten, beispielsweise Fragebögen, Studien oder Coupons.

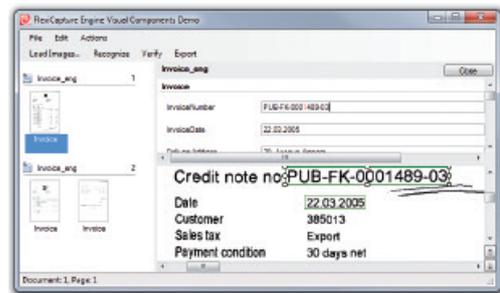
Verarbeitung von mehrseitigen Dokumenten mit verschiedenen Layouts

Geschäftsdokumente bestehen oft aus mehreren Seiten. Mit FlexiCapture Engine bekommen Entwickler eine einzigartige Technologie an die Hand, die es ihnen ermöglicht, auch mehrseitige Dokumente mit Tabellen einfach zu verarbeiten.

Datenkontrolle und Verifizierung

FlexiCapture Engine bietet effektive Möglichkeiten, die erfassten Daten zu überprüfen und Fehler aufzudecken:

- Automatisch: Daten können bei Bedarf automatisch überprüft werden, das umfasst eine einfache Bereichsprüfung und den Abgleich bekannter Werte aus einer Datenbank. Zudem können durch Skripte eigene Regeln festgelegt werden.
- Durch den Benutzer: Die API von FlexiCapture Engine erlaubt die Entwicklung eigener Benutzeroberflächen. Mit den Visual Components (Source Code Samples) können Entwickler den Anwendern eine effektive, ergonomische Benutzeroberfläche bereitstellen.



Daten- und Dokumentenexport

ABBYY FlexiCapture Engine ist in der Lage, gewünschte Daten aus einem Dokument auszulesen und es dann anhand dieser zu indexieren. Für Archivierungszwecke können zudem Dokumente in hochkomprimierte (MRC) PDFs oder PDF/A-Dokumente konvertiert werden.

Java Wrapper

Das FlexiCapture Engine enthält auch Java-Klassen, so kann die Funktionalität des SDKs auch ohne Mehraufwand aus Java-Anwendungen genutzt werden.

Auto-Klassifizierung

FlexiLayout Studio unterstützt Klassifizierungstechniken, die Raster und Inhaltsklassifizierung intelligent miteinander verbinden. Die Technologie ist einfach einzurichten und garantiert eine schnelle Verarbeitung. Dokumenten-Training ist auch über die API möglich.

Dokumententemplates

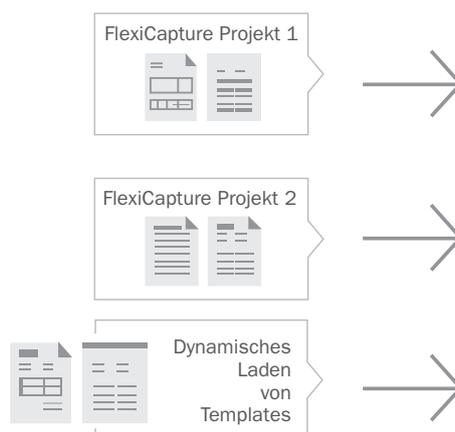
Templates für statische Formulare:
Fragebögen, Coupons, Formulare, ...



Templates für flexible Layouts:
Rechnungen, Verträge, Briefe, ...



Sammlung von Templates



Applikation



Spezifikationen und Lizenzierung

SPEZIFIKATIONEN

Systemanforderungen

- PC mit x86-kompatiblen Prozessor (1 GHz oder höher).
- Betriebssystem:
32-Bit: Windows 8, 8.1, Windows 7,
64-Bit: Microsoft Windows Server 2012, 2012 R2, 2008 R2, Windows 10, Windows 8, 8.1, Windows 7
- Speicher: mind. 1 GB RAM

Cloud Umgebungen

- Microsoft Azure
- Amazon EC2

Eingabeformate

BMP, PCX, DCX, JPEG, JPEG 2000, JBIG2, PNG, GIF, TIFF, DjVu, gescannte und am PC erstellte PDFs

Ausgabeformate

Editierbare Formate

- XLS, DBF, CSV, TXT, XML

Durchsuchbare Formate

- PDF und PDF/A mit MRC-Komprimierung, tagged PDF

Exportformate für Bilder

- BMP, PCX, DCX, JPEG 2000, JPEG, TIFF, PNG, PDF, PDF/A

Entwicklung

FlexiCapture Engine API unterstützt den COM-Standard und kann einfach mit C#, C++ (raw und COM), Visual Basic .Net, Delphi, Java und anderen Tools implementiert werden. Das SDK enthält Codebeispiele in C#, C++ (raw und COM), VB.Net, Java, Delphi und Scriptingsprachen wie VBS, JS und Perl.

Weitere Informationen online:

<https://www.abby.com/data-capture-sdk/technical-specifications/>

ABBYYs Lizenzpolitik

ABBYY FlexiCapture Engine wird über eine flexible, modulare Lizenzpolitik vertrieben, die es Entwicklern ermöglicht, die beste Kombination von Tools und Preisoptionen für ihr Projekt auszuwählen. Die Lizenzierung wird angeboten als:

Entwicklerlizenzen

Diese erlauben es, FlexiCapture Engine in eigene Applikationen zu integrieren und zu testen. Die Lizenz beinhaltet auch die Applikation FlexiCapture Standalone mit dem Tool FlexiCapture Studio. Diese ermöglicht die Entwicklung von FlexiLayouts für Dokumente mit verschiedenem Layout sowie von Templates für Standardformulare.

Runtime Lizenzen

Erlauben den Vertrieb von Applikationen, die FlexiCapture Engine Funktionalitäten beinhalten. Runtime Lizenzen unterscheiden sich in ihrer Funktionalität, in der Anzahl an Seiten die pro Jahr verarbeitet werden können, oder der Anzahl an CPU Cores (ohne Seitenzähler). ABBYY Europe liefert standardmäßig Netzwerklicenzen. Diese werden zentral über einen im LAN installierten Lizenzmanager bereitgestellt, laufende Instanzen, egal ob virtuell oder physikalisch, können dann die Lizenz nutzen.

Zusätzliche Module für Runtime Lizenzen

Runtime Lizenzen können durch das Hinzufügen von spezifischen Funktionalitäten, wie z. B. weiterer OCR Sprachen oder der PDF-Export-Funktion, erweitert werden. Auch die ABBYY FineReader Engine API für Volltext-OCR und Dokumentumwandlung kann mit FlexiCapture Engine lizenziert werden.*

Software Maintenance, Training und Professional Services

Um den Erfolg Ihrer Produkte zu sichern, bietet ABBYY zusätzliche Unterstützung, Trainings und Schulungsprogramme für alle Produkte. Wenn Sie professionelle Unterstützung möchten, kontaktieren Sie bitte ABBYY, um Details zu Professional Services zu erhalten.

* FineReader Engine API ist im Lizenzmodell von ABBYY Europe enthalten, dies gilt nicht für andere Vertriebsgebiete.

ABBYY FlexiCapture Engine Dokumenten-Verarbeitung

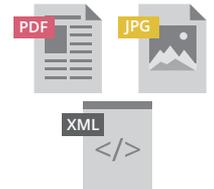
Dokumentenstapel aus verschiedenen Quellen



Eigene Applikation mit FlexiCapture Engine für Desktop, Server oder Cloud



Extrahierte Daten & Multi-Format-Dokumentenexport



Extrahierte Daten in einer Datenbank



ABBYY®

Headquarter Europa
ABBYY Europe GmbH
Landsberger Str. 300, 80687 München, Deutschland
Tel: +49 89 69 33 33 0
sales_eu@abby.com
www.ABBYY.com