

# PDF Xpansion SDK 11

## User Guide

# LEGAL NOTICE

© 2002-2015 soft Xpansion GmbH & Co. KG

Koenigsallee 45

D-44798 Bochum

<http://www.soft-xpansion.com>

PDF Xpansion © 2002-2015 soft Xpansion GmbH & Co. KG

All rights reserved.

All information, technical data and programs in this file or book as well as on the included CD were compiled with the greatest care by the authors. The publisher cannot provide a guarantee nor be held legally responsible for the effects of any faulty information imparted.

All the software and hardware descriptions and trademark names of the respective companies are protected by trademark, copyright and patent rights.

Although the text was created with diligence, the publisher and the authors do not take further responsibility for the completeness and accuracy of the content of the text, nor, for that matter, do they guarantee that following the recommendations of the text will lead to successful results for the user. Use of the text and the software is exclusively at the user's own risk.

# TABLE OF CONTENTS

- PORTABLE DOCUMENT FORMAT (PDF)..... 3
- ABOUT THE PDF SDK XPANSION GUIDE ..... 4
- THE PDF XPANSION SDK..... 5
  - Overview ..... 5
    - General Structure of the PDF Xpansion SDK..... 6
    - Experience ..... 8
    - New Generation..... 8
    - Try it? ..... 10
    - Buy it! ..... 10
  - SDK Features..... 10
    - PDF File/Document ..... 10
    - XPS File/Package/Document ..... 14
    - Viewer ..... 15
    - Global Settings and Object Factory ..... 16
  - SDK Licensing Model – Pay Only Once!..... 16
    - License Conditions..... 17
    - License Types ..... 18
    - License Agreement..... 18
    - Prices..... 19
- INTEGRATION WITH TEXT RECOGNITION FUNCTIONALITY (OCR)..... 24
- CONTACT US..... 25

# PORTABLE DOCUMENT FORMAT (PDF)

A PDF document is a file saved in the Portable Document Format (PDF). The resulting file has the extension “.pdf”. Adobe Systems Inc. proposed this format in 1993 as an open standard for document exchange. It enables consistent and uniform display of documents independently of the hardware or software used to create or view them. Each PDF file encapsulates one document with a complete description of the page content according to a fixed layout, including the text, fonts, graphics and other information needed to display it.

From 1993 to 2006 Adobe Systems actively developed the PDF format and released 7 new versions, up to version 1.7. The PDF format has spread widely through very different areas of electronic document management around the world. The PDF format is now the de facto standard in various application areas.

At the start of 2007, Adobe Systems submitted an application to the International Organization for Standardization (ISO) to register the PDF format as an ISO standard. The ISO published the approved ISO 32000-1 standard based on PDF version 1.7 in July 2008. PDF thus became the de jure standard as well.

A PDF file contains document data in a binary format. A special program is required to read a PDF file – a PDF reader. You also need a special program to save a PDF file – a PDF creator (writer). To process a PDF file, you need a PDF editor program. The PDF Xpansion SDK is a good platform for creating any of these applications or an all-in-one solution. PDF Xpansion’s features grant you the power and flexibility to handle any task involving PDF files.

The structure of a PDF document contains of a lot of interrelated objects: pages, annotations, images, fonts, etc. If you use the PDF Xpansion SDK to process PDF documents, you do not need to know or consider the peculiarities of a PDF document: the PDF Xpansion SDK ensures its internal integrity and correctness. At the same time, professional developers gain a wide range of powerful options for processing PDF documents.

# ABOUT THE PDF SDK XPANSION GUIDE

This document is intended to provide a brief introduction to the PDF Xpansion SDK:

- Overview
- Functionality and properties
- Licensing
- Prices

It is not a programming guide; it contains a general overview of the SDK and a guide to the licensing models and prices. All three types of license are described; the prices apply to the Business license, which is the most frequently used model.

For technical questions please refer to the following resources:

- PDF Xpansion SDK Reference – part of the SDK documentation
- Technical support via [pdf@soft-xpansion.com](mailto:pdf@soft-xpansion.com)

# THE PDF XPANSION SDK

## Overview

The PDF Xpansion SDK provides a set of APIs that you can use to create high-performance PDF/XPS-enabled applications and/or software solutions.

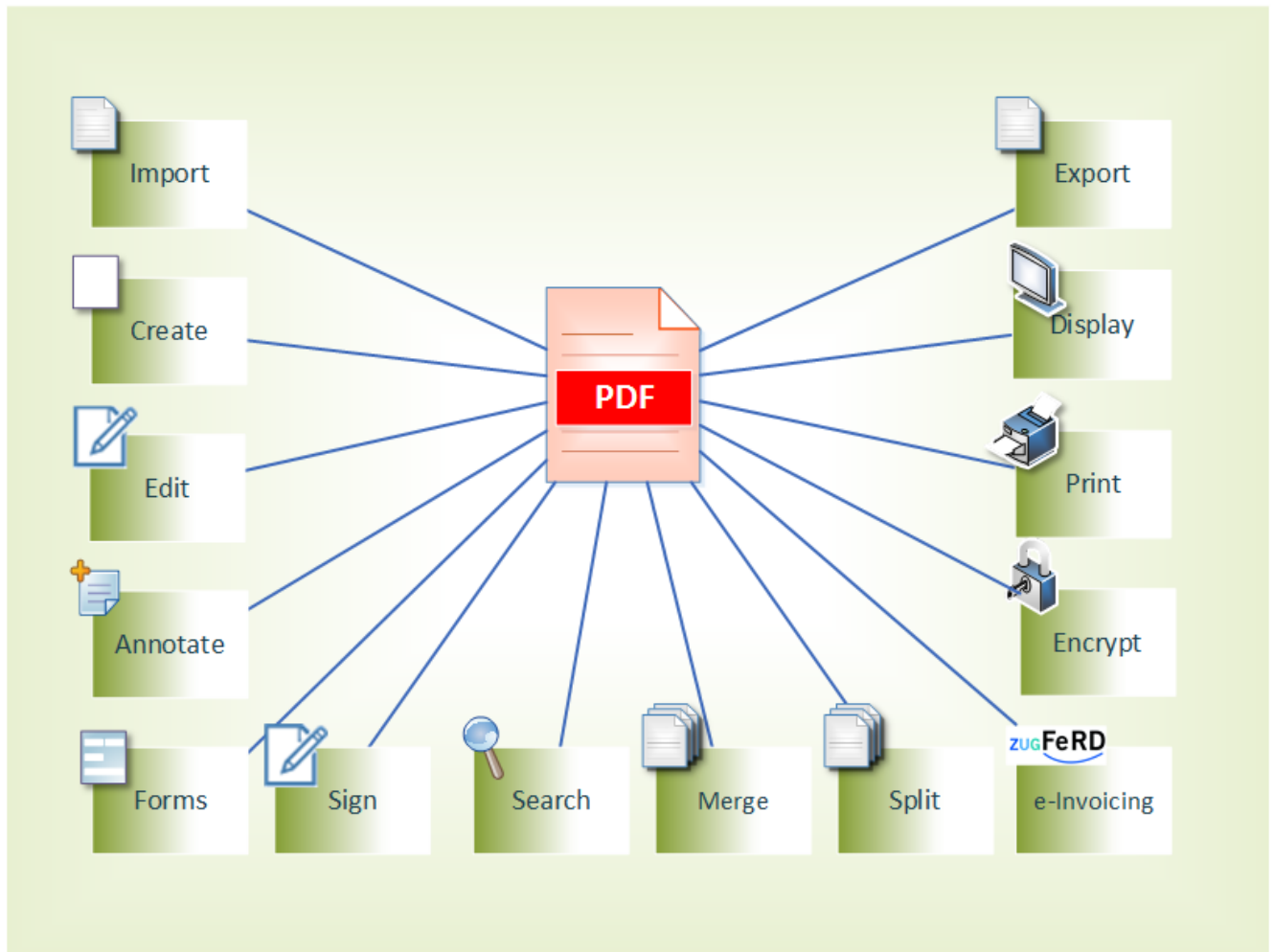
The PDF Xpansion technology grants you wide-ranging options for creating and editing PDF or XPS documents, displaying and printing them, converting between these two formats, exporting and/or importing to/from other file formats, and much more.

You'll find a detailed feature list in the [SDK Features](#) chapter.

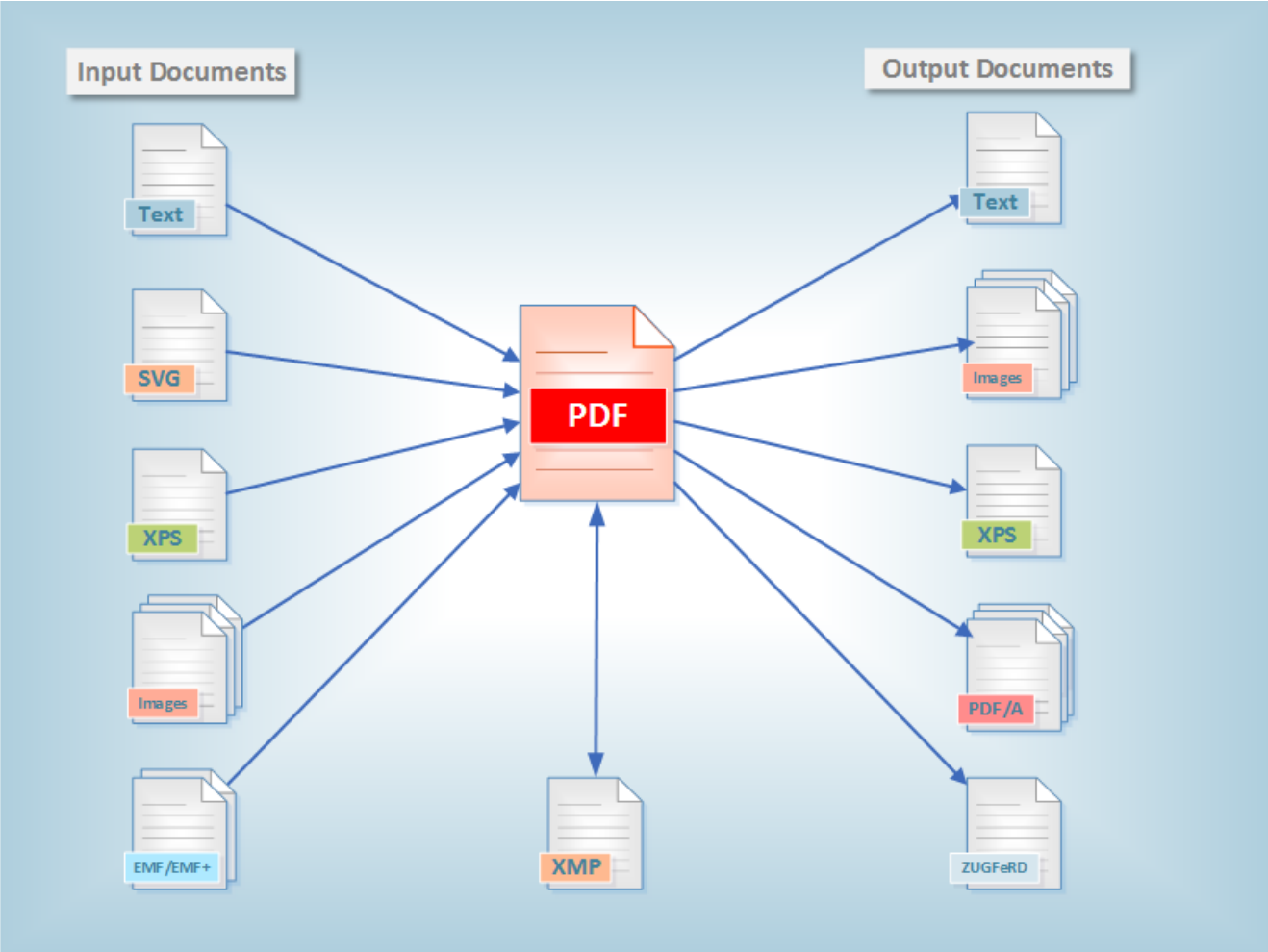
You'll find a description of the PDF Xpansion SDK API and all the required information about programming resources in the [PDF Xpansion SDK Reference](#) included in the SDK delivery package.

# General Structure of the PDF Xpansion SDK

## General Functionality

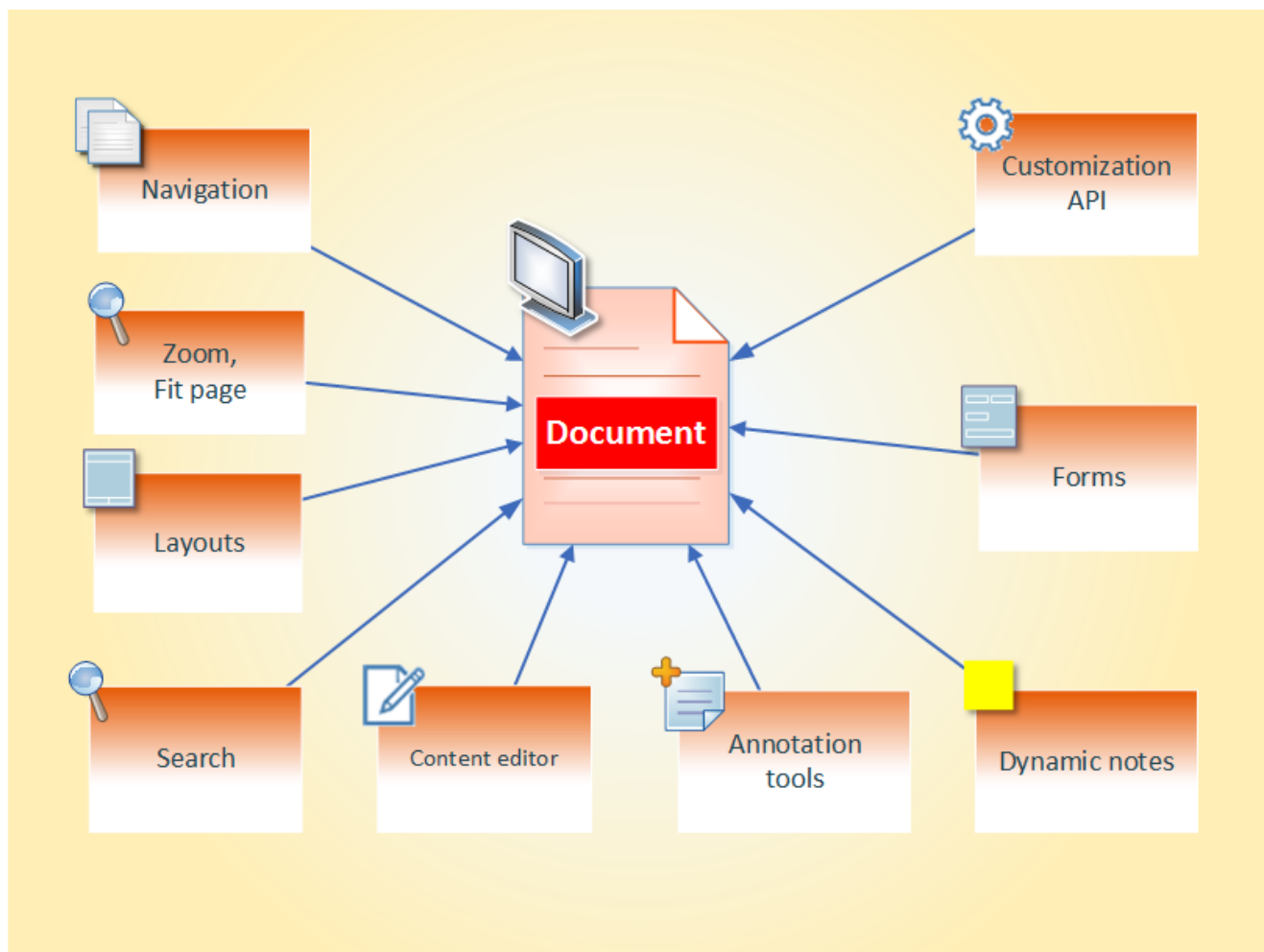


# Document Formats





## Document Viewer



## Experience

Soft Xpansion has been developing its PDF Xpansion technology for more than ten years. It draws upon our twenty years of software development experience and our deep knowledge of the PDF document format and document management techniques.

Our clients have integrated the six previous generations of the PDF Xpansion SDK into hundreds of applications, products and solutions. Using the SDK as a base, we have developed and provide:

- A wide range of consumer products (Perfect PDF, Perfect Print, etc.),
- A high-performance server solution (PDF Render Center) and
- Many customized DMS solutions.

Please consult our [Competence Center](#) for more info.

## New Generation

The new generation of PDF Xpansion technology is **PDF Xpansion SDK 11**.

## Development Platforms

This generation of the SDK provides four sets of APIs with similar structures and functionalities:

- **Windows** – applications and services written in C++ like the classic Windows x86/x64 programs
- **DotNET** – .NET Framework applications
- **UWP** – applications for all Windows 10 devices
- **WinStore/WinPhone** – Windows Store/Windows Phone Store applications for Windows 8.1

Windows Store/Windows Phone Store applications compatible with Windows 8.0 are supported by PDF Xpansion SDK 10, the previous version of the SDK. Please contact us if you require this functionality.

For Delphi and/or C++ Builder developers we provide a wrapper library with a COM interface. This wrapper enables the use of most of the PDF Xpansion SDK's functionalities in applications written in any programming language that supports COM calls and interfaces. Please contact us if you need this wrapper.

## Operating System Platforms

The PDF Xpansion SDK can be used with the following Windows versions:

1) Without any functional restrictions and third party components:

- Windows 10 and Windows Server 2016
- Windows 8.1, Windows Server 2012 R2 and Windows Phone 8.1
- For C++ and .NET applications: Windows 8, Windows Server 2012

2) With restrictions and/or third party components:

- Windows 8, Windows Server 2012 and Windows Phone 8:
  - Not possible for Windows Store and Windows Phone applications, please contact us for PDF Xpansion SDK 10
- Windows 7 and Windows Server 2008 R2:
  - For C++ and .NET applications: the [platform update](#) must be installed
- Windows Vista, Windows XP, Windows Server 2008
  - For C++ and .NET applications: very restricted use is possible, please contact us for details

You will find more information about redistributing the SDK in the [Redistribution of PDF Xpansion](#) chapter of the [PDF Xpansion SDK Reference](#).

## Hardware Platforms

The PDF Xpansion SDK contains a set of libraries for the x86 and x64 platforms as well as for the ARM processor architecture.

## Multithreading

The PDF Xpansion SDK libraries are safe for multithreaded operations with multiple document objects. You can process several documents in multiple threads without any synchronization. The interface of document object and child objects cannot be considered to be thread-safe. In practice, this means that you cannot access a document and/or child objects from two or more threads without additional synchronization.

## Unicode and Internationalization

The PDF Xpansion SDK libraries are fully Unicode-enabled and provide data types for Unicode-only programming. The libraries do not contain any localizable data or messages. All groups of world languages, incl. Asian and RTL (right-to-left) languages are covered by the PDF Xpansion SDK. Please refer to the section on [OCR possibilities](#) for different languages as well.

## Try it?

If you would like to test the performance and functionality of the PDF Xpansion SDK, you can download the current version of the SDK with a trial license from our [website](#). This license allows you to use all of the SDK's functionality but applies a demo watermark to the document pages. Use the samples included in the SDK for a quick start or read the [Getting Started](#) chapter in the [PDF Xpansion SDK Reference](#).

Try our Perfect PDF desktop applications, like [Perfect PDF 9 Premium](#), to experience their powerful features and performance as well. You will also find our **PDF Xpansion SDK-based application** in the [Windows Store](#). Please note that the SDK version used in the applications may differ from the current one.

## Buy it!

The licensing model is simple, flexible and ensures our clients a reasonable price. You'll find more information about the licensing model and prices in the [SDK Licensing Model](#) chapter.

# SDK Features

## PDF File/Document

## Create PDF Document

There are many ways to create a new PDF document from different sources:

- 1) Import or convert a PDF document from:
  - The print output of any printable application sent to the virtual printer in EMF/EMF+ format
  - XPS or OXPS documents
  - Plain text
  - Raster images: BMP, TIF, JPG, PNG, GIF, etc.
  - Vector images: SVG
  - Combinations of the sources mentioned above
- 2) Create a new PDF page from an empty page and combine the sequence of pages and metadata in the new PDF document

## Open

The SDK loads a document from any file or stream that conforms to the PDF specification (versions 1.0 – 1.7), to ISO 32000 standards (versions 1 and 2) or to any PDF subset (PDF/A, PDF/X., etc.).

## Save

The SDK saves documents to files or streams conforming to ISO 32000 standards (versions 1 and 2).

## Print

The SDK enables the printing of PDF documents – individual pages, content, comments, forms, etc.

## View/Render Pages

The SDK provides a draw procedure for displaying pages or rendering pages as raster images. This procedure uses the Direct2D API. This functionality can be used as a basis for programming your own PDF viewer.

## Combine (Merge) Pages

The content of different pages can be combined into a new page. You can use different methods and layouts to combine the content into one page. Copying selected pages from one document to another is also supported.

## Manage Document Pages

The pages in a document can be added, removed, copied, moved within the document structure, resized and rotated without changing their content.

## Search in Text, Export Text

The SDK can search within a document's text. It is also possible to export plain text.

## Encryption (Document Security)

The SDK loads not only non-encrypted PDF files but also files encrypted using a password security handler and a public-key security handler. The algorithms RC4 and AES are supported, as are different key lengths that are declared in ISO 32000. Additionally, the SDK supports the encryption algorithm specified in ISO 32000-2 and used in Adobe Acrobat X or higher.

The SDK allows you to encrypt documents using these two security handlers and any combination of supported algorithms and key lengths. You can specify user permissions for the encrypted documents.

Support for custom encryption algorithms or security models is available upon request.

## Digital Signatures

The SDK provides an API for signing documents and validating signatures according to all the signature and digest formats declared in the PDF specification. The SDK also supports timestamping signatures.

## Form Fields

The SDK provides access to data from form fields. It is possible read or write field values. The SDK also allows you to modify forms: add, remove or edit form fields. It enables use of the FDF (Forms Data Format – a file format for representing form data and annotations that are contained in a PDF form) and XFDF (XML Forms Data Format) formats for saving and exchanging form field content.

## Dynamic Forms – XFA Format

The PDF Xpansion SDK provides a restricted usage of XFA forms (XML Forms Architecture). The following operations are available:

- Previewing forms in a PDF viewer (filling is not supported)
- Printing XFA forms
- API for access to form data (read and fill form content)

## Combine and Split Documents

Using the SDK you can copy pages from one document to another and combine or split new documents. The combine function allows you to copy not only pages but also to correctly copy links, bookmarks, fields, annotations, layers and other elements of PDF documents to a new document.

## Revisions (Incremental Updates)

If a document contains several document versions (revisions), the SDK loads the document in the state of the last version. You can switch to any other revision before you begin editing the document. You can save any changes to the currently loaded document as a new version (revision, incremental update).

## Document Properties and Metadata

The SDK provides an API for accessing different document properties and metadata.

Metadata in XML format (eXtensible Metadata Platform) is supported via the XMP API. All standard XMP schemes are available.

## Document Outline (Bookmarks), Links and Named Destinations

The SDK allows you to use and/or edit the table of contents within a document as well as hyperlinks and named destinations.

## Layers (Optional Content)

The SDK provides an API for using or managing document layers.

## Embedded Files and Portfolio

Insert and explore other files that are embedded in the PDF file. Support for portfolio functionality according to the PDF specification.

## Annotations

The SDK supports most common types of PDF annotations and design elements: text markup, forms, notes, textboxes, stamps, watermarks, pen, eraser, media and 3D annotations, embedded files as comments and replies in comments. It provides an API for managing annotation properties and allows you to generate appearances after changes in properties. It is possible to create and edit annotations and to display them in Viewer ([Viewer functionality](#)) as well as to create end user tools for working with annotations in your application.

## Rich Content Functionality

Rich Content is intended for creating PDF content from the very beginning and for editing content in existing PDF files. It provides close access to the page content of PDF and XPS documents. With Rich Content you can create content for new pages (fill the blank pages) and

---

edit the content of pages created by other applications. You can also explore the content of the pages in order to reveal the placement of content objects (text, images, vector elements) on the page, format the text, etc. The Rich Content component is required for enabling editing functionality in your own application.

If you intend to create your own PDF page content editor, you can use our Page Content Editor Tool module, which supplies not only the content editing functionality but also the required user handling.

## **Convert to PDF/A**

The SDK converts PDF documents to make them compliant with the PDF/A-1, PDF/A-2 and PDF/A-3 formats. A, B and U levels of conformity are supported.

## **ZUGFeRD API – Electronic Invoice Processing**

The ZUGFeRD (Zentraler User Guide des Forum elektronische Rechnung Deutschland) data format – a universal format for electronic invoice processing and interchange – is supported by the PDF Xpansion SDK. It requires PDF/A-3 and XMP metadata processing. All three ZUGFeRD profiles (Basic, Comfort and Extended) can be used in the PDF Xpansion SDK. Using the API it is possible to read invoices with access to the metadata, verify them and create new ones according to the ZUGFeRD specification.

## **Import from XPS**

The SDK enables the quick, high-quality conversion of XPS documents to PDF documents.

## **XPS File/Package/Document**

### **Open**

The SDK loads packages from any file or stream that conforms to the XPS specification.

### **Save**

The SDK saves packages to files or streams conforming to the XPS specification.

### **Properties**

The SDK provides package properties for reading and editing.

### **Documents and Pages**

The SDK provides an API for managing documents within packages and pages within documents.

### **Print**

The SDK makes it possible to print a single document or all the documents in a package.

## View/Render Pages

The SDK provides a draw procedure for displaying pages or rendering pages as raster images. This procedure uses the Direct2D API.

## Search in Text, Export Text

The SDK provides search functionality for text search within a document or package. Exporting plain text is possible.

## Outline, Links and Named Destinations

The SDK provides an API for navigating within documents/packages and for editing outline structures, links and named destinations.

## Export to PDF

The SDK enables the quick, high quality conversion of XPS documents to PDF documents.

## Rich Content Functionality for XPS

Rich Content provides functionality for creating and editing XPS files similar to the functionality provided for PDFs.

## Viewer

The SDK provides several platform-oriented implementations of our document viewer. Using these viewer implementations, you can quickly and easily integrate powerful PDF, XPS, text and image viewers into your application. You can also display your documents in your own structure and layout using these viewer implementations.

The viewer provides a wide spectrum of traditional document viewer functions: different layouts, zooming, scrolling, etc. Using powerful and effective event mechanisms provided by the viewer, your application can extend the standard functionality of the viewer and implement custom tools (please refer to the viewer's ToolMode property in the [PDF Xpansion SDK Reference](#)) for the visual processing of PDF or other documents.

The viewer is a window or control (according to the development platform) that displays PDF, XPS or other documents and processes for end user activity (touchscreen, mouse and keyboard input).

## Navigation

The layout canvas can be scrolled by the program or by the end user. The viewer allows navigation to specific pages or exact places on the page. The viewer also supports hyperlinks in PDF and XPS documents.



## Interactive Forms

The viewer provides the possibility to fill interactive form fields in PDF files.

## Read and Edit Document Notes

The viewer implements dynamic popups for markup annotations in PDF documents – the end user can read and/or edit annotations and notes in the PDF file.

## Annotation Tools

The viewer implements multiple interactive tools. These tools provide the end user with the ability to make annotations in PDF documents: text markup, textboxes and callouts, pen markup, sticky notes, stamps, etc. as well as to edit or delete the existing annotations. The annotations can be viewed in popup windows, inline as well as in a panel view. Dynamic notes are also supported.

## Interactive End User Tools

Some common tools provide additional comfort to the end user:

- Magnifying glass and snapshot (screenshot tools)
- Tracker and rectangle zones tool for visual annotation processing
- Stylus pen

## Customization

You can customize and extend the viewer functionality using the viewer API and the flexible event model.

## Global Settings and Object Factory

### Global SDK Settings

You can get or set several common SDK settings and properties: memory usage, file caching, local time offset, etc.

### Object Factory

You can create several SDK objects using the object factory interface.

## SDK Licensing Model – Pay Only Once!

This chapter describes the licensing model and is intended to help you select the optimal type of license. It is not a license itself.

The licensing model is simple, flexible and ensures our clients a reasonable price. This model has

---

a lot of **advantages** that distinguish it from other licenses for similar technologies.

### **Advantage 1: Royalty-free license.**

You only pay for it once. The price does not depend on the number of the customers the licensee has, on the number of products produced by the licensee using the PDF Xpansion SDK or on the number of products sold. The license does not require sales reports for products using the SDK libraries nor are there royalty payments. The license does not stipulate any annual payments during the whole period of using the purchased version of the PDF Xpansion SDK, except for optional maintenance payments.

### **Advantage 2: No developer licenses.**

The license does not stipulate that you to buy "developer licenses", which means that it doesn't restrict the number of workstations at a single company location where the library can be integrated into the licensee's products.

### **Advantage 3: Pay for what you really need.**

You purchase only the functionality you need at the moment and do not have to pay for the whole SDK or for unnecessary components.

### **Advantage 4: Extend functionality – it's easy!**

The only thing you need to do is order the additional functionality. You only pay for the additional components or features. There is no processing fee.

## **License Conditions**

- The scope of the licenses is limited to service functionality: using the PDF Xpansion SDK as the main functionality of a customer's products is not permitted. This means that the SDK cannot be used as the kernel or core of an application but it can be used as a service or a component of it.
- The final decision about an applicable license type will be made by the licensor, soft Xpansion.

## **Maintenance**

You can purchase technical support service and assistance with the integration of the SDK into client products, as well as maintenance, as an add-on. This is a paid service. Purchasing the service entitles you to:

- Free SDK updates – after the release of a new SDK version you will receive a license with the permissions included in the previous one
- Professional tech support via email or by phone

- Quick, priority bug fixing with short response times
- Small improvements to the current SDK version at the client's request, following coordination with the developers and based on the technical resources required

## License Types

There are three license types for the PDF Xpansion SDK: Office, Business and Corporate.

### Office License

The license permits the use of the PDF Xpansion SDK only within the company at a single company location with less than 100 employees (Windows PC workplaces). Commercial use (i.e., including it in your own products) is prohibited.

### Business License

The license permits the use of the PDF Xpansion SDK:

- For a distribution of the products created by the licensee as **standard applications**. All products must be released by a single copyright owner. An unlimited number of products and sales (royalty free) in up to five countries worldwide are permitted.
- In your own software for use within your own company, with an unlimited number of employees (workplaces) and for multiple company locations in up to 5 countries worldwide.

The Business license cannot be used in the following cases:

- Using the SDK in several individual projects (non-standard apps) outside your company
- Using the SDK in high-productivity server solutions (processing 5,000 documents per day)
- Using the SDK as part of another SDK or development tool
- Using the SDK as part of an intranet-based or Internet web service
- Global use in more than 5 countries (incl. development, internal use, and/or distribution)

### Corporate License

The special Corporate license permits the use of the PDF Xpansion SDK in cases that are not covered by the Office and Business licenses. Contact us for our Corporate license prices and conditions.

## License Agreement

You'll find more information about the various licenses and maintenance services in the license agreement. [You can request the agreement template via email.](#)

## Prices

The price for the PDF Xpansion SDK depends on the type of license and is made up of three components:

- Required functionality – only the features you need (one royalty-free fee)
- Additional properties and services – supported platforms, languages, etc. (additional one-time royalty-free fee)
- Maintenance – optional service, annual fee

Below you will find the current price list for the individual features, properties and services. Only the prices for the Business license are listed. They are intended to serve as an orientation and may vary depending on the actual requirements of the licensee. The prices for the Office license are usually 2-3 times lower. The prices for the Corporate license depend on how the PDF Xpansion SDK will be used in the licensee's tasks and projects and are calculated individually. Please [contact us](#) to request an offer based on your needs.

Functionality	Price (€)
PDF Functionality	
Open file/stream	1,000
Save file/stream	1,000
Open/save encrypted documents (open and save permissions are required)	200
Save changes in PDF as a revision (save permission is required)	300
Split/combine documents (open and save permissions are required)	500
Compose pages (merge content of pages, copy pages within document)	500
Page management (add, move, remove, resize, etc.)	100
Search text (and export plain text)	500
Print document	500
Render and/or display pages	2,000
Signatures – verify and sign, timestamps	500
Form fields – fill and read values, create and edit properties	500
FDF, XFDF (import/export/submit), HTML (submit)	500
Links, Bookmarks (TOC, Outline), Named destinations	300
Annotations – processing using the SDK API	
Text markup	300
Notes, stamps, watermarks	300
Textbox & callout	300
Pen, eraser	300
Shapes (line, rectangle, circle, etc.)	500
Media & 3D annotations	500
Attachment annotations	200
Replies & statuses for comments	500
Layers	300
Embedded files	300
Associated files (relationships of embedded files)	200
Processing XMP metadata	200
Logical document structure and marked content (tagged PDF)	1,000
Custom crypto filter	1,500
Articles	300
Document parts	300
Portfolio	500
Create/Process	
Explore PDF content (Rich Content Editor)	2,000

Create PDF content (Rich Content Editor)	2,000
Edit PDF content = Explore + Create	
Convert to PDF/A-1 <sup>1</sup>	1,500
Convert to PDF/A-2 <sup>1</sup>	1,000
Convert to PDF/A-3 <sup>1</sup>	1,000
Import from XPS (open XPS and save PDF permissions are required)	1,500
Import from SVG as a page	1,500
Import GDI metafiles (EMF) as a page	1,000
Import GDI+ metafiles (EMF+) as a page	1,500
Text vectorization	1,000
Detect content bounds	300
<b>XPS/OXPS Functionality</b>	
Open file/stream	750
Save file/stream	750
Split/combine documents/packages (open and save permissions are required)	500
Compose pages (merge content of pages, copy pages within document)	300
Page management (add, move, remove and transform)	200
Search text (and export plain text)	300
Print document	300
Draw/render pages	1,500
Signatures (verify/sign)	500
Processing	
Explore XPS content (Rich Content Editor)	1,000
Create XPS content (Rich Content Editor)	1,000
Import from PDF (open PDF and save XPS permissions are required)	1,500
<b>Viewer Functionality</b>	
Basic functionality (open document and display pages permissions are required)	1,000
Form filling	500
Popups, thumbnail view, snapshots (viewer state)	300
Media player API	500
WPF control	300
Interactive end user tools	
Magnifying glass & snapshot (screenshot) tools	200
Tracker tool (supports creating notes, stamps & watermarks)	200
Rectangle zones tool (supports creating/editing form fields, links, etc.)	200
Text selection, text markup	300
Stylus	300

Management of PDF annotations (move, resize, delete, etc.)	300
PDF textbox & callout	300
PDF pen, eraser	300
PDF shapes (line, rectangle, circle, etc.)	300
Page content editor tool (requires the explore PDF content and create PDF content functionalities)	2,000
Dynamic notes	
Text markup	500
Comments & pen	500
Shapes	300
<b>Other Functionality</b>	
Rich text box	500
Image/text documents	300
Fragment/composite documents	500
Acquisition documents (scanners & webcams) <sup>2</sup>	300
JavaScript actions <sup>3</sup>	1,500
ZUGFeRD invoices:	On demand
Validate invoice file	
Read data from invoice	
Create conform invoice file	On demand
XFA forms:	
Preview forms	
Print form data	
Access form data (read and/or change)	--
Text recognition (OCR) <sup>4</sup>	
<b>Developer Platform/Processor Architecture/Text Processing</b>	<b>% Fee</b>
<a href="#">Windows Desktop/Service apps (C++)</a>	--
.NET apps (C#/Visual Basic)	10%
UWP / WinStore / WinPhone apps (C++/C#/Visual Basic)	20%
COM API (for C++Builder, Delphi, etc.)	10%
<a href="#">x86 architecture</a>	--
x64 architecture	10%
ARM architecture	20%
<a href="#">Support for West European languages (ISO 8859-1 encoding)</a>	--
Support for European languages (ISO 8859-2,3,4,5,7,9,10,13,14,15,16 encoding)	5%
Support for Asian languages	10%
Support for Arabic, Hebrew and other RTL languages	30%

<sup>1)</sup> If you buy two versions of PDF/A you will receive a 500 EUR discount.

If you buy 3 versions of PDF/A you will receive a 1000 EUR discount.

<sup>2)</sup> This feature uses WIA/WIA2 technologies for desktop versions of Windows.

<sup>3)</sup> The JavaScript interpreter for embedded PDF scripts. Available for desktop versions of Windows only. The document object model is restricted according to what Adobe Reader provides.

<sup>4)</sup> This feature uses the Tesseract OCR engine. It is free and does not come with support.



# INTEGRATION WITH TEXT RECOGNITION FUNCTIONALITY (OCR)

Text recognition (OCR) is included in the PDF Xpansion SDK and is implemented using the [Tesseract OCR engine](#) (HP, Google). This feature enables you to detect text in images in your PDF files and to make your scanned PDF documents searchable. The detected text can also be exported as plain text from scanned PDF pages or from images that contain text content. See the [PDF Xpansion SDK Reference](#) for more information. You do not need to use the Tesseract engine directly in order to use this feature.

Current version: 3.xx.

Number of supported languages: 60.

**This feature in the PDF Xpansion SDK is completely FREE!**

The Tesseract OCR engine is available under the [Apache License 2.0](#).

# CONTACT US

## **soft Xpansion GmbH & Co. KG**

Koenigsallee 45, 44789 Bochum, Germany

Phone: +49 234 2984171

Fax: +49 234 2984172

E-mail: [pdf@soft-xpansion.com](mailto:pdf@soft-xpansion.com)

[www.soft-xpansion.com](http://www.soft-xpansion.com)