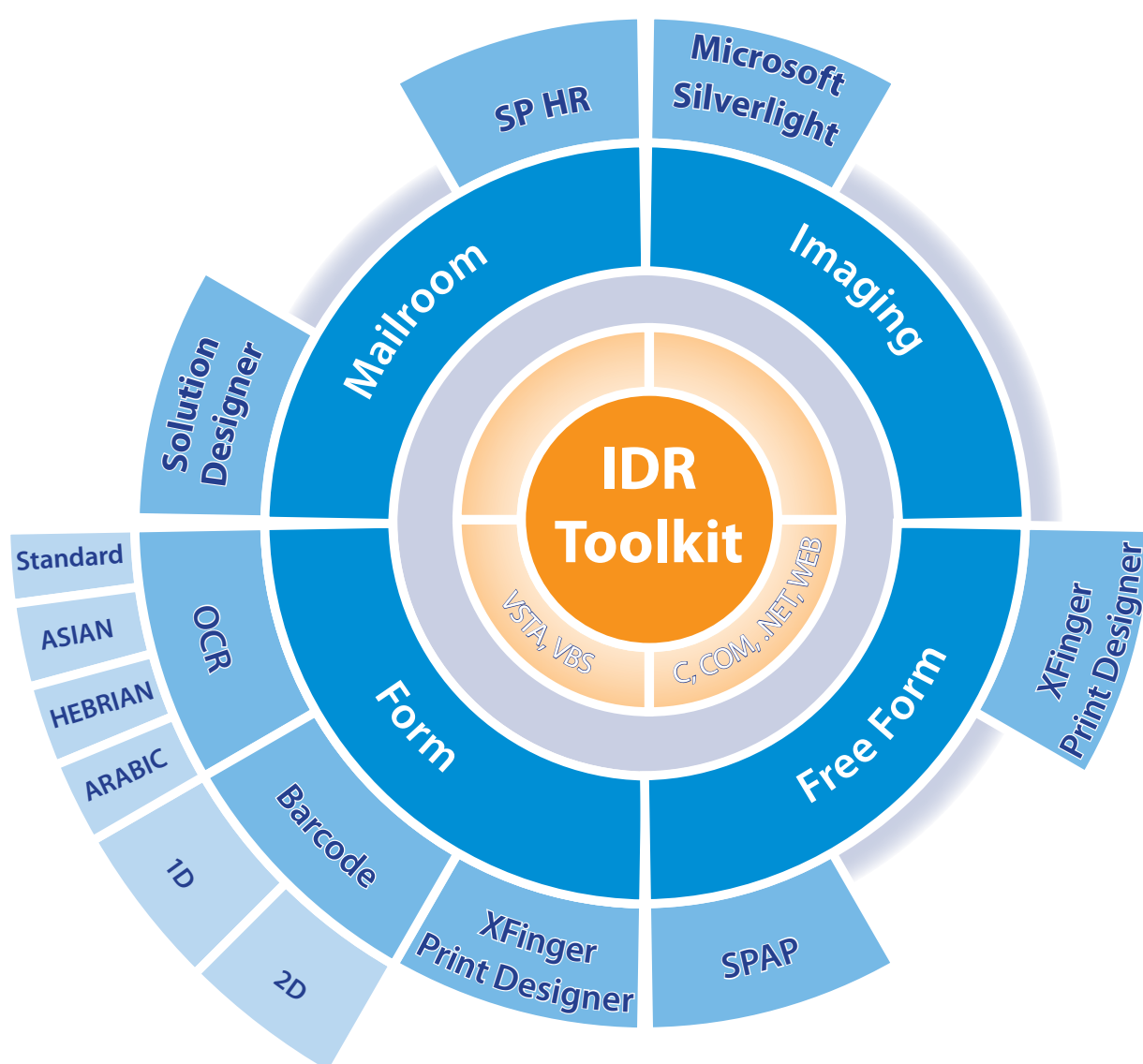


# IDR™ Toolkit

Documentation IDR Toolkit V3.0.1



## Content

1. Introduction .....	3
2. System requirements .....	4
3. Installation and activation .....	4-7
4. Application programming interface .....	8-9
5. Sample application .....	9
6. Appendix 1: result structure .....	9-11
7. Appendix 2: 3 <sup>rd</sup> party software copyright notices .....	11

## 1. Introduction

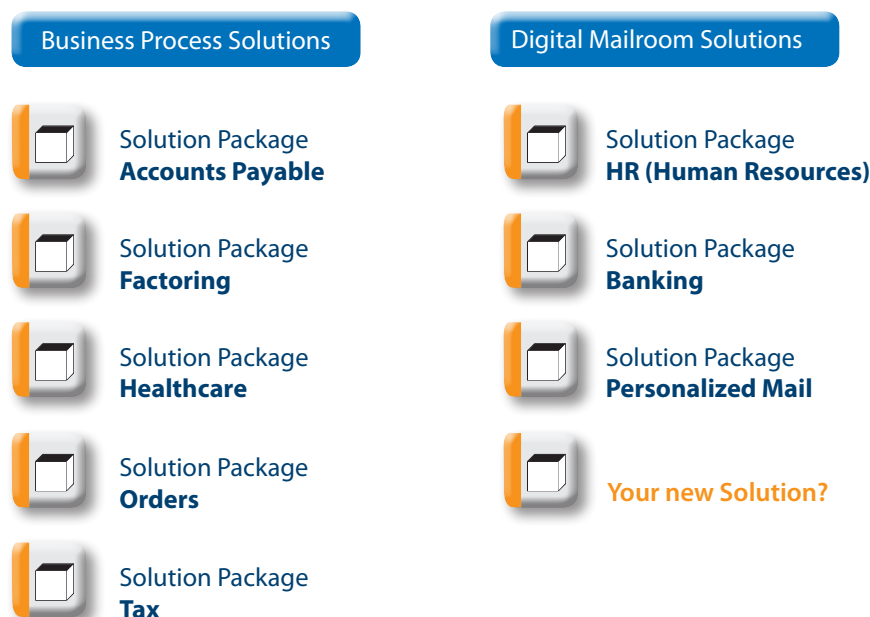
This document describes the release 3.0.1 of the IRIS IDR Toolkit. This Toolkit includes a runtime SW plus test-application (see below) and either the XFingerPrint Designer for Forms or semi-structured documents or the Solution Designer for classification of unstructured documents as development environment. These designers will be needed to develop the extraction-project for the OEM runtime solution. There are 3 different types of IDR Toolkit available:

### IDR Toolkits - ascending functional scopes/levels -



Optionally predefined solution packages (f.e. SP Accounts Payable, SP Human Ressources) which come with their own editor can be integrated as well:

### Solution Packages for the IDR Toolkit



The distribution folder contains following parts:

- ▶ Install: installation packag
- ▶ Redist:
  - Licence\_XXX.lic - license file.
  - Docutec.AnalyzeOEM\_EXE.Interop.dll - primary interop assembly for using by the hosting application
- ▶ SDK: additional SDK resources
- DemoAnalyzeSDK: sample application to test and learn the IDR Toolkit API (needs Microsoft VisualStudio 2005 or 2008)

## 2. System requirements

The following operation systems are supported by this release:

- ▶ Windows XP SP2 or SP3
- ▶ Windows Server 2003 SP2, Windows Server 2003 R2 SP2 (x86)
- ▶ Windows Vista Business SP1 (x86)

At least 2 GB free space must be available on the system drive.

## 3. Installation and activation

**To install IDR Toolkit** change into the folder Install and start *XAnalyze\_SDK.exe*.

*XAnalyze\_SDK.exe* is a bootstrapper that installs on demand some missing system components (like Microsoft .NET framework) and infrastructure components like DReco package.

After all prerequisites are installed, the package "Analyze SDK" will be installed.

**Alternatively**, individual packages can be installed separately (e.g. if packaging IDR Toolkit into a deployment package of the hosting application). In this case the following installation order must be provided:

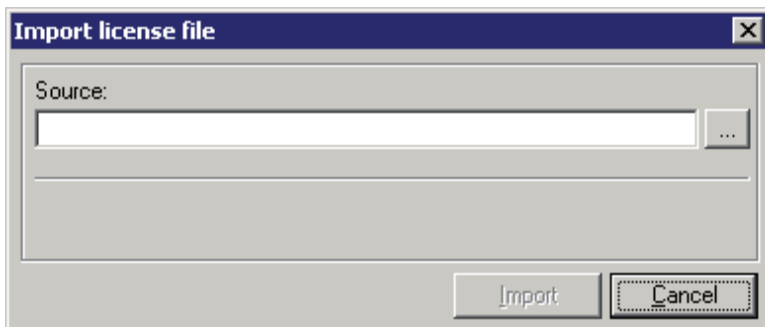
- Microsoft .NET Framework 2.0 (dotnetfx\dotnetfx.exe)
- Microsoft Visual C++ 8 SP1 redistributable (vcredist\_x86\vcredist\_x86.exe)
- DReco (DReco\DReco.msi)
- IDR Toolkit kernel (XAnalyzeSDK\XAnalyzeSDK.msi)

- Optionally: IRIS IRIS-Docutec SPAP 4.0 (SPAP\Context\_AccountsPayable\_Xtract3.msi) by running setup.exe from the subfolder SPAP

IDR Toolkit uses a software licensing technology and needs to be activated before using. Within the activation process a hardware-dependent key will be generated by the activation utility and must be transferred to IRIS-Docutec. Licensing service at IRIS-Docutec will generate an activation key, that should be entered into the activation utility.

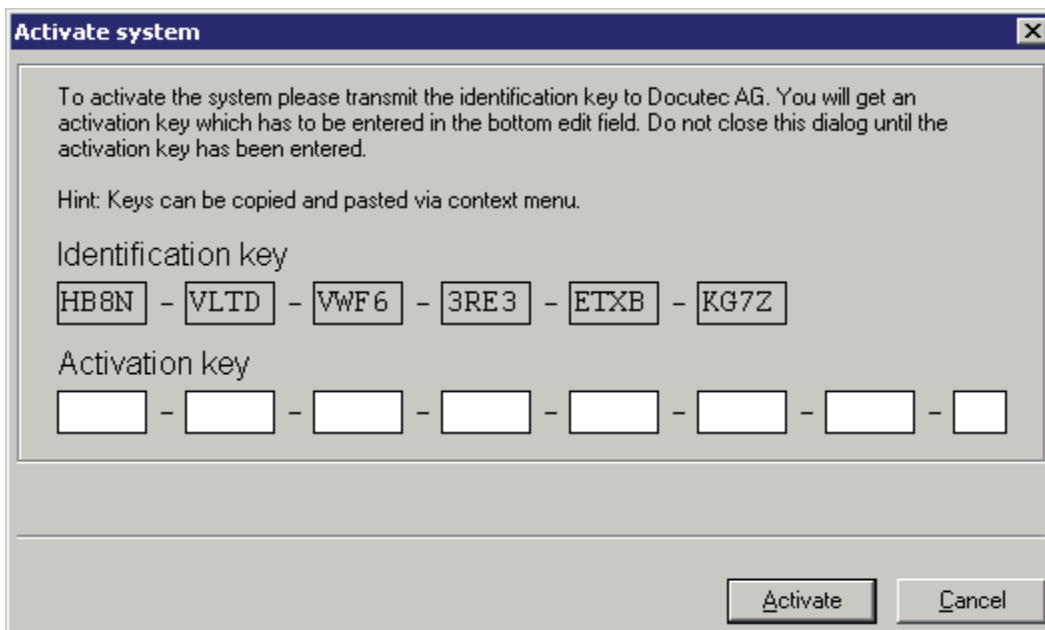
### To activate IDR Toolkit:

1. Start the activation utility *LicenseFS.exe*, usually located in C:\Program Files\Common Files\IRIS-Docutec (location may vary dependent on OS language).
2. If no license file exists an import dialog will be shown:



Click to the browse button, navigate to the license file, e.g. *2009\_01\_09\_XAnalyze\_SDK\_Test\_SPAP\_V1\_LP91.lic* and click "Import".

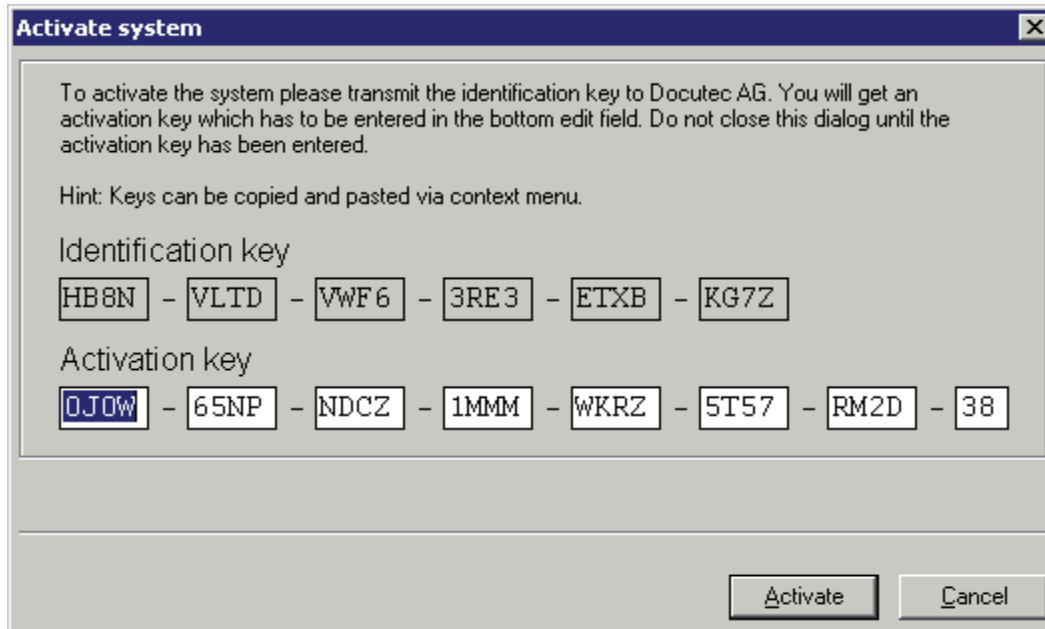
3. An activation dialog will be shown:



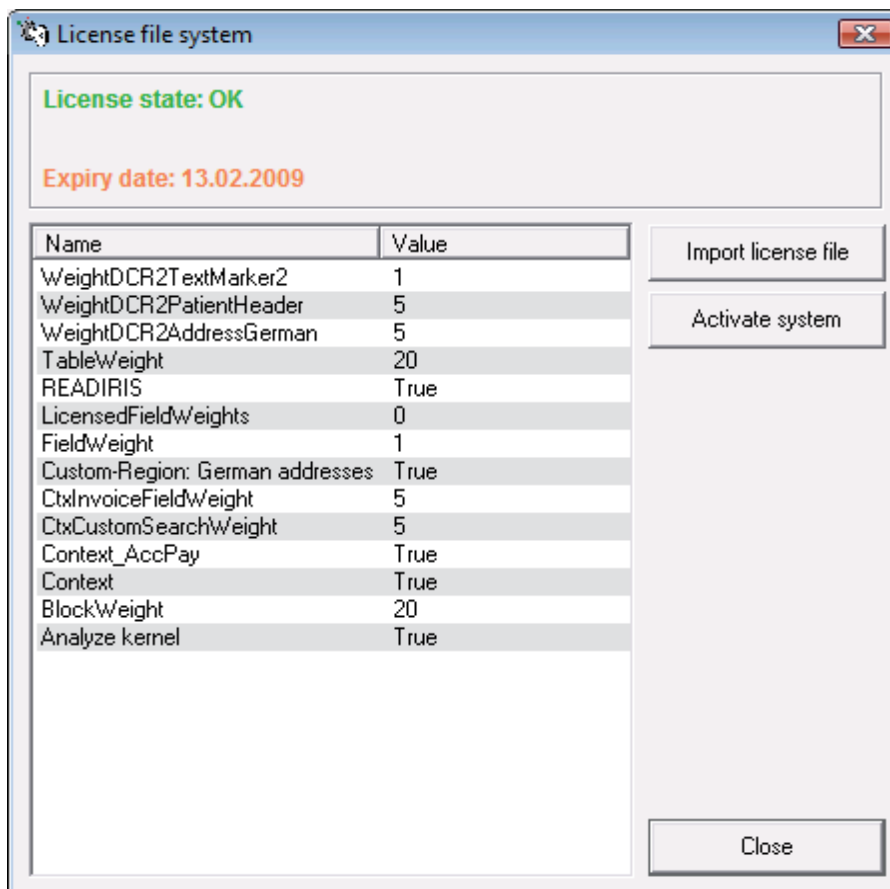
Copy the identification key (e.g. by using the context menu command to copy the key to the clipboard) and send it to [register@docutec.de](mailto:register@docutec.de).

**Important:** do not close the activation dialog until activation is completed!

- After receiving the activation key enter it into appropriate fields (you may use the context menu command “paste” here):



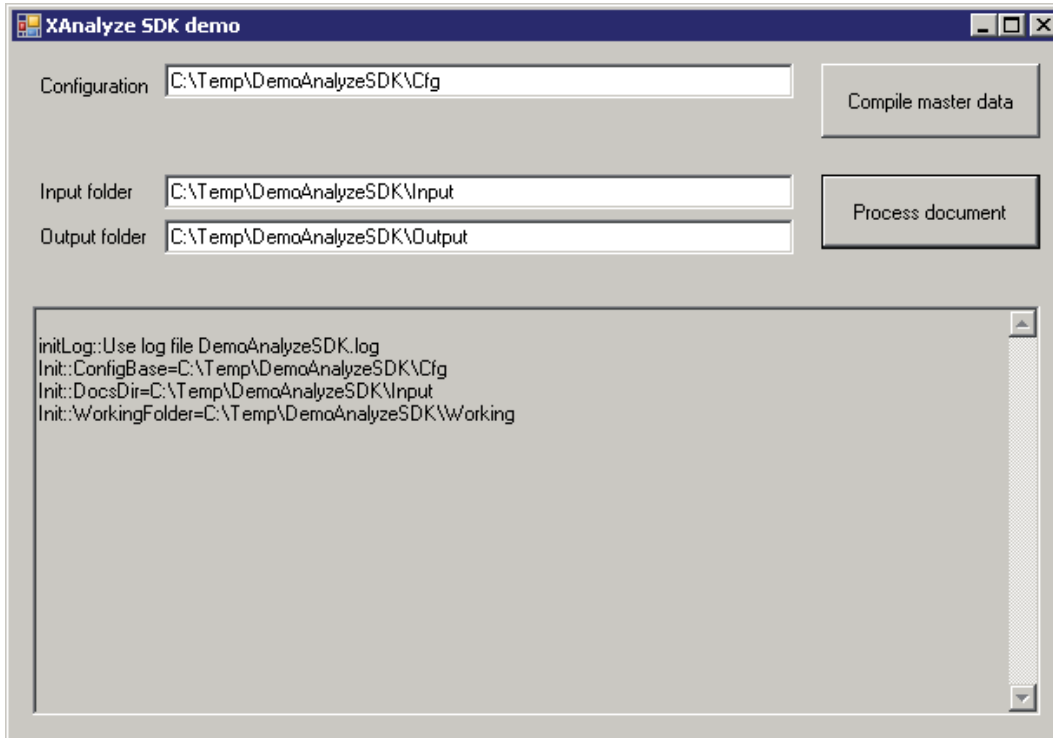
and press „Activate“. Message „Key is valid“ appear and, after clicking to „OK“ the license information will be displayed:



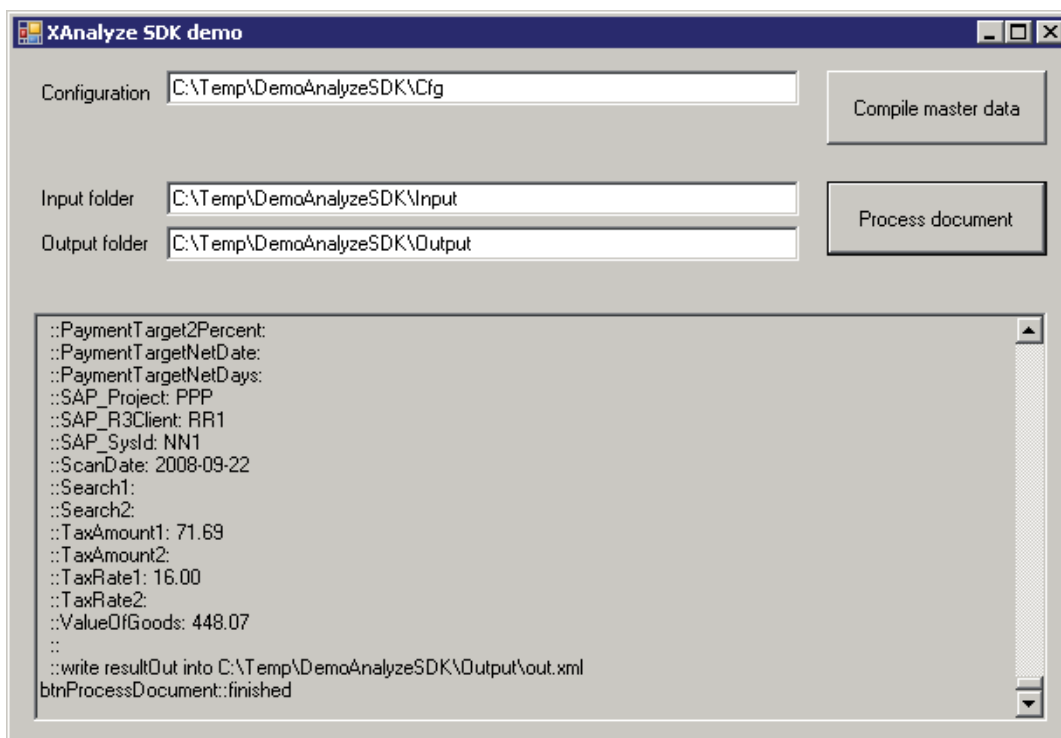
**Note 1:** The shown license is limited to 91 days.

## To test the installed SDK:

1. Copy the folder *SDK\DemoAnalyzeSDK* to a local disk location.
2. Start the demo application ***DemoAnalyzeSDK.exe***:



3. Press to "Process document". After couple of seconds, the document will be processed and fields are dumped into the log window:



## 4. Application programming interface

An application programming interface (API) for XAnalyze may be used by the COM or Microsoft .NET based clients. For the latter case, the *primary interop assembly* Docutec.AnalyzeOEM\_EXE.Interop.dll is provided in the folder *Redist*.

Following calls are available:

### ► Global initialization

```
void Init(string workingFolder, string licenseK
```

The initialization function initializes the Analyze instance.

The *workingFolder* will be used for logging purposes and as a temporary storage. Each instance of the XAnalyze kernel needs a dedicated storage folder, the content of the folder is only important for the lifetime of the instance. For the sake of performance a folder within the local file system should be preferred.

The parameter *licenseKey* must be an empty string.

### ► Handling certain configuration

```
Load Environment(string rootFolder, string clientName)
```

Set the current environment for the XAnalyze kernel. It may be called at any time to switch between environments. *rootFolder* points to the project configuration folder, *clientName* specifies name of the project.

```
CloseEnvironment(string rootFolder, string clientName)
```

Closes the current configuration environment.

### ► Handling master data

The hosting application prepares/detects changed master data and calls the function in one instance of XAnalyze kernel to force compiling of master data:

```
CompileMasterData(string rootFolder, string clientName)
```

Other instances of the hosting application may call *LoadMasterdata()* to force reloading the actualized master data:

```
LoadMasterData(string rootFolder, string clientName)
```

### ► Document processing

- For each document of a batch:

```
InitializeDocument(int command, byte[] docParams
```

The parameter *command* must be 1 for now, optional data *docParams* may contain an XML UTF8-formatted stream in a following form:

```
<?xml version='1.0' ?>
<xml>
  <PARAMS>
  <SCAN_OPERATOR>JohnDoe</SCAN_OPERATOR>
  </PARAMS>
</xml>
```



For parameters available for the Solution Package *Accounts Payable* see SP AP documentation.

- Next, for each page in the document use the call

```
LoadPage (byte[] imageBuffer)
```

pass a bitonal G4-compressed TIFF image.

- After passing the last page of the document start processing by calling

```
ProcessDocument (ref byte[] resultOut, ref byte[] paramClientOut)
```

If extraction succeeded the *resultOut* byte stream contains an XML stream with document extraction result (see Appendix 1 for information about result structure).

## 5. Sample application

The sample application is a C# project for Microsoft Visual Studio 2005

## 6. Appendix 1: Document result structure

The document result structure for Solution Package *Accounts Payable* is described in the SP AP documentation, here is a cut-out of XML representation:

```
<?xml version="1.0" encoding="UTF-8" ?>
<Root xmlns:dt="urn:schemas-microsoft-com:datatypes"
  NAME="XtractContext_INVOICE" RIDVERSION="1.1.3.1">
  <CONTROL>
    <CIDMINOR dt:dt="i4">11</CIDMINOR>
  </CONTROL>
  <REGIONS>
    <FIELD NAME="ARCHIVID">
      <STATE dt:dt="i4">5</STATE>
      <VALUE />
    </FIELD>
    <FIELD NAME="DOCTYPE">
      <STATE dt:dt="i4">5</STATE>
      <VALUE>
        <RECT>
          <R>119,476,353,507</R>
        </RECT>
        <TEXT>RMB</TEXT>
      </VALUE>
    </FIELD>
    <FIELD NAME="CreditorBankNumber">
      <STATE dt:dt="i4">5</STATE>
      <VALUE>
        <RECT>
          <R>670,3120,816,3144</R>
        </RECT>
        <TEXT>39050000</TEXT>
      </VALUE>
    </FIELD>
  </REGIONS>
</Root>
```

```

    </VALUE>
</FIELD>
<FIELD NAME="SEARCH1">
  <STATE dt:dt="i4">5</STATE>
  <VALUE>
    <RECT>
      <R>119,476,353,507</R>
    </RECT>
    <TEXT>searchresult1...</TEXT>
  </VALUE>
</FIELD>
<FIELD NAME="SEARCH2">
  <STATE dt:dt="i4">5</STATE>
  <VALUE>
    <RECT>
      <R>119,476,353,507</R>
    </RECT>
    <TEXT>serachresult2...</TEXT>
  </VALUE>
</FIELD>
...
<TABLE NAME="ITEMS">
  <COLS>
    <COL NAME="ORDERMATNR">
      <REGIONTYPE dt:dt="i4">2</REGIONTYPE>
    </COL>
    <COL NAME="ORDERNUMBER">
      <REGIONTYPE dt:dt="i4">2</REGIONTYPE>
    </COL>
  </COLS>
  <ROWS>
    <ROW>
      <RECT>
        <R>1490,2117,2399,2202</R>
      </RECT>
      <REGIONS>
        <FIELD NAME="ORDERMATNR">
          <STATE dt:dt="i4">5</STATE>
          <VALUE>
            <RECT>
              <R>124,2117,2463,2202</R>
            </RECT>
          </VALUE>
        </FIELD>
        <FIELD NAME="ORDERNUMBER">
          <STATE dt:dt="i4">5</STATE>
          <VALUE>
            <RECT>
              <R>124,2117,2463,2202</R>
            </RECT>
            <TEXT>500218402</TEXT>
          </VALUE>
        </FIELD>
      </REGIONS>
    </ROW>
  </ROWS>
</TABLE>

```

```

</REGIONS>
<RESULTTYPE dt:dt="i4">1</RESULTTYPE>
<SHEETS>
  <SHEET>
    <FRONT NAME="ConText_With_Positions_1" RIDVERSION="1.1.3.1">
      <CONTROL>
        <CIDMAJOR dt:dt="i4">4</CIDMAJOR>
      </CONTROL>
      <HEIGHT dt:dt="i4">3507</HEIGHT>
      <RESX dt:dt="r8">300.000000</RESX>
      <RESY dt:dt="r8">300.000000</RESY>
      <WIDTH dt:dt="i4">2480</WIDTH>
    </FRONT>
  </SHEET>
</SHEETS>
<STATE dt:dt="i4">3</STATE>
</Root>

```

## 7. Appendix 2: 3<sup>rd</sup> party software copyright notices

IDR Toolkit uses some open source software:

- **zlib** - general purpose compression library <http://zlib.net/>  
Copyright (C) 1995-2005 Jean-loup Gailly and Mark Adler
- **LibTIFF** - TIFF Library and Utilities <http://www.libtiff.org/>
- Copyright (c) 1988-1997 Sam Leffler  
Copyright (c) 1991-1997 Silicon Graphics, Inc.
- **libjpeg** - JPEG image compression and decompression library <http://developer.novell.com/wiki/index.php/Libjpeg>  
This software is copyright (C) 1991-1998, Thomas G. Lane.  
This software is based in part on the work of the Independent JPEG Group.



Imprint:  
© 2008-2009 IRIS-Docutec AG. All rights reserved.  
Document Version 1.04  
Date: June 9, 2009  
Author: Andrej Jantzen, Docutec AG

I.R.I.S. SA  
Rue du Bosquet 10  
1348 Louvain-la-Neuve - Belgien  
Tel.: +32 (0) 10 45 13 64  
Fax: +32 (0) 10 45 34 43  
info@irislink.com  
www.irislink.com