Optimize
what you scan!

IRISDocument™ 9

Convert your images to searchable,
structured and compressed documents

I.R.I.S.
Document to Knowledge

IRISDocument™ Global Benefits

IRISDocument™ is the most competitive and fastest professional solution on the market to convert all images into indexed and editable files, and provide multiple users with OCR services.

IRISDocument™ is:

Fast

With the Multi-CPU Add-on, spread the workload across up to four cores or more to make the best use of your computer and server resources. At the highest text recognition rates, it will process up to 60,000 images/day on a single PC. Convert large batches of documents automatically with minimal user intervention thanks to the Watched Folder mode.

Easy to use

IRISDocument™ is quickly and easily installed – no expensive experts needed.

Affordable

Although packed with features, IRISDocument™ comes at a very reasonable price.
A single license will allow you to provide your whole organization with OCR services.
IRISDocument™ has no page limitation - no other solution will give you as much value for your money.

Versatile and interconnectible - Create your own complete scanning platform!

IRISDocument™ is fully integrated with I.R.I.S.' scanning solution IRISPowerscan™.

From any high-speed scanner, scan, index, OCR and export your documents to create a complete document workflow from front-door content to ERP or Document Management Systems (like Microsoft® SharePoint®), all the while keeping your indexes!

IRISDocument™ is versatile; you can also import scanned images from many applications which have become the industry standard (see page 6).

And if that isn't enough, enjoy IRISDocument™'s simple integration into your own system through command line call and export API.

Accurate and reliable

IRISDocument™ will help you define the parameters you need to get the best OCR results on your documents thanks to a powerful OCR Accuracy Report.

Feature-rich

- Improve your scanned images with powerful image cleaning tools.
- Deskew your text, detect its orientation… and much more.
- Rename and separate your documents automatically – IRISDocument™ will detect barcode values and blank pages.
- Powerful recognition engines for languages, barcodes, banking fonts, patch codes, etc.
- Index your documents, and store the extracted data into a descriptive XML index file.
- Thanks to the revolutionary iHQC™ hyper-compression technology, share your documents easily.
- Convert your document images into many formats, among which PDF, XPS (with possible hyper-compression), etc.
- Think about the future: IRISDocument™ can convert your documents to PDF/A-1b for long-term archiving.

Enjoy The Most Productive OCR Technology

Recognition engine and full page OCR

IRISDocument™ uses I.R.I.S.' latest recognition engine and formatting technologies. Page analysis is carried out on the color image, resulting in better zone segmentation. This new engine also recognizes the color of the text and keeps it in the output file.

More than 130 languages supported

All European languages are supported, including Central European, Baltic, Greek and Cyrillic ('Russian') languages.

An Asian Add-On module is available for the recognition of 4 Asian languages: Japanese, Korean, Traditional and Simplified Chinese.

Hebrew and Arabic/Farsi Add-On modules are also available.

Multilingual recognition

IRISDocument™ recognizes up to 5 languages, even with mixed alphabets, on the same document.

Image enhancement

- **Deskew**
  The skewing of the text is analyzed on the front of the page but is applied both on the front and the back of the document, resulting in a much faster processing speed.

- **Despeckle**
  The despeckle feature cleans up images by removing noise that can appear when scanning documents.

- **Smoothening of color images**
  Allows for optimal recognition results by smoothening out the color transition in the image.

- **Page orientation detection**
  Automatic rotation and straightening of the text orientation contained in your documents for optimal OCR accuracy.

Enjoy Countless Output Possibilities

Multiple output formats

Generate several output files with different file formats in one single process and from one image file.

Wide number of file output formats

The following output formats are available: PDF, PDF/A-1b, XPS (those three formats are also available with iHQC™ compression), standard text files, OpenDocument Text, XML , etc. (see full list on page 7).

Formatting options

IRISDocument™ will answer the most common reformatting needs, e.g. for Word and RTF output files: "retain color of text", "retain color of background", "keep paper size", etc.

Standard image compression

Modify JPEG compression rate and image resolution when generating PDF files (with or without text files) and highly compress all your images with JPEG 2000 technology.

Key Features – a Closer Look

- Powerful and fast OCR engine (up to 60,000 images/day on a single PC) and high text recognition accuracy.
- Conversion of your images into many output formats. (See full list on page 7).
- Document renaming and separation features based on barcode value and blank page detection.
- Deskew and orientation detection.
- Banking Fonts are now included - no additional registration required anymore.
- With the multi-CPU Add-on, spread the workload across up to four cores to make the best use of your computer resources.
- With the iHQC™ add-on, create highly compressed PDF, PDF/A-1b or XPS files all while maintaining original text legibility and image color quality (see page 10).
- Fully integrated with IRISPowerscan™ - Scan, Index, OCR and Export your documents, creating your own complete scanning platform.
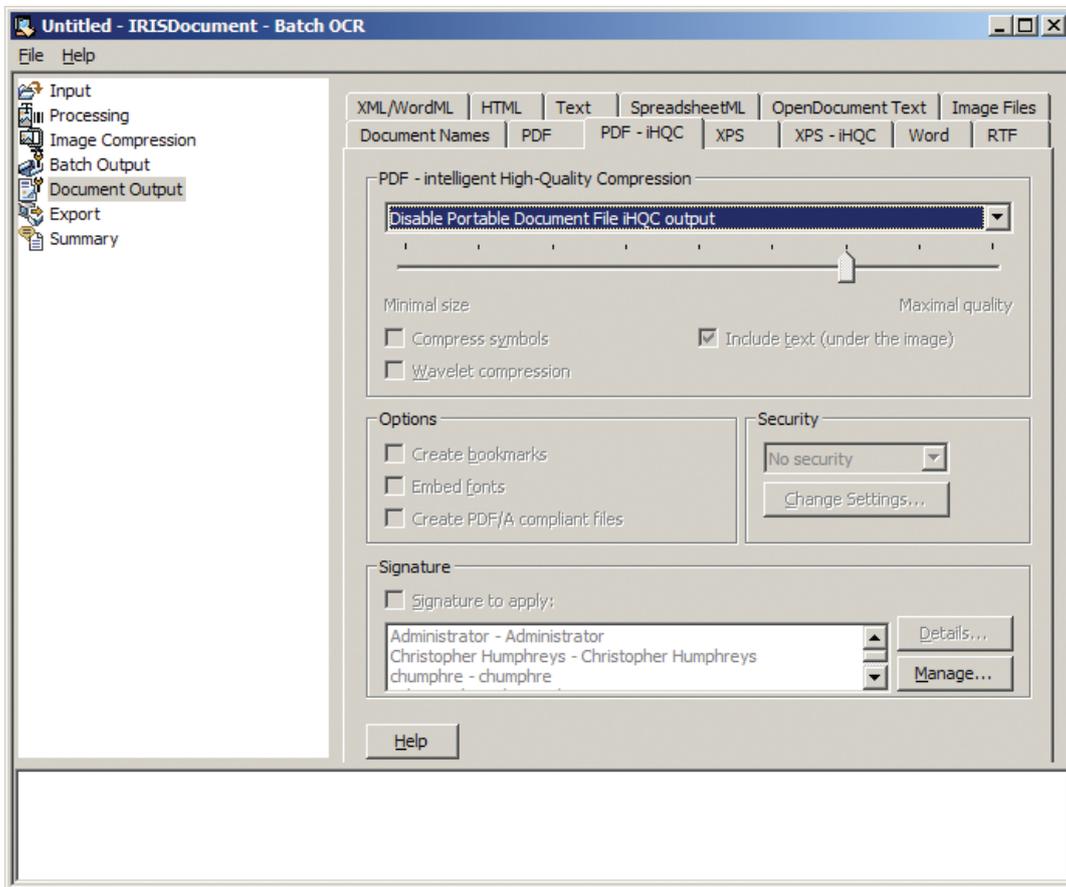
Three Operation Modes to Meet Your Needs

1 - Batch OCR mode

Large batches of scanned images are easily converted into documents with minimal user intervention. A settings file, configured once with the user interface, is loaded each time you process the images. It is possible to save and use different settings files for each different project.

2 - Watched Folder mode

Define as many watched folders as you want to convert incoming image files into fully searchable and editable documents. Each folder has specific settings such as OCR settings, output file type, output folder location, process priority level, etc.

Another application is the setup of a server with user folders, which will provide multiple users with OCR services.

Finally, you can also use the Watched Folder mode as a Windows® service.

3 - Integrated OCR module into your DMS

Simply call the executable of IRISDocument™ in your application, with the settings file as parameter, and IRISDocument™ Server is integrated. The settings file is plain ASCII and is easily created by your application on the fly.
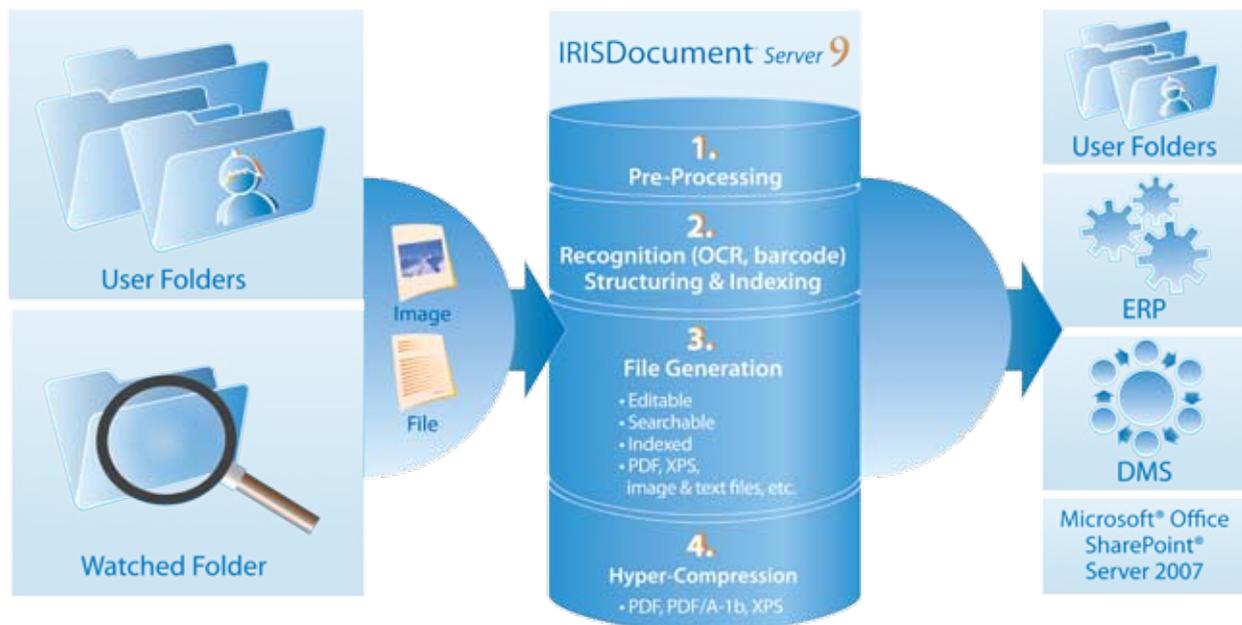
Eight Steps for a Complete Workflow, from Input to Archive and Export

1 - Input

In a few clicks IRISDocument™ will perfectly adapt to your infrastructure and practices

- **Batch OCR or Watched Folder:** Select an input folder for IRISDocument™ or let it watch a folder for you
- **Folders and subfolders:** Pick up images from the root and/or its subfolders
- **Input file types:** JPEG (*.jpg, *.jpeg), JPEG 2000 (*.j2c, *.j2k, *.jp2), TIFF images (*.tif, *.tiff), Windows® bitmaps (*.bmp), Adobe® Acrobat® PDF (*.pdf), DCX fax (*.dcx), DjVu® (*.djv, *.djvu), GIF images (*.gif), Photoshop® images (*.psd), PICT (Mac OS® format) (*.pct), Portable Bitmaps (*.pbm), Gem Paint™ (*.img), Zsoft Paintbrush® (*.pcx), Portable Network Graphics (*.png)

2 - Pre-processing of the images

Optimize your image files and take your scanner's job even further

- **Deskewing:** remove any skew of the text in the image
- **Orientation detection:** rotate the image
- **Despeckling:** remove noisy pixels
- **Smoothening:** flatten out the colors of the image, which will remove the JPEG compression artifacts and help recognition

3 - Recognition

- Choose from **137 OCR languages**:
  American English, British English, Afrikaans, Albanian, Aymara, Balinese, Basque, Bemba, Bikol, Bislama, Brazilian, Breton, Bulgarian, Byelorussian, Catalan, Cebuano, Chamorro, Corsican, Croatian, Czech, Danish, Dutch, Esperanto, Estonian, Faroese, Fijian, Finnish, French, Frisian, Friulian, Galician, Ganda, German, Greek, Greenlandic, Haitian (Creole), Hani, Hiligaynon, Hungarian, Icelandic, Ido, Ilocano, Indonesian, Interlingua, Irish (Gaelic), Italian, Javanese, Kapampangan, Kicongo, Kinyarwanda, Kurdish, Latin, Latvian, Lithuanian, Luxemburgh, Macedonian, Madurese, Malagasy, Malay, Maltese, Manx (Gaelic), Maori, Mayan, Minangkabau, Nahuatl, Norwegian, Numeric, Nyanja, Nynorsk, Occitan, Pidgin English, Polish, Portuguese, Quechua, Rhaeto-Roman, Romanian, Rundi, Russian, Samoan, Sardinian, Scottish (Gaelic), Serbian, Serbian (Latin), Shona, Slovak, Slovenian, Somali, Sotho, Spanish, Sudanese, Swahili, Swedish, Tagalog, Tahitian, Tok Pisin, Tonga, Tswana, Turkish, Ukrainian, Waray, Wolof, Xhosa, Zapotec, Zulu, Bulgarian - English, Byelorussian - English, Greek - English, Macedonian - English, Russian - English, Serbian - English, Ukrainian - English, Hebrew - Simplified Chinese (optional), Traditional Chinese (optional), Japanese (optional), Korean (optional), Arabic and Farsi (optional) + Moldovan, Bosnian (Cyrillic and Latin), Tetum, Swiss-German and Kazak.
- Asian (Chinese, Japanese, Korean), Hebrew, Arabic and Farsi are available as add-ons.
- IRISDocument™ also features accurate **multilingual recognition** (up to 5 languages on the same document).
- **MICR:** OCR-A, OCR-B, E13B and CMC7.
- **Barcode:** Codabar, Code 128, Code 39, Code 39 extended, Code 39 HIBC, Code 93, Datalogic 2 of 5, Discrete 2 of 5, EAN-13, EAN-2, EAN-5, EAN-8, IATA 2 of 5, Interleaved 2 of 5, MSI pharmaceutical, MSI-Plessey, PDF-417, PostNet, PostNet 32, PostNet 52, PostNet 62, QR Code, Reverse 2 of 5, UCC-128, UPC-A, UPC-E and UPC-E extended.
- **Dot Matrix.**
- Define a **User Lexicon.**
- Fine-tune the **Recognition Vs. Speed** ratio.
- Choose to apply recognition to selected pages only.

4 - Document Structuring

Define the content of your output documents and the structure of your batches

- Create a document per Batch, Image Folder, Input File, Input Document (e.g. from IRISPowerscan™).
- Split a document based on detection of barcodes or blank pages (which you can decide to keep or not).

5 - Indexing

- Generate a full-page text index for every image.
- Recognize the value of barcodes.
- Export that information as an index file in XML format.
- You can also use the indexed information for file renaming.

6 - Document creation

- Output formats:
  - **PDF** (4 different file modes: Image, Image-Text, Text, Text-Image, with many options: PDF/A-1b compliance for long-term archiving, certification, encryption).
  - **PDF with iHQC™ hyper-compression** (3 different compression levels, many options: PDF/A-1b compliance, certification, encryption).
  - **XPS** (4 different file modes: Image, Image-Text, Text, Text-Image).
  - **XPS with iHQC™ hyper-compression** (3 different compression levels, many options).
  - **Common file formats** such as TXT, RTF, DOC, OpenDocument text, HTML, XML, WordML, SpreadsheetML.
  - **Image file formats** such as TIFF (single or multi-page), BMP, JPEG, JPEG2000, or PDF image (also available with iHQC™ hyper-compression).
- IRISDocument™ offers you the possibility to activate multiple output formats simultaneously. This allows you for instance to convert an image into PDF/A-1b for long-term archiving, and at the same time to PDF with iHQC™ hyper-compression for sharing by e-mail within your organization.
- Various settings specific to the output format are also available, for instance layout retention, document protection by password, electronic signature with or without embedded fonts, etc.

7 - Document naming and file output

IRISDocument™ will make sure you get the right documents into the right folders

- **Document naming:** IRISDocument™ offers many options to name your documents: automatic incremental number, indexing field (available only with IRISPowerscan™, eCopy ShareScan® or Kodak Capture Software®), the name of the first image, the name of an image folder, the content of barcode, or the first sentence of recognized text.
- **Mirror input folder structure:** don't lose time reorganizing your documents, IRISDocument™ will make sure the structure you defined upstream is the one you get downstream.
- **Copy source images to output:** decide whether you want to get rid of your original scanned images, or keep them for further reference.

8 - Export

Call another application to process the created documents

- **Send the documents to any external application**
  Call any external executable directly from IRISDocument™ (in both Batch OCR or Watched Folder modes); it will receive the XML index file as an argument.

- **Connect and send the documents to Microsoft® SharePoint®**
  IRISDocument™ connects directly to any Microsoft® SharePoint® server and exports your documents automatically. For optimal SharePoint® connectivity, use IRISDocument™ in combination with I.R.I.S.' scanning application IRISPowerscan™. Based on settings defined prior to scanning, IRISPowerscan™ will recognize the type of document and capture indexes. According to their type, documents are then exported with their indexes directly into the right SharePoint® library.

IRISDocument™ at the Heart of Your Own Scanning & Archiving Platform

When starting from paper, IRISDocument™ can connect to a scanning application and work off of the freshly scanned images directly.

Depending on your needs you may use IRISDocument™ synchronously with the scanning program, or connected to it through an asynchronous process allowing you to completely free up the scanning platform from any heavy duty processing needed by IRISDocument™. In this case, IRISDocument™ becomes IRISDocument™ Server.

IRISDocument™ for IRISPowerscan™

I.R.I.S.' scanning platform IRISPowerscan™ comes standard with an embedded version of IRISDocument™. This version is ideal for limited volumes of paper, and will provide the user with searchable and compressed documents instantly. In this mode IRISPowerscan™ and IRISDocument™ share the computer resources.

IRISDocument™ Server for IRISPowerscan™

A special stand-alone version of IRISDocument™ was developed for IRISPowerscan™, in order to make optimal use of your resources.

IRISDocument™ Server connects seamlessly to any number of IRISPowerscan™ scanning stations.

Once the scanning is over, the IRISPowerscan™ station exports the scanned images to your server for IRISDocument™ Server to process.

A settings file is sent along with the images so that IRISDocument™ Server knows what to do with images coming from scan station A or B.
For example, scan images to hyper-compressed searchable PDF documents if they come from Accounts Receivable, to Word documents if you want to update existing contracts, etc.

This combination makes full usage of all the stations dedicated to the dematerialization process.

You could push it even further with the Multi-CPU Add-on, making sure any multi-core computer in your organization develops its full capacities to reach off-the-chart processing speeds.

On top of this "in-house" integration, several connectors were also developed to provide I.R.I.S. technology with the industry standard scanning platforms, such as Kodak Capture Software®, Kofax® Ascent Capture™ and eCopy ShareScan® OP.

To determine which version meets your specific needs, you can find your nearest reseller at www.irislink.com or contact us at var.opt@irislink.com

Add-On Modules

Language add-ons

IRISDocument™ offers you the possibility to add recognition of Traditional and Simplified Chinese, Korean, Japanese, Hebrew, Farsi and Arabic.

iHQC™ add-on

Take full advantage of the new compression technology with this page limited or unlimited add-on (see page 10 for more details).

Multi-CPU add-on

With the Multi-CPU Add-on, spread the workload across the dual/quad cores of your CPU(s) to make the best use of your computer and server resources.

For instance, with this Add-on and on a quad-core CPU, IRISDocument™ Server will process up to a stunning 60,000 images/day on a single PC!

SharePoint® add-on

IRISDocument™ connects directly to any Microsoft® SharePoint® server to export the generated documents with their indexes.

Moreover if you combine IRISDocument™ with IRISPowerscan™, the different document types and their indexes you defined prior to scanning will be exported to SharePoint® along with your files, automatically into the right document libraries, based on these document types.
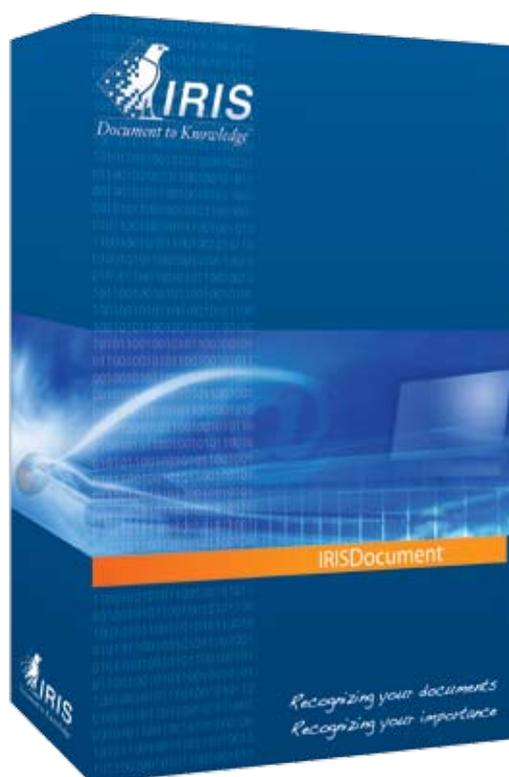
iHQC™: Intelligent High Quality Compression Technology

iHQC™ is the revolutionary hyper-compression technology from I.R.I.S., which will dramatically reduce the size of your PDF, PDF/A-1b and XPS files (color or black and white) while maintaining perfect image quality and text legibility. iHQC™ generates better quality than JPEG 2000 with much smaller files.

iHQC™ offers several levels of compression. IRISDocument™ comes standard with the basic one, for unlimited volumes of documents.
You will also enjoy the best levels of compression for 1,000 images/year without having to purchase the add-on.

Images	Color (Uncompressed)	Color compressed JPEG		Black & White TIFF Group 4		PDF+ iHQC™ Good Size (image + text)		PDF+ iHQC™ Good Quality (image + text)	
	Kbytes	Kb	compression	Kb	compression	Kb	compression	Kb	compression
byte_a	19.978	1.445	14	119	168	57	350	86	232
sojoumc_son	19.158	824	23	91	211	59	325	88	218
eos03noise	33.836	1.024	33	89	380	58	583	100	338
fnac01	7.795	474	16	30	260	25	312	35	223
lion_a	8.196	759	11	35	234	30	273	44	186
park_a	4.303	436	10	30	143	25	172	32	134
scenic_b	13.540	619	22	76	178	43	315	61	222
time_france	19.429	904	21	106	183	57	341	80	243
invoiceTDI	26.092	1.259	21	29	900	52	502	85	307
Total	126.236	7.744	16	605	209	406	311	611	207

This table is showing that iHQC™ allows to generate fully-searchable PDF, PDF/A-1b and XPS files (containing both the color image and the full-text index) that are more than 200 times smaller than the original color image and smaller than the black & white TIFF Group 4 image (which doesn't contain the full-text index).
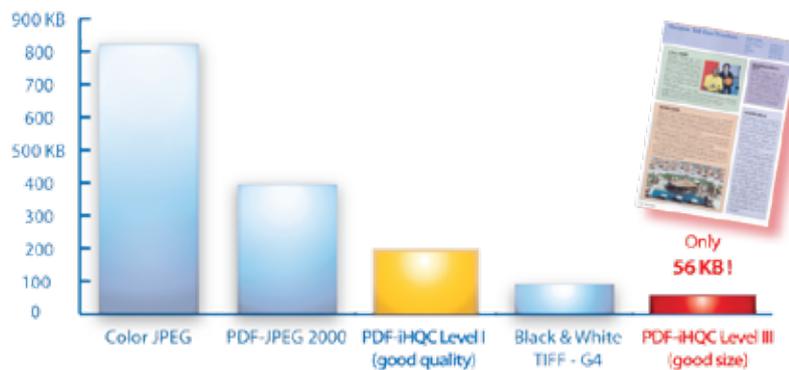
Available iHQC™ compression levels:

- iHQC™ - Level I – Good Quality - INCLUDED
- iHQC™ - Level I – Good Size
- iHQC™ - Level II – Good Quality
- iHQC™ - Level II – Good Size
- iHQC™ - Level III – Good Quality
- iHQC™ - Level III – Good Size

INCLUDED: 1.000 pages only.
Over 1.000 pages: Optional iHQC™ Add-on.
Prices are volume based. See I.R.I.S. price list.

System Requirements

Recommended

Intel® Pentium® IV 2 Ghz PC

1 GB RAM is recommended to process grayscale and color images

2 GB recommended when using the multi CPU add-on

USB port for hardware key

2 GB available disk space

Windows® 7 ready, Vista® (64 bits), XP (64 bits), Server® 2008 and 2003 (32 or 64 bits)

I.R.I.S. Company Profile

Image Recognition Integrated Systems (I.R.I.S.), a Belgian company founded in 1987 and listed on Euronext Brussels (IRI), is a leader in the "Document to Knowledge" market, and provides extremely high-quality solutions for converting paper documents into electronic formats for archiving, storing and sharing digital information.
For more information on our company, our solutions or our products, visit the I.R.I.S. website at www.irislink.com.

I.R.I.S. s.a.: Rue du Bosquet 10 • 1348 Louvain-la-Neuve • Belgium
Phone: +32-(0)10-45 13 64 • Fax: +32-(0)10-45 34 43
info@irislink.com • www.irislink.com

I.R.I.S. Inc.: Delray Office Plaza • 4731 West Atlantic Avenue • Suite B1 et B2 • Delray Beach, Floride 33445 - U.S.A.
Phone: +1-(561)-921 0847 • Fax: +1-(561)-921 0854
info@irisusa.com • www.irisusa.com

I.R.I.S. HK Ltd. : Unit 805-806, Hung Tai Industrial Building • 37-39 Hung To Road, Kwun Tong • Kowloon, Hong Kong - China
Phone: +852 22646133 • Fax: +852 23895733
info@irislink.com • www.irislink.com

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

I.R.I.S. Inc.: Delray Office Plaza
4731 West Atlantic Avenue
Suite B1 et B2
Delray Beach, Floride 33445
U.S.A.
Phone: +1-(561)-921 0847
Fax: +1-(561)-921 0854
info@irisusa.com
www.irisusa.com

I.R.I.S. HK Ltd. : Unit 805-806, Hung Tai Industrial Building
37-39 Hung To Road, Kwun Tong
Kowloon, Hong Kong
China
Phone: +852 22646133
Fax: +852 23895733
info@irislink.com
www.irislink.com

C090710-09