



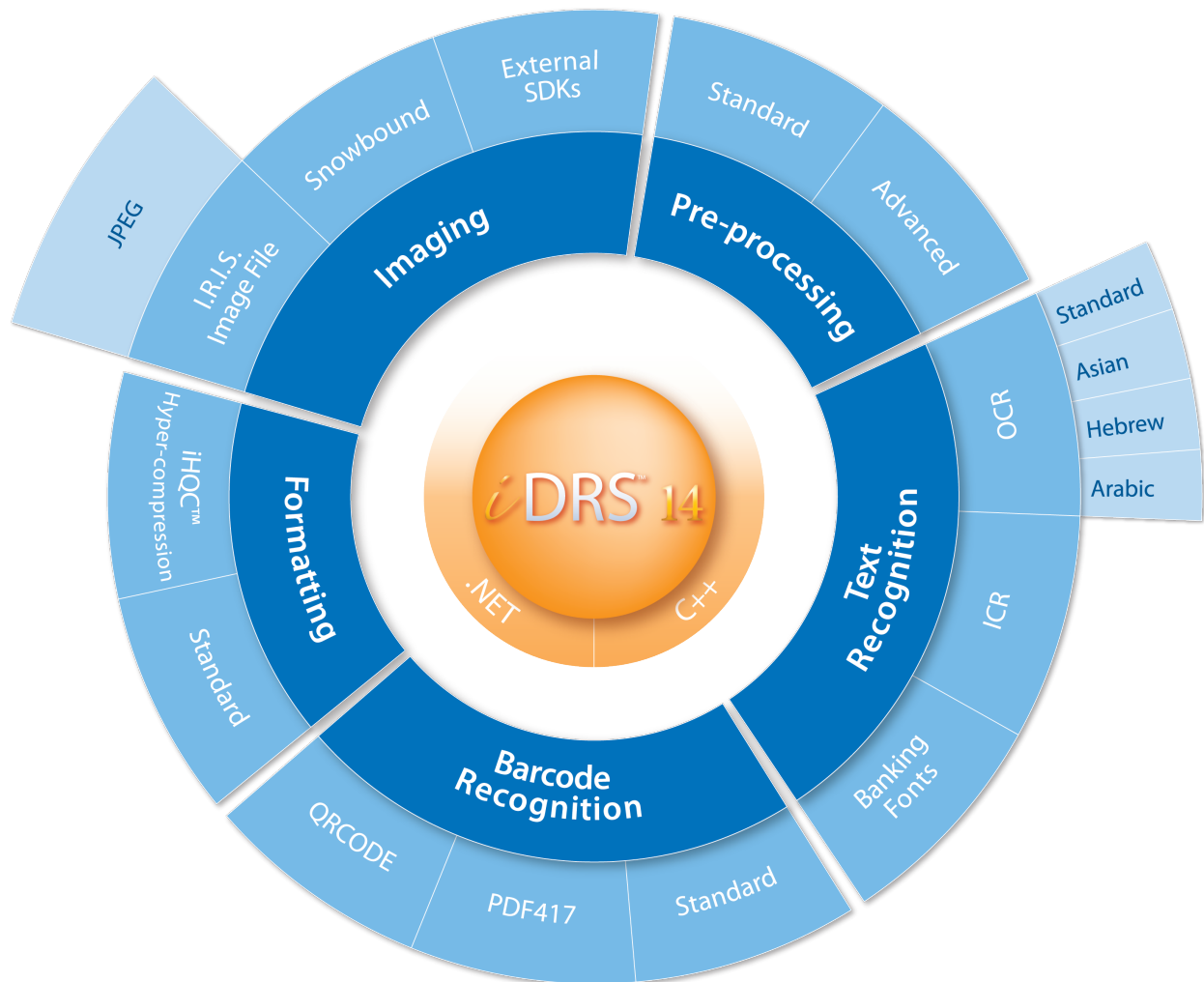
Datasheet

iDRS™ 14

You developed your own application.
You deserve recognition...



*Bring the power of recognition to your application
with more than 20 years of I.R.I.S. technology*



iDRS™ 14 is a compilation of more than 20 years of I.R.I.S. technology and expertise in data recognition and output. All proprietary I.R.I.S. software is based on iDRS™ - it is what we use!

iDRS™ 14 includes all our latest advances in **preprocessing** (deskew, despeckle, binarization, etc.), **recognition** (OCR, ICR, etc.), **formatting** (PDF, RTF, PDF/A-1b, XPS, etc.) and **hyper-compression** (iHQC™, an I.R.I.S. proprietary technology) and will allow developers to create their own specific solutions for the recognition of any kind of document. The toolkit is compatible with the most commonly used operating systems (Windows®, Linux®, UNIX®, Mac OS®) using a simple and complete C++ interface.

iDRS™ 14 is also available for .NET environments! allowing J#, C# and VB.NET developers to use iDRS™ 14.0 in Microsoft® Windows®.

iDRS™ 14 is a modular toolkit that will fit your project requirements perfectly. You just use what you need.

iDRS™ Global Benefits

iDRS™ 14 is the most competitive and fastest professional solution on the market to enable your application with the ability to convert all images into indexed and editable files, and/or provide multiple users with document recognition services.

iDRS™ is:

Fast

The recognition speed of the OCR engine is the fastest technology available on the market. iDRS™ technology embedded in I.R.I.S. products has been used to process hundreds of thousands of documents per day on a single machine!

Easy to install and use

An integrator can create a complete application in a few hours. Just have a look at the many included sample codes and you will be able to integrate all the iDRS™ technologies into your C++, C#, J# or VB.NET applications.

Demo applications are also available as examples showing how the features can be used and integrated.

Affordable

Not only does iDRS™ come without any page or volume limitation, but it is composed of modules which can be purchased separately – you only pay for what you need.

The pricing structure of iDRS™ is also flexible, and takes into account factors such as run-time, machine and even your business model (workflow, purpose, server/desktop application, heavy duty or not, etc.).

Versatile - Create your own solution, from entry-level OCR software to a complete document management platform!

Did you see something that caught your attention in one of our products and you would like to develop your own version? Or have you already taken care of the scanning and archiving parts but would like to throw OCR and indexing into the mix? In any case, iDRS™ is all you need.

Feature-rich

iDRS™ is packed with features which will be described later on in this datasheet, but here is a glimpse of what it can bring to your product line:

- Improve scanned images with powerful image cleaning tools.
- Deskew text, detect its orientation, and much more.
- Rename and separate documents automatically by detecting barcode values and blank pages.
- Powerful recognition engines for languages, barcodes, banking fonts, patch codes, etc.
- New OCR engine with 137 recognized languages - unlike our competitors, every language we recognize is fully supported, including a full dictionary.
- Index documents, and store the extracted data into a descriptive XML index file
- iHQC™ (hyper-compression technology by I.R.I.S.) provides a much higher compression rate than the best available TIFF G4 compression for black and white images.
However iHQC™ is much more than an image compression, it is a document compression technology that combines off-the-chart rates with perfectly preserved readability. (more details on page 8).
- Convert document images into many formats, among which PDF, XPS (with possible hyper-compression), etc. (see full list on page 7).
- Think about the future: give your customers the ability to convert documents to PDF/A-1b for long-term archiving.

A Closer Look at the Modules

iDRS™ is composed of different modules you can pick from to implement into your own application.

*1. Imaging module**

- JPEG extension

2. Pre-processing module

- Advanced extension

3. Text Recognition module

- Asian extension
- Hebrew extension
- Arabic extension
- Banking Fonts extension
- ICR extension

4. Barcode Recognition module

- PDF417 extension
- QRCODE extension

5. Formatting module

- iHQC™ extension

* The Imaging module can be replaced by:

- SNOWBOUND
- External SDKs

1) Imaging

The imaging module offers support for the compression and decompression of images in various formats. Decompression is to open various image file formats while compression is for embedding images inside your documents.

- **Standard**
This standard Imaging module allows you to open and save the TIFF (CCITT G4).
- **JPEG extension**
Allows you to open and save JPEG.
- **Snowbound, External SDKs**
iDRS™ can use external imaging toolkits to extend its imaging capabilities. They integrate seamlessly into iDRS™ in a “plug and play” fashion. Other external SDKs can also be used. For a full list of formats supported by those SDKs, please refer to their respective documentation.
- **Input file types**
 - TIFF CCITT IV monochrome I/O.
 - JPEG JFIF I/O.

2) Pre-processing

Provides image pre-processing capabilities allowing image enhancement to increase the accuracy and interpretability of the image resulting in optimal recognition results.

- **Standard pre-processing**
The standard pre-processing module offers several image enhancement techniques to improve the OCR result:
 - **Deskewing:** remove any skew of the text in the image up to an angle of 10.5°.
 - **Orientation detection:** Automatic rotation and straightening of the text orientation contained in your documents, by 90°, 180° or 270°.
 - **Despeckling:** remove noisy pixels with customizable cleaning tools.
 - **Inversed Despeckling:** remove white noise from a black background.
 - **Smoothing:** flatten out the colors of the image, which can for example remove the JPEG compression artifacts and help recognition.
 - **Adaptive and threshold based binarization.**
 - **Pixel inversion, etc.**
- **Advanced pre-processing**
The advanced pre-processing module offers more advanced pre-processing capabilities: color dropout after binarization, advanced despeckling, dark border removal, line removal and interpolation.

3) Text Recognition

iDRS™ enjoys the most productive optical recognition technology, and includes various recognition engines for machine print (OCR, MICR), handprint (ICR) and barcode recognition.

Recognition engine and full page OCR

iDRS™ includes the latest I.R.I.S. 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.

- **OCR**

The I.R.I.S. OCR engine currently supports 137 languages under which all Latin languages, Greek, Cyrillic, etc.

Add-ons offer recognition of Asian languages, Arabic, Hebrew and Farsi. (see below).

Next to pure recognition the OCR module also offers page analysis, access to the recognized data and formatting information, possibility to load user lexicons and many other recognition related functions such as:

- Page-layout analysis: automatic detection of text, graphic, table and barcode zones.
- Detailed recognition information: access to the character confidence level, alternative solutions, the coordinates of each character, etc.
- User Lexicon support: a user lexicon containing specific terms can be loaded into the engine. This is very useful for the recognition of very specific information like for example pharmaceutical names.
- Formatting information: paragraph attributes, typefaces, font size, typestyle, color, drop capital.
- Character recognition: recognition of ligatures ("joined" characters), broken and degraded characters (6 to 72 point size).
- Multilingual recognition: combine up to 5 languages on the same document (selective combinations).
- Dot Matrix.

- 137 recognized 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, Luxemburg, 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 add-on**

The Asian add-on extends the OCR languages with 4 additional languages: Traditional Chinese, Simplified Chinese, Japanese and Korean.

- **Hebrew add-on**

The Hebrew add-on extends the OCR languages with the Hebrew language.

- **Arabic & Farsi add-on**

Offers recognition of Arabic and Farsi.

- **ICR**

I.R.I.S. has developed its own Intelligent Character Recognition (ICR) engine for all Latin based languages. Supports spaced and even touching handprint.

- **Banking Fonts - MICR**

The following banking fonts can be recognized: OCR-A, OCR-B, E13B and CMC-7.

- **Define a User Lexicon.**

- **Fine-tune the Recognition Quality Vs. Speed ratio.**

- **Choose to apply recognition to selected pages only.**

- **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.
- Use the indexed information for file renaming.

4) Barcode Recognition

- **Barcodes**
 - **1D** - The barcode module will identify any barcode located anywhere on a page. 24 different 1-dimensional barcodes are supported covering the most popular ones as code 39, EAN, Code 128, etc.
 - **2D** - The I.R.I.S. barcode engine also supports the recognition of PDF 417 and QR barcodes. 2-dimensional barcodes containing much more information than the 1-dimensional barcodes.
 - **Full list:** 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.
- Fine-tune the **Recognition Vs. Speed** ratio.
- Choose to apply recognition to selected pages only.
- 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.
 - Use the indexed information for file renaming.

5) Formatting

The formatting module allows the developer to export the recognized data to various output formats. Several formatting options are available: retain formatting information or not, retain text color, set page size and many more:

- **Standard**

The following output formats are available in this module: PDF (text, image-text, text-image and image), PDF/A-1b, HTML, XML, RTF, TXT, UNICODE, UTF-8, ODT, WordML, SpreadsheetML and CSV.
- **Advanced**

This module offers additional compressed output formats like for example PDF hyper-compressed with iHQC™. (see page 8).
- **Full list of 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). (see page 8).
 - **XPS** (4 different file modes: Image, Image-Text, Text, Text-Image).
 - **XPS with iHQC™ hyper-compression.** (see page 8).
 - **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).
- The iDRS™ engine 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.

iHQC™: Intelligent High Quality Compression Technology

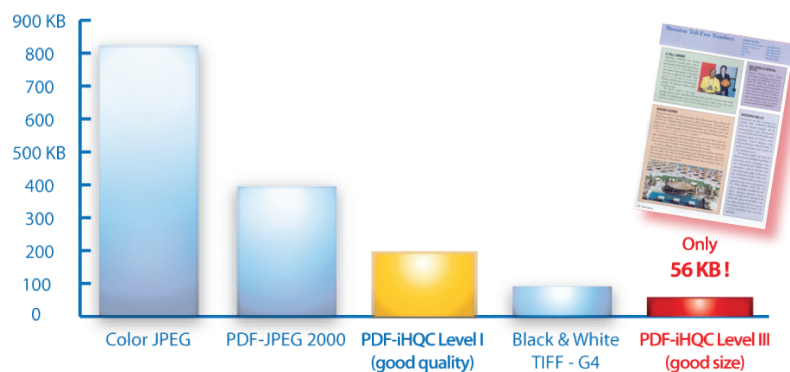
iHQC™ is the revolutionary hyper-compression technology from I.R.I.S., which will dramatically reduce the size of your PDF, PDF/A-1b and XPS files (color or black and white) while maintaining perfect image quality and text legibility. iHQC™ generates better quality than JPEG 2000 with much smaller files.

iHQC™ offers several levels of compression.

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

Example:

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).



A multi-platform toolkit

iDRS™ has been developed on a multitude of platforms (Windows®, Mac OS® and Linux®).

All have the same API, which means you could write your code once and simply recompile on the different platforms.

I.R.I.S. is working on extending the number of supported platforms. As their number is continuously growing, feel free to contact us for more details.

System Requirements and Platforms

Development Environments:

- Windows® - Microsoft® Visual Studio®
- Mac OS® – Xcode®
- Linux® - GCC

OS supported:

iDRS™ is built in such a way as to reduce OS dependencies to an absolute minimum. It works on all current desktop or server declinations of Windows®, Mac OS® and Linux® OS. This does not however cover embedded versions such as Windows® CE.

- Windows® 32 bit – Windows® 7, 2008, 2003, XP, 2000, NT, ME, 98, etc.
- Windows® 64 bit – (note the SDK is not native 64 bit)
- Mac OS® 32 bit
- Mac OS® 64 bit (note the SDK is native 64 bit) – Snow Leopard™
- Linux® 32 bit
- Linux® 64 bit

Miscellaneous

The size of the images the iDRS™ OCR engine is able to recognize are 4961 x 7016 or 5100 x 6600 pixels at all resolutions, which is equivalent to:

- A4 – 600 dpi
- A2 – 300 dpi
- A0 – 150 dpi

General Benefits of doing OEM business with I.R.I.S.

1. We own all the intellectual property of our technologies
 - We are commercially competitive as we don't need to pay licenses to anyone.
 - We have complete control of our technologies.
 - We provide long-term support.
 - We react quickly on issues, requests and suggestions.
 - We are more up-to-date with market trends.
 - We offer great flexibility since we have got all IP's and we can adapt to your needs.
2. We cover all the technology that you need
 - Avoid paying different license costs (different policies, protection mechanisms, limitations, etc.).
 - Avoid multiple negotiations resulting in time loss and cost ineffectiveness.
3. We offer simple integration (both for the integrator and the user).
4. We have many different ways/proposals to do OEM, from very quick and easy deployments to the most demanding integrations.
5. We provide close support.
6. We are in constant development to satisfy demanding market requests.
7. We are cross-platform (Windows®, Mac OS®, Linux®).
8. We propose a modular approach; you choose and buy only what you need.
9. Etc.

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

iDRS™ Global Benefits	3
A Closer Look at the Modules	4
iHQC™: Intelligent High Quality Compression Technology	8
A multi-platform toolkit	9
System Requirements and Platforms	9
Miscellaneous	9
General Benefits of doing OEM business with I.R.I.S.	10
I.R.I.S. Company Profile	11

© Copyright 2010 I.R.I.S. s.a./n.v.
 All rights reserved for all countries. I.R.I.S., I.R.I.S.' product names, I.R.I.S.' logos
 and I.R.I.S.' product logos are I.R.I.S. trademarks. All other products and names
 mentioned are trademarks or registered trademarks from their respective owners.

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