iText launches iText pdfOCR, a powerful open-source product enabling text recognition in scanned documents and conversion into editable PDFs

Tuesday 30 June, 2020

Ghent, Belgium – June 30, 2020. iText Group NV, a globally recognized thought-leader and innovator in PDF libraries and solutions, today announced the launch of iText pdfOCR, the newest addition to their award-winning software offering.

iText pdfOCR, which is part of the renowned iText 7 PDF SDK, offers Optical Character Recognition (OCR) functionality to convert printed text in scanned documents and images into a fully searchable PDF/A-3u compliant format (PDF version 1.7) and make accessing those texts easier and faster. Without machine-readable text, printed or scanned documents cannot be searched, indexed or interpreted. Logical follow-up actions could be data extraction with iText pdf2Data, secure content redaction with iText pdfSweep, or multilingual document recreation with iText pdfCalligraph. With repurposing data with the low-code document generator iText DITO® often being the final cherry on the cake.

The iText pdfOCR add-on is built on the Tesseract OCR engine technology. Tesseract supports over 100 languages and was originally developed by Hewlett-Packard (‘85), and was released under the Apache open source license in 2005. Since 2006, its development has been sponsored by Google.

“With COVID-19 urging companies to accelerate their digital transformation projects, organizations are forced to explore new ways of accessing and managing their data – existing and new. Being a leader in the digital documents space, we’re pleased to be at the forefront of this new era. As such, I am very proud to announce the latest addition to our PDF library for today’s new world: thanks to the OCR capabilities of iText pdfOCR many new opportunities will open up for users and enterprises that want to maximize their data potential.” Yeonsu Kim, CEO at iText Group NV stated.

“Staying true to our open-source roots, we’ve decided to build iText pdfOCR upon the open-source Tesseract OCR Engine. With this, we wish to reconfirm our positioning as an open-source company - a value which is appreciated by our millions of users and clients.”

“With this new addition to our PDF library, developers will now be able to leverage data locked away in documents which until now weren’t accessible. Our latest product enables them to enlarge their digital workflow capabilities by accessing the data buried in scanned files and deploy it for any action or purpose they or their end-user would like.” Tony Van den Zegel, VP of Products & Marketing at iText Group NV and General Manager at iText Software Belgium, said.

The applications of iText pdfOCR are various: for instance, archiving of historical documents, translations of legal documents, automatic data entry while processing all sorts of physical applications or claims, and sorting of otherwise not editable printed or scanned documents.

Please tune in for live demos on 9 July 2020. More information can be found here.

About iText

iText is a global leader in innovative PDF software. Its award-winning products are used by millions of users, both open source and commercial. The diverse customer base includes many of the Fortune 500 companies - ranging from technology, financial, travel to healthcare companies, as well as small companies and government agencies. Headquartered in Belgium, iText also has offices in Asia (Singapore and South-Korea) and in the USA (Boston).

www.itextpdf.com

For more information, please contact:

Karien Pype

iText Software

karien.pype@itextpdf.com

Media:



Related Sectors:

Business & Finance :: Computing & Telecoms :: Consumer Technology :: Government :: Manufacturing, Engineering & Energy :: Medical & Pharmaceutical :: Personal Finance :: Public Sector & Legal ::

Related Keywords:

PDF :: OCR :: IText :: SDK :: PDF Library :: Software :: Developer :: Document :: Java :: .Net :: Open Source :: Code :: API :: Workflow :: Archive :: Digital Signing ::

Scan Me:

+32 473 93 22 79

Press release

Company Contact:

iText Software

T. +32473932279
E. marketing@itextpdf.com
W. https://www.itextpdf.com/

Additional Contact(s):
Karien Pype, Global Lead Marketing / karien.pype@itextpdf.com
Rien Verbrugghe, Marketing Manager EMEA, APAC / rien.verbrugghe@itextpdf.com

View Online

Additional Assets:
https://itextpdf.com/en/products/itext-7/pdfocr
Press release - Russian
Press release - German
Press release - Korean
Press release - Dutch
Press release - French
Press release - Japanese

Newsroom: Visit our Newsroom for all the latest stories:
https://www.itext.pressat.co.uk