

ByteScout Improved OCR Features - Auto Repairing of Damaged Text in PDF, Support for Multiple Languages

ByteScout recently enhanced OCR features for damaged PDF text auto fixing & support for multilingual recognition in scanned PDF documents.



WILMINGTON, DE, UNITED STATES,

December 21, 2021 /EINPresswire.com/ -- ByteScout recently improved OCR features for auto repairing damaged text in PDF, support for multiple languages in scanned documents recognition, and others.

These new complex features enhance the experience for developers using ByteScout SDK and web components for data extraction and PDF processing.

ByteScout offers an exclusive [PDF Extractor functionality](#) for cloud and on-premise data extraction. ByteScout is well-known and trusted by TOP companies that continuously use its SDK and Web API to get unstructured data from scanned PDF documents or images. Many users have some number of documents with damaged text.

Damaged text is the text that appears OK but when you select and copy it you see some damaged symbols. This may be caused by some legacy apps or automated PDF generators.

Recently, ByteScout added a new OCR mode called Repair Fonts powered by AI. In this mode text on each page is checked automatically and if the text is damaged then the PDF Extractor engine restores the original text.

At the moment, the only English text is supported but we are happy to help customers to implement automatic text repair into the current workflow for other languages and for mixed languages PDFs.

This new feature is fully available in the cloud version ([PDF.co](#)) and on-premise versions of SDK and API Server. The latter is a self-hosted version available as a separate product called "[ByteScout API Server](#)" that can be easily deployed into a server in a private cloud or in an offline environment with minimal requirements for server hardware.

ByteScout SDK and API allow extracting structured data from unstructured or scanned documents, invoices, orders, statements, and other documents. They can accurately split and merge PDF files, convert PDF forms and tables into CSV, XML, Excel, TXT, search for text, add or remove security, fill and sign PDF forms.

ABOUT BYTESCOUT

ByteScout provides data extraction solutions for companies of every size from small businesses to Fortune 500 companies in the Insurance, Risk Management, and Banking industries since 2006. Offerings include on-demand API, on-premise Enterprise API Server, and low-level on-premise Software Development Kits (SDK). Enterprise customers are also provided with on-premise solutions ensuring secure and privacy-friendly data processing.

Media Relations

ByteScout, Inc.

+1 888-908-2357

[email us here](#)

Visit us on social media:

[Facebook](#)

[Twitter](#)

[LinkedIn](#)

This press release can be viewed online at: <https://www.einpresswire.com/article/558928473>

EIN Presswire's priority is source transparency. We do not allow opaque clients, and our editors try to be careful about weeding out false and misleading content. As a user, if you see something we have missed, please do bring it to our attention. Your help is welcome. EIN Presswire, Everyone's Internet News Presswire™, tries to define some of the boundaries that are reasonable in today's world. Please see our Editorial Guidelines for more information.

© 1995-2021 IPD Group, Inc. All Right Reserved.