New PDF Tools, Data Extraction and Validation from External Data Sources, and Microform Processing Coming to Grooper Version 2.8

OKLAHOMA CITY, SEPTEMBER 17, 2019 – Grooper, the leading software platform for discovering unstructured data and turning it into actionable information, announces the release of version 2.8. New capabilities include automated generation of PDF document elements like signature, checkbox, and textbox widgets, as well as better shade, shape and scratch removal. The new version also features improved Optical Character Recognition (OCR) and PDF document text extraction, advanced microfilm processing, easier integration with outside data sources, and a new visual studio integration.

New PDF Elements

Grooper's ability to produce PDF documents gains even greater versatility with 2.8. The following PDF elements have been added:

- Highlighting – edit border style, color, and transparency
- Signature widgets
- Checkbox widgets
- Textbox widgets
- Radio group widgets
- Position widgets relative to any data element in the document's content model

New Recognize Activity

In previous software versions there were two separate Activities for OCR and PDF text extraction. These two activities have been combined into a single Activity called Recognize. This new activity uses computer vision to understand semantic content and enables users to target an entire document layout, including text, barcodes, lines, checkboxes, and other layout information. Recognize processes PDFs that contain images and native text.

4 New Image Processing Commands

1. Shade Removal – removes shaded headers
2. Shape Removal – improves OCR by effectively removing shapes
3. Scratch Removal – detects / removes / repairs scratches on scanned images
4. Extract Page – processes and repairs images taken by phone cameras

Advances in Microfilm Processing

Grooper enables complete microform processing:

- Process all forms of microfilm and microfiche
- Fast and accurate digitization
- Correct image warping, scratches, and other defects
- Perform OCR on digitized documents
- Validate extracted data with information on external databases
- Verify and correct original document indexing by grouping documents based on content

Better Data Integration

Grooper version 2.8 builds upon a history of data integration tools and techniques with five updated features:

1. New connectors for outside data sources
2. Expression-based field mapping on import and export
3. Expanded database lookup capabilities using SQL query formatting
4. Robust transcript, invoice, and inventory processing
5. Validate and populate data based on external database information

New and Improved Visual Studio Integration

Grooper's visual studio integration has been re-built with the following improvements:

- Integration with any version of Visual Studio from 2012 forward
- No dependency on VSTA (Visual Studio Tools for Applications)
- Supports C#
- Templates for common Grooper Objects and Activities
- Generate custom reports
- Create custom Activities that allow REGEX to run against batches and flag matches
- Run live queries against databases when validating extracted data

Learn more about Grooper – visit www.grooper.com.

About Grooper

Grooper enables rapid innovation for organizations processing and integrating large quantities of unstructured data. Created by a team of courageous developers frustrated by limitations in existing solutions, Grooper is an intelligent document and digital data integration platform. Grooper leverages patented and sophisticated capture technology, image processing, OCR, machine learning, computer vision, and natural language processing.

Grooper – available, accessible, and transparent artificial intelligence.

###