Kodak

Info Input Solution

Release Notes
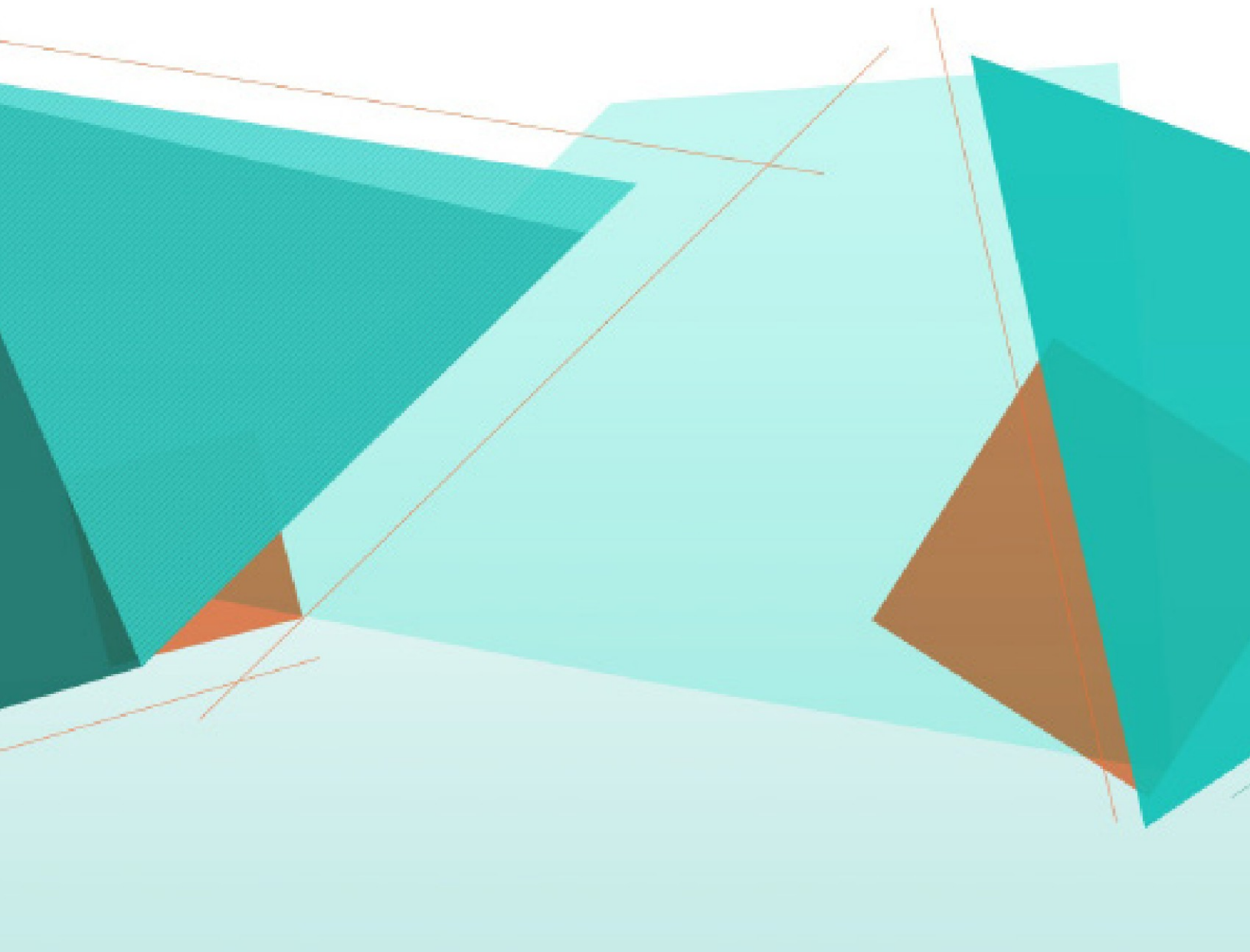
Contents

1 New Features / Enhancements

Info Input Solution is a web-based scanning software solution that enables users to scan documents from any remote location and deliver the captured information to an Enterprise Content Management system, for storage or further processing.

Info Input Solution v6.1.7:

HTML Client:

- Language files updated.

Extraction:

- Add support for `achine nly` flag and `Layout UUID` in the HyperScience extraction profile configuration.

Authentication Service:

- Add support for custom authorization logic after successful user authentication.

Info Input Solution v6.1.6:

Extraction engine:

- Added an `Auto-create mappings` option for intelligent OCR Extraction step.

Indexing:

- Added global navigation keyboard shortcuts that can be used during Indexing.

Scripting:

- Add API method `getCurrentScanProfile()`.

Export Service:

- Added Export Service parameter to process only specific Job(s).
- SharePoint Online Export: Added support for application-level authentication and more SPOnline metadata types.
- Documentum DFC Export: Added configuration option for custom Folder Type, other than the default dm_folder.
- Documentum DFC Export: Added return property for the exported DocObjectId to be used by next Export configuration steps.
- Filenet P8 Export: Added the ability to export files as "Unfiled Documents".

Info Input Solution v6.1.3:

Import Service:

- Added context-sensitive search capability to the Index Class / Form Type combo box.

Scan/Import:

- Added XFAF Support for processing PDF with XFA forms as images.

Export Service:

- Scripting Export: Added PDF utility function to split the output PDF documents by size (output file size

threshold).

Extraction engine:
- Added support for using Microsoft Form Recognizer and Hyperscience engine as OCR Extraction profiles.

Info Input Solution v6.1.2:

Indexing:
- Added context-sensitive search capability to the Index Class / Form Type combo box.

Import Service:
- Added Kodak INfuse file source option.

Export plugins:
- SharePoint Online: Added support for dynamically creating folders in SharePoint Online.
- SharePoint Online: Added Remote Procedure Call method `getFolders()`. This method can be used to fetch the list of available target Folders during Indexing.
- Documentum DFC: Added an extra configuration parameter/field to define Folder Type of the Folder object in DCTM.

Info Input Solution v6.1.1:

Extraction:
- Added support for Hyperscience engine in Intelligent Extraction Workflow step.

HTML Client:
- Added the Global Script function `ready()` to HTML Client scripting runtime, triggered once the Client application has been successfully initialized and all required data has been loaded.

Info Input Solution v6.1:

Extraction:
- Added new Intelligent Extraction Workflow Step, which supports Google DocumentAI and Microsoft Azure Form Recognizer.
- Added support for Google DocumentAI and Microsoft Azure Form Recognizer as OCR Extraction engines.

HTML Client:
- Updated language localization. Supported languages are: English, Arabic, Chinese, French, German, Japanese, Polish, Spanish, Turkish.
- Enhanced the Batch Manager with `Filter` and `Group by` features. An additional license is required for this HTML Client-specific enhancement.
- Advanced Indexing Form: Added Form Script methods to run Database Lookup queries.

Scan/Import:
- Added support for processing PDF with XFA-F forms as images.

Indexing:
- Added the ability to specify on an Index step whether the operator is allowed to change the assigned Index

Class / Form Type. The Index Class / Form Type combo box will be disabled, provided that the current value is not null.

- Added support for sorting or filtering the Index Class / Form Type drop-down list. Added function `customizeAvailableTypes()` to the Indexing Workflow step configuration script. This function takes a list of all available Form Types or Index Classes and returns an array with the desired items in the right order.

Authentication Service:
- Added support for SAML identity management protocol.

Job Workflow:
- Added Fork/Join Job Workflow design logic for Document or Folder level processing

Export Service:
- Images Export: Added support for setting PDF properties and bookmarks to the output PDF documents.
- Images Export: Added support for running Images Export for the whole batch or per Folder node.
- Images Export: Added support for creating searchable PDF documents using the OCR results from Extraction Workflow step.

2 Fixed Issues

Info Input Solution v6.1.7:

Oracle Security Alert for CVE-2021-44228 log4j2 vulnerability
- Update all log4j2 libraries to latest non-vulnerable version.
  Please refer to Patch Deployment instructions section, library files update step.

HTML Client:
- Db Lookup action is not executed correctly when hitting Enter.
- Primary key violation during scan when using `IPage.detectBarcodes()` in `pageArrived()` without `.awaitUploaded()`.
- Batch can be closed at Indexing step, when using Table fields with required index fields but still empty (no rows).
- Updated the variables selector for Create Batch Copy action in Batch Manager.
- Login error messages from HTML Service show in bad format.
- HTML Client user session re-login (automatic session recreation) not successful after a communication lost event between the HTML Service and Core Service.

Thick Client:
- The digest script does not run if digest.txt is corrupted (or empty).

Authentication Service:
- User login error when the payload that comes from authentication service is too big.
  Added a new configuration named `principal-filter` which can be filled with a regex value to filter the principals (not only groups) provided by the identity provider.

Job Setup:
- IndexZone Test Dialog throws Null Pointer Exception when a dependent zone is sent without the "use from" zone.

Extraction:
- File not found error in Intelligent OCR Extraction step when the Save Payload is selected. This is reproducible on Linux hosted app server, only.
- Google Document AI fails when processing a document with more than 10 pages.
- Hyperscience integration code fails with `DateTimeParseException: Text '' could not be parsed at index 0`

Import Service:
- The Import Administration does not report to the user or the log, if it is not possible to save configuration changes to a config file.

Export Service:
- CMIS Export: GC Overhead limit when the target CMIS directory structure (Destination Folder) is requested at Job Setup configuration step.

Database:
- Manual DB upgrade using ant script db-config.xml fails for SQL Server target DB with user permissions error.

Info Input Solution v6.1.6:

Client:
- Open Batch / Select Batch to Index dialogs don't show the batches created by current User in bold.

Indexing:
- The arrow navigation buttons don't update the Indexing Viewer properly when using some custom Indexing scripting logic.

Job Workflow:
- Unable to delete a Document node after Fork step, if the Document has been already processed for an Export Workflow step.

HTML Client:
- Image Operations are not available in HTML Client scripting.
- When the promise from method `IPage.awaitRendition()` is fulfilled, the image may not be available for further processing.
    - Batch Level index fields are displayed, when a forked Folder/Document is loaded for Indexing.

HTML Service:
- Cluster application restart failure with `GlobalStorageResource initialization error` when all cluster nodes are restarted at the same time.
- Prevent the HTML Service from starting if `GlobalStorageResource` fails to create.

Extraction engine:
- Image Registration is not performed in Extraction step, when the OCR data from a previous intelligent OCR Extraction workflow step are used.
- An Intelligent OCR Extraction profile fails to process a document when used in Form Type Extraction zone.
- A Source field dropdown list in Mapping configuration for an Intelligent OCR Extraction step cannot be expanded using the arrow button.
- Google Document AI Invoice OCR processing fails with `java.lang.IndexOutOfBoundsException: Index: 0` error.

Export Service:
- SharePoint Online Export: improvements for upload error handling and subfolder creations.
- Filenet P8 Export: cannot upload to root folder or connect over SSL encryption.
- Filenet P8 Export: cannot fetch document classes names for document subclass levels (paths with more than 4 levels).
- Documentum DFC Export: Index fields mappings does not show all DCTM fields/properties.

Installer:
- Azure DB configuration scripts added.

- Database upgrade error: `Could not parse mapping document from resource event.hbm.xml`

Info Input Solution v6.1.4:

Export Service:
- CMIS export: Error while retrieving CMIS repositories.

Info Input Solution v6.1.3:

HTML Client:
- Batch Manager: Group by "Created" or "Locked on" column displays a wrong group title.
- The required Index fields, when empty, should show with red border.
- The word 'optional' is appended to the default CL-HTML indexing label for non-required Index fields.
- Ranger scan profile: The HTML Client fails to launch when the use_check_scanning parameter is enabled.
- Blank page detection during Scan with CCT (when enabled in Job Setup separation rules) does not save the result as page property (`page.isBlank()` returns undefined).
- ScanService enumeration is not available in HTML Client scripting.

HTML Service:
- The HTML Service may randomly fail to create the image thumbnails from an uploaded multi-page PDF file, only when running on Linux, with error: `Failed to get thumbnail from server: Failed to open (code 2) PDF file`.
- The Custom Indexing Form may fail to load when moving quickly between Indexing nodes that trigger Database Actions.

Scan/Import:
- Some digitally signed PDF documents fail to be imported.

Indexing:
- Batch displayed in multi-page View mode when opened from Select Next Indexing Task or Get Next Task to Index dialogs.

Extraction engine:
- Extraction processing step fails with `java.lang.IllegalArgumentException: Illegal bounding box` error caused by invalid Image Registration.

Import Service:
- Email Import fails to process txt or html files with `java.io.IOException: File type is not specified`
- Kodak INfuse: Import Administration Utility error when reloading a Kodak Infuse file source configuration.
- Kodak INfuse: "Test connection..." using invalid credentials will log the error to import service log file but doesn't notify the Job administrator.

Export Service:
- Documentum DFC Export: Cannot map index fields for Double and Integer field types.

Info Input Solution v6.1.2:

HTML Client:
- The scanned Pages may not be added to the Batch structure in the correct order, when scanning with CCT Scan Profile and also using the pageArrived() hook function in asynchronous mode (using Promises with callback functions) for blank page detection/deletion.
- When using the Next>> action button to proceed to the next Indexing node that requires Indexing, the Single Page Viewer window is not refreshed accordingly.
- When using the Batch Tree to select an Indexing node, the Single Page Viewer window may not be refreshed accordingly.
- When scanning with CCT Scan Profile and the Job uses Extraction profile for Barcode detection, the Extraction results are not always sorted and saved in the same order.
- A row selection in Batch Manager window is lost when refreshing the Batch Manager content.

Thick Client:
- User login with Windows integration fails with security exception.

Scripting:
- Page.isBlank property is not updated after an execution of the page.detectIsBlank() method in HTML Client scripting runtime.
- The Page.getRectX() and getRectY() methods in HTML Client scripting runtime return absolute coordinates, instead of relative to the Page width.
- The objects returned by function Page.getBarcodeData() in HTML Client scripting runtime are expressed in absolute coordinates and size.
  After this fix, only if the global parameter use_relative_page_coordinates_and_point_size is set to false, they will be expressed in absolute coordinates and size.

Core Service:
- The Core service stops logging to the application log file after adding a third party JDBC driver.
- Batch Audit information is not logged to tables StatsBatchSessions and StatsImportedItems in the Database.

Export Service:
- An Images Export processing task returns with error when the "Skip first page" configuration option is enabled and the working (document) node has only one page (or is empty).

Import Service:
- File Import source fails to process PDF or TIF files with error message: java.io.IOException: File type is not specified.
- Email body inline images are not imported to the batch when the email also contains at least one attachment.

Extraction:

- Rotating a page with highlighted index zone from an intelligent Extraction processing step, does not update the zone in the GUI.
- When duplicating a batch in Batch Manager after an intelligent Extraction processing step, the batch fails to be loaded for Indexing step.

Info Input Solution v6.1.1:

HTML Client:
- Chrome browser only: Local Storage / IndexedDB is not cleared periodically, which can potentially result in browser crash after long HTML Client sessions.
- The return value `Const.AbortOp` in `pageScanned()` function does not stop the scanning procedure.
- Switching between a CCT Scan Profile and a Folder (Default) Scan Profile may cause the CCT Scan Profile options to be disabled.

Export Service:
- The Images Export "Add auto-annotation" configuration option fails for PDF output file type..
- Images Export fails to create searchable PDF using the Nuance OCR engine at Export runtime.

Job Setup:
- Under specific conditions the PDF Options menu for Images Export configuration does not open, when the Add auto-annotation option is also enabled.
- Authentication (testing) of the Extraction profile is always successful for the Microsoft Form Recognizer OCR Extraction engine option.

3 Patch Deployment Instructions

The cumulative patch 6.1.7 can be deployed on top of any v6.1 patch level Application Server.

Steps for applying the patch:

a) Stop the application:

Windows: Use the Administration Utility (Services tab) to stop all running Windows Services, in the following order. Finally, close the Administration Utility window.

Linux: Stop the server-side processes, using the stop scripts or System V init-compatible scripts provided, in the following order.

1. Info Input Solution Import Server (if installed)
2. Info Input Solution Export (Release) Server
3. Info Input Solution Apache Tomcat Server

b) Take a backup copy of the installation directory.

c) Copy the patch distribution archive to the application server and extract the patch distribution archive to the installation directory. This step will overwrite some files and folders of the existing installation. When prompted, the user should confirm file (folder) replace.

d) Some library (jar) files have been updated to latest version with Cumulative Patch 7, so the old version of these files must be removed from the installation directories.
We provide the libs-cleanup.bat file (for Windows) and libs-cleanup.sh (for Linux) under <root_installation>\tools directory.
Windows environment: Open a new Command Prompt window with administrator permissions, change directory to <root_installation>\tools, run the libs-cleanup.bat file and follow the steps to verify the deletion.
Linux environment: Change directory to <root_installation>\tools, make sure the shell script file is executable (run chmod +x libs-cleanup.sh), run the cleanup file (./libs-cleanup.sh) and follow the steps to verify the deletion.

The deprecated jar files that must be deleted after patch deployment are the following:
1st set: fontbox-2.0.20.jar, pdfbox-2.0.20.jar, xmpbox-2.0.20.jar from folders
   <root_installation>\client\lib
   <root_installation>\importd\lib
   <root_installation>\released\lib
   <root_installation>\server\WEB-INF\lib
   <root_installation>\service-html\WEB-INF\lib
2nd set: log4j-core-2.13.0.jar, log4j-api-2.13.0.jar, log4j-1.2-api-2.13.0.jar from folders
   <root_installation>\client\lib
   <root_installation>\importd\lib
   <root_installation>\released\lib
   <root_installation>\server\WEB-INF\lib
   <root_installation>\service-html\WEB-INF\lib
   <root_installation>\tools\lib
3rd set: log4j-core-2.13.0.jar, log4j-api-2.13.3.jar, log4j-1.2-api-2.13.0.jar, log4j-slf4j-impl-2.13.0.jar from folder

          <root_installation>\auth-service\WEB-INF\lib

e)   This patch may require a Database Schema upgrade.
       Start the Administration Utility: at startup, it automatically checks the active Database configuration and performs the required Database Schema upgrade.
       Note that, after a Database Schema upgrade, importing Jobs from a previous patch level is not supported.

f)   Start the application: start the Windows Services / server processes in the reverse order.
       RHEL: Start the server-side processes, using the start scripts or System V init-compatible scripts provided, in the following order.

1. Info Input Solution Apache Tomcat Server

2. Info Input Solution Export (Release) Server

3. Info Input Solution Import Server (if installed)

4 System & Environment Requirements

System requirements for the Info Input Solution Application Server:

CPU:

- Minimum requirement is 4 cores at 2GHz.

- For high volume systems, the recommended configuration is 12 cores @2.5+ GHz, or 12 cores of older / slower CPUs.

- When the Info Input Solution HTML Client is used by end-users, additional CPU is required for the HTML supporting service/process. The hardware specs depend on the use case and estimated load.

- When the Info Input Solution Classification and Extraction server-side engines are used in the Job Workflow(s), additional CPU is required for the Info Input Solution Core service/process. The hardware specs depend on the use case and estimated load.

- For virtual environments, features like Intel VT-x / VT-d and AMD V / Vi must be supported by the processors, activated on the machine, enabled in the Virtualization product and activated for the virtual machine.

Memory:

- Minimum memory size is 4 GB.

- Recommended memory size for high volume systems is 12 GB.

Operating System:

- Windows Server 2012 / 2016 / 2019.

- Red Hat Enterprise Linux 7 / 8.

- SUSE Linux Enterprise Server 12 / 15.

Disk Storage:

- Single server installation: The recommended configuration is a dedicated disk array (RAID 5 or 10) either physically installed on the server, or exported over a dedicated high-speed Storage Area Network (SAN). The minimum configuration is a single disk, again either physically installed in the server, or exported over a SAN.

- Cluster installation: The recommended configuration is a dedicated Network Attached Storage (NAS) unit, using a file system native to the OS of the application servers, connected to the server over a dedicated high-speed network. The minimum configuration is a NAS that can be used by all nodes in the server cluster.

- SMB v3 compatible Windows Server file storage systems are supported.

Network:

- Single server: The recommended configuration is separate networks for HTTP, Database and storage traffic, if external storage is used. All networks should be at least 1 Gb and all nodes should be attached to the networks through switches. The minimum configuration for low volume systems is a single 1 Gb network.

- Server cluster: In addition to the single-server recommended configuration, the recommended configuration for a cluster is to provide additional networks for cluster heartbeat and reverse proxy to cluster node HTTP traffic. The minimum configuration for low volume systems is a single 1 Gb network.

Software requirements for the Info Input Solution Application Server:

Apache Tomcat: version 8.5.68 64-bit (included in the installation package).

Database:
- SQL Server 2014 / 2016 / 2017 / 2019.
- Oracle 11.2 / 12.1 / 18c / 19c.

Bundled Database in installer:

If during installation, you choose to install a new instance of SQL Server 2019 Express, the operating system must meet the minimum requirements:

- Operating System: Windows 10 TH1 1507 or greater and Windows Server 2016 or greater.

For more information, see Hardware and Software Requirements for installing SQL Server 2019 at the Microsoft website.

System requirements for user workstations running Info Input Solution HTML Client:

- Operating Systems: All supported Windows OS versions (Windows 8.1, 10, Server 2012, 2016 and 2019). All supported Mac OS versions.
- All contemporary web browsers are supported: Google Chrome v40+, MS Edge, MS Internet Explorer v11, Mozilla Firefox 40+, Mozilla Firefox ESR 40+, Apple Safari (on Mac OS).

System requirements for user workstations running Info Input Solution Client:

- Operating Systems: All supported Windows OS versions (Windows 8.1, 10, Server 2012, 2016 and 2019).
- All contemporary web browsers are supported: Google Chrome, MS Edge, MS Internet Explorer v11, Mozilla Firefox, Mozilla Firefox ESR.
- The Info Input Solution Client application is also offered through web browsers that support java plug in: MS Internet Explorer (v8 up to v11), Mozilla Firefox (v3.6 up to v51), Mozilla Firefox ESR (up to v52). Java Runtime Environment version 8, 32-bit and 64-bit, is also required in this case.

Kodak

Info Input Solution