

Intella Connect 2.1 Release Notes

Highlights

- Added new cloud sources: **Dropbox, Office 365 (incl. OneDrive), SharePoint, Gmail.**
- **Email threads** are now detected and visualized. This includes the determination of the **inclusive emails**: together these cover all the content in the thread. This can reduce review time and effort. **Missing emails** are highlighted in the thread.
- Added an **integrated OCR** option. All Intella users can now OCR documents and images without requiring additional software, licenses, or systems.
- Items can also be OCR-ed by a connection to an **ABBYY Recognition Server** or by using a manual import and export method.
- Completely redesigned **user interface for importing load files.**
- **Custom columns** let one extend Intella's data model with new columns, populated by load file columns.
- A **Connect Grid** allows for a single point of access when multiple Connect servers are used.
- Added **regular expression**-based detection of text patterns, e.g. bank account numbers. A **Regular Expression Assistant** is included for constructing the expressions, together with a library of example expressions.
- Improved the **presentation of instant messages** by bundling them in day-to-day conversation items.
- Added **recovery of deleted files** in NTFS disk images using the MFT.
- Added support for the **Ext4** file system.
- Added functionality for **removing sources** from a case.
- Added support for indexing non-encrypted **iTunes backups.**

General

- Added a mechanism for dumping server diagnostics and usage information to a file. This can aid technical support as it reduces the need for submitting entire log files.
- The dongle ID, dongle type and related licensing information is now logged.
- Several performance and stability improvements in reading and writing case files, for both local and network file systems.
- Resolved an issue with browser connections closing prematurely due to unnecessarily strict time-outs.
- Resolved an issue with users logging out and not being redirected to the login page.
- Improved temp folder management when indexing.
- Changed the log file names to clarify the distinction between the logs produced by Intella, Intella Connect and Intella Node, as well as the distinction between case logs and server logs.

Connect Grid

- Connect servers can now be organized in a grid. In this setup, when a user logs in on the Connect dashboard, both the local Connect server and the other Connect servers in the grid will be queried for the cases that the user is authorized to access. This allows for a single

point of access for all reviewers in an organization, regardless of where the cases are stored; reviewers do not need to be aware which Connect server is hosting the case. For this to work, all Connect servers involved need to use the same LDAP instance.

Authentication and Authorization

- Made the availability of instant messaging in the Connect UI controlled by a permission.
- Made the LDAP connector more forgiving towards incorrect LDAP setups, e.g. an LDAP group that has missing entries.
- Resolved an issue with passwords containing non-Latin characters not being processed correctly.
- The size of the LDAP cache has been made configurable through a system preference.
- Resolved an issue with the permissions of a role remaining visible after the role was deleted.

Case Management

- The Add Existing Case page now lets the case admin browse the Connect server's local file system to locate the case.xml file.
- Cases that can only be opened in "review only" mode are now labeled as such in the Cases list.
- Improved the source configuration view to cover all details of the configuration.
- Improved folder validation when creating a new case and assigning a case folder to it.
- Resolved an issue with accounts with a case creation permission not being able to retrieve the list of connected Intella Nodes.
- Resolved an issue with the Cases page getting frozen when the sources of multiple cases were being edited simultaneously.
- Resolved an issue with reviewers getting HTTP 404 errors when accessing a shared case that had not been used for a long time.

Indexing - General

- Sources can now be removed from a case.
- Added support for indexing non-encrypted iTunes backups. This was tested on iTunes 12 with a variety of iOS versions. Other iTunes versions are being tested.
- Item IDs now stay the same when re-indexing a case.
- Many stability and performance improvements for indexing PDF documents. This results in more and better extracted text and images, faster extraction times and improved resilience to broken data.
- Improved indexing speed on large SQLite files.
- Improved the processing of TNEF attachments (winmail.dat files).
- Added detection of Apple icon (.icns), Radiance High Dynamic Range RGBE Format (.hdr) and DjVu (.djvu, .dvj) images.
- One can now edit the set of mailboxes that is extracted from an EDB file.
- Geolocation references embedded in Google Maps URLs are now extracted.
- Improved paragraph hashing, resulting in better detection of duplicate paragraphs. Several search features benefit from this.
- Improved modeling and normalization of the sender and receiver information of instant messages and phone calls.

- Resolved a concurrency issue that could occur when indexing PST files.
- Improved type identification of Bloomberg plain text documents.
- Improved processing of hierarchical generic Notes documents.
- Improved determination of the Source IP address of emails.
- The NSF document UID is now logged before processing that NSF item. This can help diagnose NSF items that fail to index.
- Improved resilience for NSF files containing items on which Notes crashes.
- Resolved an issue with MBOX and EML files originating from MacOS platforms not indexing correctly due to their end of line encodings.
- Resolved an issue with the longitude and latitude properties of items in an XRY phone dump not being extracted.
- Resolved an issue with the sorting on primary or family dates that could produce an incorrect sort order when the primary date preferences were changed and the user canceled the subsequent recalculation of these dates.
- Resolved an issue with incorrectly processed MMS messages in XRY phone reports.

Indexing – Disk Images

- Intella can now extract deleted items from disk images. File recovery is currently restricted to NTFS file systems and is based on traces of the deleted files found in the Master File Table (MFT). Intella tries to recover as much as possible of the file content and metadata. Whether a full or even partial recovery is possible depends on how the disk was used after the file deletion. Note that this functionality does not scan the unallocated space or slack space.
- Added support for the Ext4 file system.
- Added support for MacQuisition disk images (IMG format).
- Added disk image validation. When adding a disk image source, this verifies that the files of a multi-part image are in fact belonging to the same image, whether the file system found in the disk image is supported, etc.
- Cellphone reports and IBM Sametime dumps can now also be indexed when they are contained in a disk image. Previously, they had to be present in the local file system.
- Resolved an issue with indexing file system roots in AD1 disk images.
- Resolved an issue with the indexing of disk images mounted as a virtual drive with EnCase.
- Various stability improvements for indexing disk images containing an NTFS file systems.
- Resolved an issue with DD disk images consisting of more than 99 parts that would not index properly.
- Resolved an issue with the Find Parts function producing duplicate entries.

Indexing – Load Files

- This release offers a completely redesigned user interface for importing load files and load file overlays, bringing it in line with the Desktop edition. The new user interface offers more control options on how the load file is imported, shows a live preview of how the data is parsed, and gives better validation and error messages.
- When importing a load file, it is now possible to define custom columns. This allows one to extend the data model of an Intella case with new columns. The new columns are populated with selected columns from the load file, allowing any type of load file to be imported fully

into an Intella case.

Custom columns are typed, e.g. as a string, number, or date. This ensures a proper sort order when sorting on that column. The data in custom columns is generally searchable using keyword search. Custom columns that use the date type can be found and queried in the Date facet.

- The unit of the Size column can now be bytes, kilobytes, megabytes or gigabytes.
- Improved importing speed through better utilization of CPU cores.
- Improved the importing speed of load files containing custodian columns.
- When processing the content of binary items bundled with a load file, one can now specify the same configuration options as when adding a source, e.g. whether item recovery should be used, whether archives should be expanded, etc.
- The import configuration is now logged when importing a load file.
- Items with empty content (zero byte files) no longer get a message hash.
- Resolved an issue with documents getting typed incorrectly when they had an incorrect file extension and the binaries were available for proper file type detection.
- Resolved an issue with load files containing MSG files getting incorrect message hashes.

Indexing – Cloud Sources

- Added support for indexing Dropbox accounts. Both personal Dropbox accounts and Dropbox for Business accounts can be indexed. File versions can be extracted and are presented chronologically.
- Added support for indexing Office 365 accounts. Retrieved information includes users, user groups, emails, attachments, folders, instant messages and files stored in the associated OneDrive. Intella can also connect to the cloud SharePoint instance associated with the Office 365 account.
- Added support for indexing SharePoint instances. Retrieved information includes basic site and owner information, users, hierarchically nested sites, conversations, posts, attachments, document libraries, and files and folders from the document libraries. Both standalone and on-premise instances can be indexed, as well as instances hosted in the cloud as Office 365 SharePoint services. The following authentication protocols are supported: OAuth2 (for cloud instances), Kerberos, NTLM and basic authentication.
- Added support for indexing Gmail mail accounts. The retrieved information can optionally be restricted to a specific date range. Benefits of using the Gmail connector over the generic IMAP connector are: better performance, more accurate data representation (e.g. folders vs. Gmail's Labels, threads), and a read-only data connection ensuring that no data is altered.
- All cloud connectors support multi-threaded access to the cloud source to make use of the provided bandwidth as best as possible. All cloud connectors feature automatic retry logic, in case retrievals fail due to intermittent network problems or when exceeding the service's throttling policy limits.
- Improved error messages when connecting to an IMAP source fails.

Content Analysis

- Added regular expression-based detection of text patterns, such as bank account numbers, monetary values, shipment tracking numbers, patent numbers, etc. This functionality comes

with a Regular Expression Assistant for constructing and testing regular expressions on sample text. This Assistant also offers a library of example regular expressions.

- The item field(s) on which content analysis is applied can now be specified. Before, this could only be done on the document text.
- Improved content analysis speed. Typically, computations can be up to twice as fast depending on the CPU type, except for skin tone analysis.
- Content analysis is now performed as a background task, meaning that it no longer blocks the Search UI.
- When launching content analysis, the user can now choose which categories are to be calculated.
- Added support for various image types in skin tone analysis.
- Put the three skin tone analysis categories in a proper order: Weak, Medium, Strong.
- Resolved an issue with the skin tone categories not updating after the content analysis task completed.
- Resolved an issue with the “Content Analyzed” flag not being cleared when running content analysis with the clear option switched on.

Email Threading

- Added functionality for threading a set of emails. This process determines the “reply”, “reply all” and “forward” relationships between emails, based on metadata found in the email headers, the email container or embedded in the email body.
- The resulting sequence or tree of emails is displayed in the Email thread tab in the Previewer, with an indicator of where the current email is located within its thread.
- Mails that are referenced in the email metadata but that could not be found in the evidence data are marked as “missing emails”. An example is a mail with an In-Reply-To header that refers to another mail that is not present in the current evidence set.
- The “inclusive” mails are determined and highlighted in the Email Thread tab. These are the mails that *together* contain all content present in the thread. Having read all inclusive mails implies having read the entire thread. This can be used to improve the time needed to review a large collection of emails.
- The determined threads are listed in the new Email Thread facet and can be used as queries.

Case Tasks

- Increased power of the Case Tasks framework:
 - Each task can now have multiple conditions and actions.
 - Each task can have a set of filters that is applied on the items that match the conditions, for example deduplication, finding top-level parents, etc.
 - The date filter can now be applied independently, rather than it being applied on the set of items produced by another filter.
- A user interface for editing, importing/exporting and executing case tasks has been added to the Search tab. Previously, such case tasks could only be entered when defining a source.
- A task condition called “OCR Candidates” has been added. This can be used to gather e.g. all documents and non-embedded images. This way they can be conveniently OCR-ed, tagged or exported. The condition can be configured to focus on specific types of documents and images, whether it should be limited to empty documents (not containing any text), etc.

- Resolved an issue with date conditions not using the correct time zone.
- Tags placed by a case task used to have “admin” as their creator. This is now the creator of the task. This lets this user delete the tag afterwards, rename it, etc.

Background Tasks

- Improved error reporting.

OCR

- Added functionality for OCR-ing items. This can be done in three ways:
 - Using an internal, fully integrated OCR engine (ABBYY FineReader). This lets users OCR items directly from within Intella Connect, without requiring any additional software, systems, or licenses.
 - Using an external ABBYY Recognition Server, to scale up to larger data sets.
 - Using a manual export and import method, letting users apply the OCR tool of their choosing to produce the OCR output.
- The OCR text is now shown in a separate tab in the Previewer; it is no longer part of the Contents tab. It is still subject to full-text search, this is only to make it clear where the text originates from.

Searching

- The divider in between the facets and the results table can now be dragged horizontally, letting the user control how much screen space is given to the facet area.
- One can now configure whether the Location facet’s trees need to be collapsed or expanded by default when opening this facet.
- Added an “Has Attachments” category to the Features facet.
- Resolved an issue with saved searches referring to items that no longer exist after a re-index.
- Resolved an issue with keyword lists, hash lists and other types of search lists not displaying properly in their respective facet when the file name of the list contains non-Latin characters.
- Resolved an issue with reviewers not being able to search using the selected item ID list.

Results

- Added a Geolocation column, reporting the longitude and latitude of items such as photos.
- An “Analysis” table column group has been added, containing columns that indicate items whose content has been processed by content analysis, email threading and OCR.
- The CSV export option can now export arbitrary fields from the Raw Data section of an item. An example use case is exporting of the PR_... MAPI properties of the selected items.

Previewer

- Instant message types such as SMS, MMS, iMessage and the various chat clients supported by the cellphone extraction tools are now processed similar to how Skype messages are processed and displayed: all messages between two people or in a group chat are combined into items that cover the messages of a single day, with the option to navigate to the previous and next day in the conversation. This improves the ease of review of such instant message types.

- Many improvements in rendering PDFs. This also affects the Preview tab for other item types, such as MS Office documents.
- Added support for rendering all pages of a multi-page TIFF. These images were already correctly imported and exported, but the Previewer would only show the first page.
- Added support for natively previewing OpenDocument documents.
- CSV and XLS files now render with auto-fitting of columns in the Preview tab (optional for XLS).
- For XLS files, text is no longer truncated using scientific notation in the Preview tab, unless the cell in the original file is set to use scientific notation.
- The Headers tab now preserves the indentation in the header text.
- The PR_MESSAGE_FLAGS value in the Raw Data tab now shows a human-readable value.
- Resolved an issue with the Previewer not showing the Attachments tab when viewing calendar items.
- Resolved an issue with tags in the Previewer being truncated and not revealing the full text in a tooltip.

Insight

- Added an info box that shows basic case information such as the case folder, case size, creation date, etc.

Batching

- The batch algorithm now sorts the items by Family Date before dividing them into batches. This ensures that item families are always put in the same batch in their entirety and that all items in the family are presented consecutively.
- Column improvements in the All Batches list:
 - The set of columns that are shown is now configurable.
 - Added “Created” and “Completed” date columns.
 - The list can now be sorted on any column.
 - The list can be filtered on user-entered text and can optionally hide the archived batches.
- Resolved an issue with case managers not being able to assign batches to LDAP accounts.
- Resolved an issue with certain batches becoming too large when the “Keep families together” option was selected.

Coding

- The item list now shows the index numbers that the items have in the current batch.
- The buttons in the Coding user interface now show the keyboard shortcuts for those actions in the button tooltips.
- The widths of the main panels in the Review page (item list, item content and coding panel) can now be resized.
- When a batch that is actively being coded is removed by the batch manager, the coding UI will now immediately show that the batch no longer exist and go into read-only mode. Before, one could continue coding and get an error at the end when the batch was updated.
- Various other minor usability improvements.

Redaction

- Changed the settings in the default redaction profile so that only limited metadata is contained in exported redacted items.
- Resolved an issue with redaction rectangles not being painted correctly on rotated PDF documents.

Exporting – Original Format

- File names are no longer truncated to 120 characters when exporting on Windows 10, as the limits for file name lengths have been increased on that platform.
- Resolved an export error that would occur when exporting items typed as “Email Headers”.

Exporting – PDF

- Added the ability for custom designations, e.g. texts like “Confidential” and “For attorneys’ eyes only”, to be placed in one of the corners of the produced PDF. The designations that are added to a specific item are controlled using tags.
- It is now possible to have headers and footers centered at the top and bottom of a page, in addition to the four page corners.
- The pages can now be numbered automatically.
- Added support for Open Type font (OTF) files.
- Added support for exporting various image file types: Windows icons (ICO files), HDR, Apple icons (ICNS), IFF, PCX, Photoshop (PSD), SVG, WMF/EMF (partial).
- Various fixes and improvements for exporting MS Office files to their native rendering.
- Resolved an issue with redacted attachments and embedded items being exported incorrectly.
- Resolved an issue with (partially) transparent pixels in PNG files being rendered incorrectly.

Exporting – PST

- One can now export directly to the PST root folder.
- Resolved an issue with MSG files contained in a ZIP file that could not be exported to a PST file.
- Resolved several issues where specific types of calendar files could corrupt the PST file they were being exported to.

Exporting – Load Files

- All PDF exporting improvements apply to the exporting of load files as well.
- Added “first Bates number” and “Last Bates number” columns to the Exports table, showing the Bates range used for a specific export.
- Added an “Has extracted or OCRed text” field.
- When exporting items as images in a non-PDF format (e.g. TIFF or PNG), PDFs can now optionally be generated as well.
- Added support for exporting to Multi-page TIFFs.
- One can now export arbitrary fields from the Raw Data section of an item. An example use case is exporting of the PR_... MAPI properties of the selected items.
- When skipping items, the document type can now optionally be mentioned in the placeholder text.

- The unit of the Size column can now be bytes, kilobytes, megabytes or gigabytes.
- Added the Page Count field to Opticon (OPT) files.
- Resolved an issue with an incorrect load file being created when export errors occurred.

Upgrade Notes

Case compatibility – Intella Connect 2.1 can directly open cases made with Intella 2.1 and Intella Connect 2.1.

Cases made with the 1.9.x and 2.0.x versions of Intella and Intella Connect can be opened, but these cases first require conversion.

Case conversion will create a copy of the case in which all evidence is re-indexed and all tags, comments and flags are imported. The existing case will not be altered in any way and can afterwards still be opened in the older Intella version.

Caveats concerning the case contents:

- Case conversion will not transfer the geolocation metadata extracted from emails when the "Determine geographic location of emails" option was used. Re-indexing of the converted case is required to restore such metadata.
- Multi-page TIFFs will be displayed in the converted case as if they were single-page TIFFs. Exporting and printing of the item report does reveal the other pages. Re-indexing of the converted case will make all pages displayable again.

Case conversion takes considerable time, comparable to what it took to index the original case.

Case conversion will also require sufficient disk space. As a rule of thumb, please reserve twice the amount of the evidence size for your case folder.

Cases made with Intella 1.8.x or older are not supported.

Cases made with beta versions are not supported and should be recreated.

Coding layouts – When a coding layout has been previously imported in another case, Intella Connect 2.1 will apply a one-time conversion of the underlying XML model stored in INTELLA_HOME/coding-layouts/conversion/file-mappings.xml. This will happen upon startup. A backup of the mentioned file with a *.bak extension will be created in that same folder. If for any reason you wish to roll-back the Connect version migration, this file will have to be recreated manually from the backup file.

Case sharing limitations – The supported number of concurrently active cases that Intella Connect can share is four cases at any given time. Previously this has been managed via a fair usage policy. The definition of an active case is one that is shared with a reviewer logged in or reviewing that case. A case that is shared but does not have any active reviewers logged in does not count towards the four case limit. Connect administrators needing to have more than four active cases at any one time will need to purchase a second Intella Connect license and set up another dedicated server for it. Optionally, administrators can take advantage of the "Intella Connect Grid" feature to supply their reviewers with unified access to all cases shared by the servers forming the grid.