



Published on xpriori.com (<http://xpriori.com>)

[Home](#) > [Printer-friendly PDF](#) > Printer-friendly PDF

FirstCull Frequently Asked Questions

Answers to the most Frequently Asked Questions...

1. [WHAT IS THE PURPOSE OF FIRSTCULL?](#)
2. [WHAT BASIC REPORT ATTRIBUTES DOES FIRSTCULL PRESENT?](#)
3. [HOW IS FIRSTCULL DELIVERED AND USED?](#)
4. [HOW IS FIRSTCULL LICENSED?](#)
5. [WHAT DOES THE USER RECEIVE WHEN HE/SHE PURCHASES A LICENSE?](#)
6. [HOW ARE MICROSOFT™ PST CONTAINER FILES HANDLED?](#)
7. [HOW ARE LOTUS NOTES™ NSF FILES HANDLED?](#)
8. [HOW ARE OTHER CONTAINER FILES HANDLED?](#)
9. [WHAT TYPE OF DUPLICATE FILE ANALYSIS/REPORT IS PRESENTED?](#)
10. [HOW ARE IMAGE FILES HANDLED?](#)
11. [HOW MANY FOLDER/FILES CAN I ANALYZE AT ONCE?](#)
12. [HOW FAST DOES IT WORK WHAT FACTORS WILL AFFECT SPEED OF OPERATIONS?](#)
13. [WHAT IS THE DIFFERENCE BETWEEN INSTANT AND DEEP ANALYSES?](#)
14. [IS ANY CONFIDENTIAL OR PRIVILEGED INFORMATION ON ANY OF THE REPORTS?](#)
15. [DOES XPRIORI STORE OR COLLECT ANY PROJECT SENSITIVE INFORMATION FROM MY PROJECTS AND STORE IT ON ITS SERVERS?](#)
16. [WHAT IS THE DIFFERENCE BETWEEN THE FIRSTCULL ANALYSIS PROCESS \(EARLY CASE ASSESSMENT - ECA\) AND ELECTRONIC DISCOVERY \(EDISCOVERY\)](#)
17. [WHAT DATA SOURCES CAN YOU COLLECT INFORMATION FROM?](#)
18. [WHAT IS OCR \(OPTICAL CHARACTER RECOGNITION\) AND WHY IS IT IMPORTANT?\[FROM WIKIPEDIA...\]](#)

WHAT IS THE PURPOSE OF FIRSTCULL?

The purpose of FirstCull is to provide an analysis of the attributes of large amounts of documentary or “unstructured” information prior to its being reviewed for content considerations in litigation, compliance, research or other instances where discovery of content will take place.

It is intended to give the user:

- Assistance in predicting the amount of time and effort which will be required to manage the information in the dataset;
 - An idea as to the size and file content of the dataset to be managed;
 - The opportunity to identify and exclude files (to identify such files) and/or not load them to a project repository where, for example, they would not contain substantive content useful in the project;
 - the identity of parties to email content and the subjects of email threads
-

WHAT BASIC REPORT ATTRIBUTES DOES FIRSTCULL PRESENT?

What attributes does FirstCull present?

FirstCull presents a series of reports on metadata and certain other information on all files included in a dataset including emails and their attachments. It presents the following:

- a numerical count for all file extensions, custodians, authors and other metadata;
- analysis of files in containers such as PST (outlook) and NSF (lotus notes/domino);
- analysis of potential duplicate files;
- email to, email from, and email subject matter;

Reports are delivered in either PDF or CSV format. The CSV format is delivered to facilitate the introduction use of the data to spreadsheet (Excel and others) and other reporting software such as Crystal Reports for presentation in other custom reports. (See white paper on use of FirstCull CSV reports with MS Excel).

HOW IS FIRSTCULL DELIVERED AND USED?

FirstCull is delivered by download to the user's computer from the FirstCull site; except for the corporate license which is handled offline.

FirstCull is used from the user's computer and can analyze files contained in multiple folders on the computer or on the network which the user can choose according to his own plan. FirstCull does not require the movement of files to a new repository or database but analyses them where they reside.

Files analyzed are always referenced to folder locations on the computer or network.

There is no limit to the number of sessions or "runs" when using FirstCull during the annual subscription period.

HOW IS FIRSTCULL LICENSED?

Lite Version (FREE for 30 days)

- 1 User
- All Features Enabled
- Process up to 5000 files per run
- Requires REGISTRATION using a legitimate e-mail address

- *Intended to be a trial or demonstration license

Pro Version (\$699/year)

- 1 User
- All Features Enabled
- Process unlimited files per run
- Annual subscription or permanent license available
- Can be used simultaneously on up to 3 computers per licensed user

Enterprise Version (\$4799/year)

- 10 Users
- All features enabled
- Process unlimited files per run
- Administrator has to register as the account holder
- A license key activates FirstCull. The product can be installed and used in up to 3 simultaneous locations per licensed user

Corporate Version (call for quote)

- It makes sense to purchase the Enterprise license when one anticipates the need for 10 or more users
- Call for a Quote
- Negotiable Terms

WHAT DOES THE USER RECEIVE WHEN HE/SHE PURCHASES A LICENSE?

- He/she receives a copy of the application software
- He/she receives all upgrades and fixes created during the license period
- Renewal annual subscription licensees will renew at the same rate

HOW ARE MICROSOFT™ PST CONTAINER FILES HANDLED?

PST files contain multiple email files and their respective attachments. FirstCull opens the containers and does its analysis on the files and attachments.

HOW ARE LOTUS NOTES™ NSF FILES HANDLED?

NSF files are similar to PST files but with respect to files contained in Lotus Notes or Domino data stores. FirstCull opens those files and does analysis on the emails and attachments and any other files contained in the NSF files.

HOW ARE OTHER CONTAINER FILES HANDLED?

The following container or compressed archive type of files are handled in all versions of FirstCull: .ZIP, .CAB, .RAR, .TAR. Others will be added in future releases. The compressed

files are expanded and metadata is extracted from the files for presentation by file types etc. in the FirstCull Reports.

WHAT TYPE OF DUPLICATE FILE ANALYSIS/REPORT IS PRESENTED?

- FirstCull runs standard hash table analysis on all files and presents the number and identity of potential duplicates.
 - Potential duplicates area presented in a CSV report.
-

HOW ARE IMAGE FILES HANDLED?

Image files – pixel based files – are accounted for by file type and all other metadata normally associated with the files. Such files are also files which will likely require Optical Character Recognition Treatment (“OCR”) before they can be indexed for text search. FirstCull assesses each such file and prepares a report on files which will require OCR treatment.

HOW MANY FOLDER/FILES CAN I ANALYZE AT ONCE?

- There are no determined limits. In tests, FirstCull has analyzed as many as 8.6 million files across multiple folders, including large PST files and NSF files.
 - Users can designate multiple sources(folders, directories, systems, networks, etc.) to analyze in the same project (there is no set limit)
-

HOW FAST DOES IT WORK WHAT FACTORS WILL AFFECT SPEED OF OPERATIONS?

In large part, the speed is determined by the size of the dataset being analyzed – the smaller, the shorter the timeframe. FirstCull assessing 8.6 million files, and operating on a dual core processor Dell™ computer, the analysis was completed in less than 30 hours. In analysis of files sets of a few hundred thousand files, it is often completed in the matter of an hour or two.

The speed is also dependent on the complexity of the container or archive files being assessed and the OCR assessment. As container files require multiple processing steps, the process is slower than handling files not found in such containers.

Speed is also affected by the OCR analysis. FirstCull determines the number of pages required for OCR and, therefore, must look at each page.

If you do not want to do analysis of container files or OCR analysis, you may choose not to do so during the product set during each run.

If users are exporting email files from Outlook to PST containers, it is better not to use single container containing large numbers of files. The better practice is to export to a series of container files that perhaps reflect the subject matter or folder structure contained on the server or computer housing the files. This is also the case where files are being introduced to a repository where they will be indexed for free text search.

WHAT IS THE DIFFERENCE BETWEEN INSTANT AND DEEP ANALYSES?

As emails and OCR analysis require more processing steps, FirstCull offers Instant analysis for situations where email analysis is not required. Thus, if users have a dataset that does not include emails, Instant analysis should be sufficient.

IS ANY CONFIDENTIAL OR PRIVILEGED INFORMATION ON ANY OF THE REPORTS?

FirstCull no analysis on the issues of confidentiality or privilege which are issues that are determined by the content of the files or the context in which they were created or delivered. Where the reports only show the file extension count etc., there is no possibility of disclosure.

The only place where disclosure of privileged or confidential information could be issue is in the email sender, receiver and subject matter analysis. For example, if there is disclosed there a communication between a husband and wife, between a lawyer and client, or between other parties to a privilege arising from a relationship, there could possibly be a disclosure. If the analysis is being done in connection with litigation support or in anticipation of litigation or compliance activities, users should work with the owner of the information and his counsel to assure no unintended disclosure takes place.

DOES XPRIORI STORE OR COLLECT ANY PROJECT SENSITIVE INFORMATION FROM MY PROJECTS AND STORE IT ON ITS SERVERS?

No. Xpiori or FirstCull does not store any information. It is stored on the computer of the user or at some other place they designate. Xpiori neither stores nor collects information.

WHAT IS THE DIFFERENCE BETWEEN THE FIRSTCULL ANALYSIS PROCESS (EARLY CASE ASSESSMENT - ECA) AND ELECTRONIC DISCOVERY (EDISCOVERY)

FirstCull does information assessments that are useful in the E-Discovery Process but does not represent the whole process. It can be a part of the whole process and is particularly useful to the IT consultant or user wanting to assess the amount of document review that needs to be accomplished.

WHAT DATA SOURCES CAN YOU COLLECT INFORMATION FROM?

FirstCull can collect information of files that are located in any folder on the network. In wide area networks, FirstCull works well. In other words, the information need not be aggregated in a single location for FirstCull to be fully operational. For example, users may want to do an assessment on folders as part of determining whether it is appropriate to copy and aggregate information to a project based repository.

WHAT IS OCR (OPTICAL CHARACTER RECOGNITION) AND WHY IS IT IMPORTANT? [FROM [WIKIPEDIA](#)^{[1]...}]

“Optical character recognition, usually abbreviated to OCR, is the mechanical or electronic translation of images of handwritten, typewritten or printed text (usually captured by a scanner) into machine-editable text. OCR is a field of research in pattern recognition, artificial intelligence and machine vision. Though academic research in the field continues, the focus on OCR has shifted to implementation of proven techniques. Optical character recognition (using optical techniques such as mirrors and lenses) and digital character recognition (using scanners and computer algorithms) were originally considered separate fields. Because very few applications survive that use true optical techniques, the OCR term has now been broadened to include digital image processing as well.”^[1] Unstructured information often contains documents of this character – TIFs, faxes, documents that contain diagrams or drawings, some PDF files and the like. For such documents to be susceptible to free text search, they must be converted through OCR. This time for this process is often material in preparing documents for content review. Thus, FirstCull includes OCR assessment.

Xpiori, LLC ? 2864 South Circle Drive, Suite 401 ? Colorado Springs, CO 80906 ? 719-425-9840 ? Fax 719-203-6496?© 2014-2019

Source URL: <http://xpiori.com/content/firstcull/firstcull-frequently-asked-questions>

Links:

[1] http://en.wikipedia.org/wiki/Optical_character_recognition