

# Introduction

C4P stands for Categorizer for Pictures. There are two applications that use the name “C4P” – an EnCase EnScript, and a Windows End-User Application.

This document applies to the EnScript, and is intended to provide clarification about the many options and features of the script itself. Any questions regarding the operation of EnCase, or EnScript in general should be directed to the owners of EnCase – Guidance Software, Inc.

The C4P EnScript’s author is Trevor Fairchild. Mr. Fairchild is a civilian Programmer / Analyst with the Ontario Provincial Police, based out of Orillia, Ontario, Canada.

For information about what the C4P EnScript does, or how C4P works, please explore this website:

<http://www.e-crime.on.ca/software/C4Pintro>

If you wish to contact the author, please see this section:

# Main Settings Page

The Main Settings Page allows users to set the most common options in the script. Each section is outlined in detail below:

## C4P Disclosure Package Options

### Create Package here:

The purpose of the C4P EnScript is to find and extract pictures from EnCase evidence files. The path specified under “Create Package here:” will be the *starting* location used by the script during runtime.

The script will build upon the *starting* location by adding a folder called “C4P”. From that point on, the script will write all reports and files out to the “C4P” folder.

### Wrap extracted files in DVD-sized Folders (instead of CDs)

Historically, the C4P script was designed to allow the extracted files to be easily burned to a CD and then delivered to an investigator. Although this medium of disclosure has become antiquated, and is in fact harmful to the use of C4P, it is still the standard within the script’s settings.

If this box is checked, the script will extract all files into DVD-sized folders as opposed to CDs. The only purpose in doing so is to cut down on the overall number of CD folders produced (a visual consideration more than anything else), *or* to actually facilitate the transporting of these files using DVDs.

### Create a Bookmark for every file that is extracted

Checking this box will result in every file extracted to be bookmarked as well. Unless there is a specific reason to do this, it is strongly recommended to leave this option *unchecked*.

Creating hundreds, or thousands, or millions of bookmarks in an EnCase case file can have a serious impact on the case file’s performance or, worse, cause corruption in the case file.

## Selected Files Options

### Search All Files

Select this option if you want to search everything possible in the EnCase evidence files. This is the default option.

### Search Selected Files Only

Select this option if you wish to only search those files and folders that have been blue-checked in EnCase. If this option is selected, the “Search File Slack” checkbox becomes enabled.

### Exclude Selected Files

Select this option if you wish to search everything in the EnCase evidence files *except* those files and folders that have been blue-checked. If this option is selected, the “Search File Slack” checkbox becomes enabled.

### Search File Slack Regardless

Check this box to search the file slack of excluded files. This option is not available when the “Search all files” option is selected.

This option is useful for users working with hash sets of “Known” pictures. Hash values that are computed via EnCase are based on Logical Size only, meaning the File Slack is not considered part of the file. If a file is excluded based on the logical data, any data in the file slack (which isn’t “Known”) could be overlooked.

## Pre-Categorization Options

An optional feature of C4P is the use of a Hash Database. This database is unique to C4P, and does not make use of the EnCase hash sets in any way. For more information the hash database, see [here](http://www.e-crime.on.ca/software/C4Pintro/timesavers.html).

There are three different database platforms available in the C4P EnScript:

### No Database

Check if you do not have a Hash Database.

### MS Access

Check if you are using an Access database file. Although present in the script, Access is *not* a recommended platform to host the Hash Database. When you select the Access option, you will be asked to browse to the C4P\_Hash.mdb file you are using.

### SQL Server

Check if you are using SQL Server, full or Express. When you select this option you will be prompted to identify the named instance of your database. *Examples: “srvDb”, “srvDb \SQLEXPRESS”*

### mySQL

Check if you are using mySql. When you select this option you will be prompted to identify the named instance of your database. *Examples: “srvDb”, “srvDb\mySQL”*. Alternately, you may enter the complete ODBC connection string for your database platform. This method requires knowledge of connection strings – for help see here.

### Populate Picture Library

Check if you want to add all pictures found by the script to your Picture Library. This is also an optional feature of C4P and should *only* be checked if you are actively maintaining a picture library.

## Search Options

### Mount Email (pst)

Check if you want the script to look for and mount Email Archives. This feature is based on file extension only. It is not recommended to use this option.

For further details, see [C4P Recommended Practices](http://www.e-crime.on.ca/software/C4P3/download/presentation/Recommended_Procedures_for_C4P.pdf).

### Mount Zip Archives

Check if you want the script to look for and mount Zip Archives. This feature is based on file signature analysis. It is not recommended to use this option.

For further details, see [C4P Recommended Practices](http://www.e-crime.on.ca/software/C4P3/download/presentation/Recommended_Procedures_for_C4P.pdf).

### Mount Thumbs.db Archives

Check if you want the script to look for and mount Windows Thumbs.db volumes. This feature is based on file name. It is not recommended to use this option.

For further details, see [C4P Recommended Practices](http://www.e-crime.on.ca/software/C4P3/download/presentation/Recommended_Procedures_for_C4P.pdf).

### Search for Base64 Headers

Check if you want the script to look for Base64 encoded pictures. If this option is selected, the script will extract the decoded pictures, and bookmark the original evidence to provide traceability. The *comment* field of the bookmark will match the extracted name of the decoded picture.

### File Format Options

Clicking this button will open a new set of windows in relation to the various file formats targeted by the C4P EnScript. See here for more.

## Debugging Options

### Run in Debug Mode

Check this box to run the script in Debug Mode. If this option is checked, then the script will create a “Debug” folder in the output folder, and place Debug.txt files inside. The Debug.txt files will be broken into 50MB chunks to allow them to be opened.

*Do not* check this option unless you are troubleshooting. Using Debug will slow down the script’s performance significantly.

# Case Report Values

The Case Report values are used to create the “Case Report.xml” file, which is part of the C4P Master Disclosure Package. The purpose of the Case Report is to make the case unique against all other C4P packages.

In addition to the three fields listed below, the Case Report will include the following:

* EnScript start date
* EnScript start time
* EnScript version
* EnCase version

These values are important to uniquely identify the case as well as to troubleshoot problems after-the-fact.

In the EnScript dialog, you will see the following three fields:

### Case Name

This value is taken direct from the EnCase case file. It can be changed as desired. Names in C4P are important, given that they may be used as part of a report file structure. As such, it is suggested to use the same naming conventions that apply to Windows file names.

**Avoid using the following characters: \/:\*?”<>|**

### Forensic Examiner

This value is taken direct from the EnCase case file. It can be changed as desired.

### Comments

This value is pre-populated with the version of EnCase that the EnScript has been tested against. This is also helpful when troubleshooting, but is not required. The comment field is a ‘free’ area for you to enter whatever additional information may be important later on.

# Resume Settings

“Resume” is an option of the C4P script that allows you to pick up the trail of a previous, failed attempt, to run the script. There are a variety of reasons and situations that will result in the EnScript crashing, or shutting down abruptly. Sometimes, this may happen many hours, or even days into the script’s execution, and rather than running the script from “square one” again, the Resume option lets you try to continue the previous running…

## Important!

***If the script crashed once it will likely crash again. The Resume option is useless unless the original problem has been identified and either fixed or bypassed.***

## Environmental factors required for success

Because of the complexity of C4P, the ability to Resume a previous script execution is tricky. In order for it to work, the following must be *identical* between the first and second attempts:

* The ROOT of your case path must be identical
  + This is the value found on the “Main Settings Page” tab, under “Create Package here”
* The option to use CD or DVD folders must be identical
  + This option is also on the “Main Settings Page” tab
* Alternately, if you have chosen to export files by their MD5 hash value, that selection MUST be replicated here
  + This option is under the “Advanced Options” tab
* You must set the Starting Entry in the “Resume Settings” tab.
  + This procedure is covered below
* You must identify the assigned number of the last picture extracted
  + Also covered below

## How to use the Resume option

To use the Resume option, you must first know where the original script execution stopped. To do this, you would open the C4P Index.xml file that was originally created. Be sure to open this file in a Text Editor, and NOT a web browser. It is possible that the Index file is too large even for a Text Editor to work with, in which case you are advised to contact the EnScript [author](mailto:c4p@e-crime.on.ca?subject=C4P%20EnScript:%20Resume%20Settings).

### Select the Starting Entry

Once you know the Full Path of the last successfully searched entry, you need to select (blue-check) it in EnCase.

**Be sure that the *only* selected entry is the one you are looking for.**

With that single entry blue-checked, you would then hit the “Capture” button in the EnScript’s Resume Settings dialog. After hitting the “Capture” button, you will see the entry’s Full Path inside the “Starting Entry” box.

**Note:** as specified above, this may not be the best place to resume the script – previously the script crashed after the last successfully searched entry…. some investigative work may be needed here.

### Last CD / DVD folder

In the original disclosure package, look at the numbered CD or DVD folders, and enter the last number present. This option is unimportant if you are exporting files by their MD5 hash value.

### Last Picture Extracted

If you go to the last sequential CD or DVD folder, and then to the last sequential folder inside, and then go to the last file in that folder, you will find the number required for this field. The files are extracted and sequentially names, so the last file extracted is the last ID number.

This number is **important regardless** of the “export files by their MD5 hash value” option.

## Desired Result of Resume

Once all settings have been entered, when the script is started, it will skip over every evidence file in the case until it hits the one “Captured” by you. From that point on it will then search the data. The files will be extracted to the proper folders to carry on the numbering system previously used, and the C4P Index.xml file will be **added to** instead of replaced.

# Advanced Options

This tab exposes two new options in the v4.0.2 EnScript, the purpose of both being to cut down the amount of space required to store the C4P Master Disclosure Package, as well as to reduce the size of the C4P Index.xml file.

## Pre-Categorization

### Ignore a Category

This option will only have an impact on the extracted data if you are using a C4P Hash Database for pre-categorization.

Every C4P Hash database requires the presence of a “junk” category. This “junk” category is used to store all the hash values of those pictures which are of no investigative value, such as windows icons.

If your organizational policy permits, you can identify what category is your “junk” category. When the script finds a picture in your evidence that is a match to that “junk” category, it will NOT extract that file, and will NOT add a record to the C4P Index.xml file.

### Bookmark these pictures instead

Checking this box will enable the forensic analyst to review those pictures which were found, but NOT extracted as bookmarks internal to the EnCase case file.

### Benefits

* The script will complete faster
  + data extraction takes time, and extracting fewer pictures = faster completion
* There will be less data to manage on your RAID/external hard drive/etc
  + Fewer pictures extracted = less drive space demands
* Creating this case in C4P will be faster
  + Fewer pictures means fewer records to load into the C4P database
* C4P cases will be smaller, thereby limiting exposure to the “2GB limit” of an Access database

### Considerations

Some organizations like to give percentages of a case “15% of the suspect’s drive held criminal material”. These types of statistics will be unreliable if you are not extracting EVERYTHING from a case

Additionally, mistakes in the C4P Hash Database (pictures mistakenly marked as “junk” and added to the Hash Database) would never get exposed to the Investigator – they would never make it out of EnCase. In this situation, the bookmark option described above would be a good mitigating action, allowing the Forensic Analyst to at least scan through the “excluded” files as a precaution.

## Unique Pictures

### Store Extracted Files by their MD5 Hash

Checking this box will cause all pictures extracted to be named after their MD5 hash value, as opposed to the ‘standard’ sequential number.

i.e.: *18A03EFEB873C312EE1F9C1B4B3CC9B7.png* instead of *00002894.png*

Additionally, the files will be stored in nested folders, two levels deep, based on the first two characters of the MD5 hash value.

i.e.: *18A03EFEB873C312EE1F9C1B4B3CC9B7.png* becomes

..\1\8\*18A03EFEB873C312EE1F9C1B4B3CC9B7.png*

### Benefits

Drive space. This is the only benefit to storing files by their MD5 Hash.

*Example:*

A case produces 100,000 files. 40,000 of these files are duplicates (by hash value). All 100,000 files represent 14.6 GB of disk space. By storing these files by their hash value, you are only storing 60,000 files – one copy of every different file. In this example, you would save 5.84GB of data.

### Important Clarification

This has NO impact on the data in the C4P Index.xml file. In the example above, ALL 100,000 records would be referenced in the case.

### Considerations

This feature demands unequivocal trust of the MD5 Hash calculation algorithms used in EnCase: both the algorithm itself, and the implementation used within EnCase. For several years now there has been documented proof of MD5 Hash collision (<http://en.wikipedia.org/wiki/MD5>). A collision means that two different files were proven to produce the same MD5 Hash value.

To be clear, this trust in the MD5 Hash algorithm is built into the framework of C4P, and has been for years. This new option, however, means only one copy of a picture is extracted from the evidence, whereas in the original method, all copies would be extracted. Although you would never see the duplicate files using C4P, you *could* manually review each copy through Windows Explorer…. That is unlikely but it needs to be stated here for the record.

# Mounting Options

The two check boxes on this tab *only* apply if you have checked the option to “Mount Zip Archives” (on the Main Settings Page). It should be noted that the [C4P Recommended Practices](http://www.e-crime.on.ca/software/C4P3/download/presentation/Recommended_Procedures_for_C4P.pdf) does not recommend using the C4P EnScript to mount anything. Rather, the “GSI File Mounter” should be utilized prior to running the C4P EnScript.

### Mount ‘Unknown’ Zip Volume Types

The ZIP file header contains a byte value which is used to indicate what *type* of zip file the item is. While there are 256 possible values for that byte, the C4P EnScript has only been programmed to account for the most common ones.

Attempting to mount a zip file which is of a type that has not been accounted for specifically could result in the EnScript failing, or EnCase crashing.

This option should only be checked when you know there are files that should be mounted, which aren’t.

### Copy Protected Zip Volumes

Attempting to mount a password protected zip volume within EnCase will cause either the EnScript to fail, or EnCase to crash. If this option is checked, and the C4P EnScript encounters a protected zip volume, then it will be extracted to the “Protected Zips” folder in the export location. It will be up to you to manually crack this file and then analyse the contents.

# File Format Options

The file formats targeted by the C4P EnScript have their own settings which can be changed at the start of the EnScript. These settings should be left at their default state unless there are specific circumstances which dictate otherwise.

To access the File Formats, you need to hit the “File Format Options” button. This button can be found at the bottom of the Main Settings Page tab.

## ART Settings

### Search for ARTS

Check this box to search for ART files. This option is turned *off* by default, as C4P Windows is unable to show ART files. See this file for more details: [C4P: working with AOL](http://www.e-crime.on.ca/software/C4P3/download/C4P_Working_with_AOL_ART.pdf)

### If no ART Footer found…

The ART file format is very simplistic, and has no internal structure which the C4P EnScript can work against. As such, if a footer cannot be found in a reasonable amount of space, the script must blindly extract data, and the amount of data extracted is dictated by this value.

If your ART files are being unfairly truncated by the EnScript, increase this value.

### Set a maximum size…

The C4P EnScript is unable to make an educated guess about the validity of an ART file – it can only rely on the presence of a header and a footer within a reasonable amount of space. The value of “reasonable” is specified in the “Set a maximum size” field.

## BMP Settings

### Search for BMPs

Check this box to search for BMP files.

### Minimum extracted BMP size…

This value imposes a minimum size for every BMP extracted.

## GIF Settings

### Search for GIFs

Check this box to search for GIF files.

### Maximum allowed GIF size

This value imposes a maximum size for every GIF extracted.

## JPG Settings

### Search for JPGs

Check this box to search for JPG files.

### Assume a JPG Compression Ratio…

This value is a percentage and is used when no JPG footer can be found. The image data of a JPG is compressed, and as such cannot be properly validated by the EnScript. In the absence of a JPG footer, the C4P EnScript will take the image dimensions (contained in the JPG header) and determine the decompressed size of the picture. It will then take the JPG Compression Ratio value and arrive at a “best guess” for when the picture data should end.

Changing this value will only affect those pictures without footers.

### Maximum allowed JPG size

This value imposes a maximum size for every JPG extracted. Mostly it is required to control the data length that is “guessed” at by the EnScript.

## PNG Settings

### Search for PNGs

Check this box to search for PNG files

# Troubleshooting

### A note from the author:

I’ve been working with EnCase and EnScript since 2003. In all that time, I have come up with only one strong and steady rule: expect trouble.

Extracting data from evidence is tricky and troublesome. There are an infinite number of possibilities and configurations in digital forensics, and it is simply not possible to account for all of them. A fix for one problem may create a dozen new ones.

If you ever run into trouble, ask for help. Chances are you’ve caught something that other people are frustrated by, and you may end up helping a lot of other people in the process. Every bug report, error report, question about accuracy, etc helps make the C4P EnScript stronger in the long run.

I can be reached by email at: [C4P@e-crime.on.ca](mailto:C4P@e-crime.on.ca)

I can be reached by phone at: 1-705-329-6471

### Submitting an Error Report

When submitting an error report, please try to include the following:

* The version of EnCase you are using
* The version of the C4P EnScript you are using (compile date for older versions)
* If you are asking about Pre-Categorization:
  + The database platform you are using
* The version of Windows you are running on – especially 32-bit versus 64-bit

Answering these questions at the beginning will help eliminate a number of back and forth emails about the basics.