
D3
Dark Data Discovery

Table of Contents

Executive Summary 3

Dark Data Defined 3

Manage Your Data 4

Enterprise Profiler 5

Red Flags 5

eDiscovery 6

Analyze Targeted Data 6

Impossible Recovery 7

Files within Files 8

The Exception Bin 9

First Responder Triage 10

Architecture 11

Conclusion 12

Company Overview 13

Executive Summary

While digital forensics is imperative to both understanding what has happened at the desktop as well as any legal proceedings that may result from these investigations, the larger concern will always be in preventing valuable data (intellectual property or sensitive information) from ever leaving the premises. The Network and Security teams at companies and federal agencies typically deal with this later issue. FID^{3} has products that span the usage models of both these data usage concerns. FID^{3} focuses on discovering the hard to find data as well as anomalous data storage behavior.

Dark Data Defined

The size of the world's digital data collections is growing at an explosive rate. The only way to keep up with this growth is to increase the efficiency of managing this data. Whether we use a proactive approach (Cyber Security, Data Redaction, and User Monitoring), or a reactive approach (Electronic Discovery, Digital Forensics, and Incident Response), the same questions about the data need answers:

- "Where is the data?"
- "What is the data?"
- "When was the data changed?"
- "How does the data affect the current situation?"
- "Who is responsible for manipulating the data?"

At FID^{3}, we target the '**Where'** and the '**What'** about your data with such a passion that our products far outperform the rest of the industry. Why are these questions so important? If you do not know "Where" all of your data is, or "What" all of your data is, then you will miss valuable evidence that can make a difference in the outcome of any legal dispute or you could lose valuable intellectual property that can greatly affect your company's stability.

We have classified data currently overlooked by services and products available in the industry as **Dark Data.** This term is quite appropriate since it classifies hidden data that

is in the dark going undiscovered. We developed Enterprise Profiler specifically to bring this data into the 'light'. Whether you are working on discovery for legal processes or trying to stop the exfiltration of sensitive data, FID^3 focuses on recognizing and reporting on more data types, redacted data and non-contiguous data than any other product on the market.

Manage Your Data

The first step for the user is to gain an understanding of what the data is and where it is located. Our Enterprise Profiler product provides a high-level view of exactly 'Where' all of your unstructured data is sitting and 'What' the data contains. Enterprise Profiler enables you to view the data in a number of categories so you can easily pick one or more categories that are of interest and quickly drill down to the department(s) or even individual users that are of interest. Enterprise Profiler even provides you the ability to analyze individual files one at a time as well as the embedded objects within each file. Enterprise Profiler has persistent filters that will allow you to maintain a strict focus on just the content that is pertinent to your investigatory process.

What other features differentiate FID^3's products from the industry?

Enterprise Profiler goes deeper and broader than any other solution available on the market today. While the majority of the industry stops at recognizing 800 file types, we support 4,000+ file types and can process them accurately and efficiently.

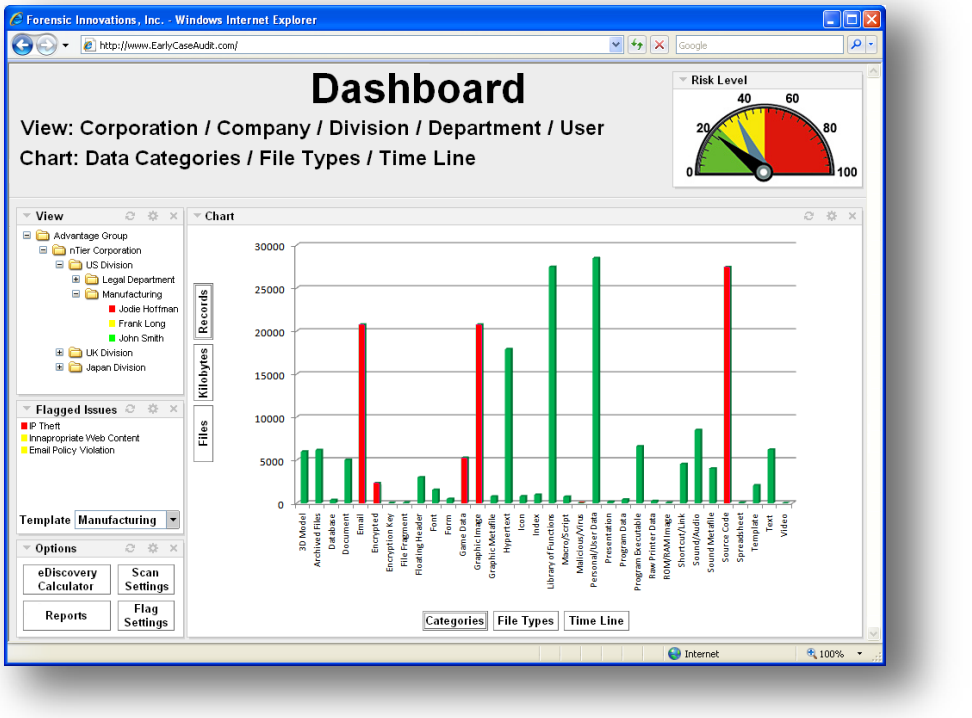
Why is that important?

Enterprise Profiler goes further than any other automated solution to recover fragmented files from unallocated disk space. In any investigation, that space is the first place to look. Not only will you find potential evidence from yesterday, but also data that has been there for many months. The typical user will have assumed that the data no longer exists and there is no evidence left.

FI – DARK DATA DISCOVERY

Enterprise Profiler

Enterprise Profiler performs a statistical analysis on all of the visible and hidden files for each hard drive, user, department, and geographic location of a company. While performing this analysis, Enterprise Profiler is able to perform additional tasks. When it encounters specific types of files, it can copy or move the files to another location (i.e. network server, external storage device) as part of a collection event. A collection event ensures the company's intellectual property is located on servers that back-up regularly. Automated shortcut replacements is a feature that allows the user to still access a file even though it has been moved.



Red Flags

When 'bad' files, as defined by the user, are found, Enterprise Profiler can automatically alert management, move the file and/or wipe the file from existence, so that it can never be recovered. A Potential Threats Report creates and contains a list of files that require review to determine their level of threat. As an example, when a receptionist's computer suddenly contains a number of files that represent the company's intellectual property (IP), would represent a situation that needs to be logged for future review and an urgent alert is sent to the Information Technology staff/Information Security Officer (ISO). The best-case result is that the IP is stopped while it is still under control of the

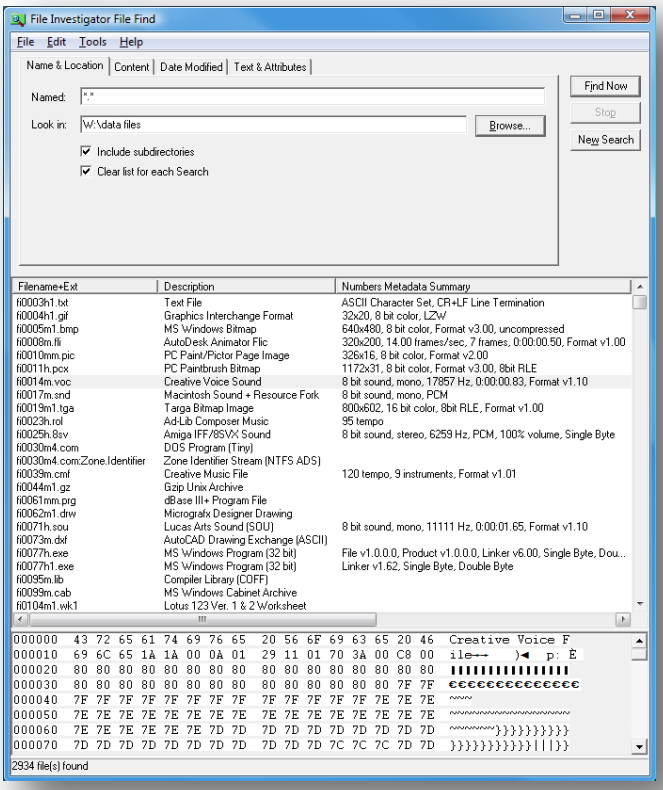
company. If discovery of the breach occurred after the data was already in motion, the exact contents of the breach can be determined.

eDiscovery

In the context of eDiscovery, the preceding weeks before a company enters litigation proceedings, they typically know that a lawsuit is imminent. With Enterprise Profiler, the relevant data of the employees involved in the pending case are analyzed to produce a Pre-Case Assessment report. Within minutes of receiving details about the case, this report can provide an assessment for how much the investigation will cost. These details include the number of records/pages of related file types, documents, emails, etc. belonging to the selected employees, as well as the costs of collecting, processing, and reviewing the data.

Analyze Targeted Data

The second step for the user would be to analyze the data that has been uncovered.



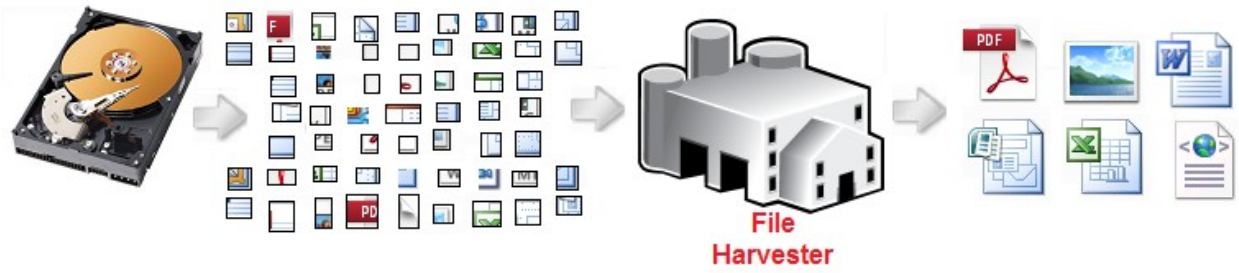
File Investigator will provide the additional details on every file that is analyzed. This includes the standard metadata from the operating system, as well as the details extracted from inside the file. When the edit dates do not match, you would want to know that. When the file extension is wrong, File Investigator will provide you with a list of possible extensions that should appear with that type of file. In addition to this data, there is also the background information that can

help you understand what this type of file is typically used for, and where it originated.

Switching the bottom preview pane between Details, Text, Hexadecimal, and Byte Value Distribution (BVD) mode lets you see the file in multiple ways, which can help to eliminate the file from your investigation or zero in on the content that you need to find.

Impossible Recovery

When a file is lost or deleted, and too much time has passed to undelete it, "Where does it go?". "What if you could get those files back?" Enterprise Profiler includes those files in its statistics. When you choose to focus on these lost files, we bring in the File Harvester to recover them. Based on our patent pending File Harvester technology, File Harvester automatically pieces these file fragments back together file



by file. "Why don't other companies offer this?" When files are fragmented, then lost or deleted, there are no links left to connect their fragments together. The task of putting those fragments back together is like sorting through a pile of jigsaw puzzle pieces from

hundreds of different puzzles. The task is daunting, and recovering any more than a few file types from fragmented space is nearly impossible. The current suite of available file/data carving tools on the market today produce scrambled results when they hit fragmentation. When recovering images, you may still be able to see parts of the pictures, but when used on documents and emails, the

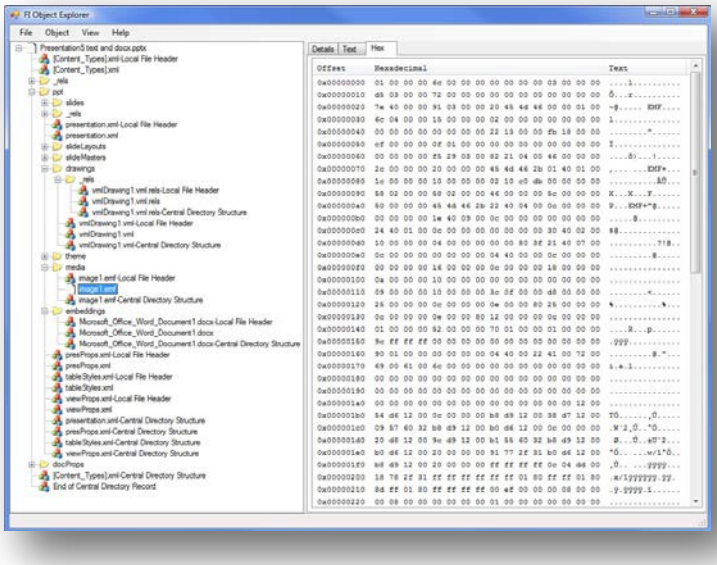
results are corruption that produces useless results. At FID[3], we have tackled this impossible task and can currently recover hundreds of different file types.

What this means to our customers, is that they will see files that no other solution on the market can produce. When an investigation fails to find the necessary evidence, this capability may be what saves the case. In a criminal case, where the defendant is likely to be hiding evidence, this would be the first place to look.

This is also the case for any threat perpetrator. The flat file used with an email application is one of the most fragmented files on a computer. If the computer used by the perpetrator is available then Enterprise Profiler would be able to recover emails even if they were fragmented, deleted and lost to the unallocated disk space.

Files within Files

What about when criminals hide their files within other files that look completely innocent? Would your investigation catch that? When an illegal image is stored within a Word document, and your search warrant only allows for the search of image files, would you catch that evidence? We developed Object Explorer to search through files looking for embedded file objects that do not belong, and data that has been deleted but not redacted. Just as you can have file slack

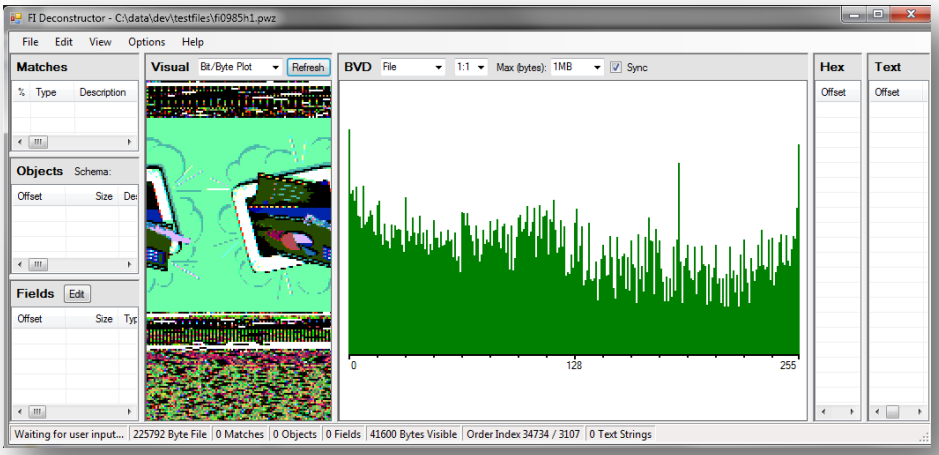


on a hard drive, there is also stream/object slack inside your files. A criminal does not even need to resort to steganography to hide their data in these places. In a simple 100KB Word Document, you can easily hide a 1 GB image file. The document still loads in MS Word with no errors. MS Word will not even be aware of the large embedded file. Historically, the only way to find this hidden data has been to hire an expert to walk through the file structure manually. With Object Explorer you can automatically extract all of the hidden areas. Today we are performing these operations on hundreds of popular types of files, and we continuously update the underlying technology to handle more and more types of files.

The Exception Bin

When the tools used in an investigation fail to identify files, they end up in the exception bin. Whether revealed to the user or not every product and service on the market has an exception bin. Even Enterprise Profiler has an unknown category, although it is an order of magnitude smaller than the other available market solutions. So, what do you do with these files? Typically, people try to ignore them. However, if your investigation comes up dry or if for some reason, you need to find everything possible, then you need to make a best effort in processing these files as well.

At FID^{3}, we encounter unidentified file types on a regular basis. That is how we find new file types to support, and how we have reached the point of supporting over 4,000+ types of files with the highest accuracy in the industry. We offer a powerful visual analysis tool to help investigators and reviewers analyze these difficult files. File Deconstructor combines our company's ability to identify files and objects with visual tools that establish whether each file requires further investigation. In the screen shot on the next page, there is a list of file type matches and object type matches provided on the left, which suggest what the file, or portions of the file, looks similar to. When you need to dissect a file down to its individual fields of data, the field organization is in the lower left. The Visual window provides multiple visualization methods, developed though academic research, to identify file types visually as well as where certain objects start and stop in the file. In the example above, the file analysis clearly shows that it

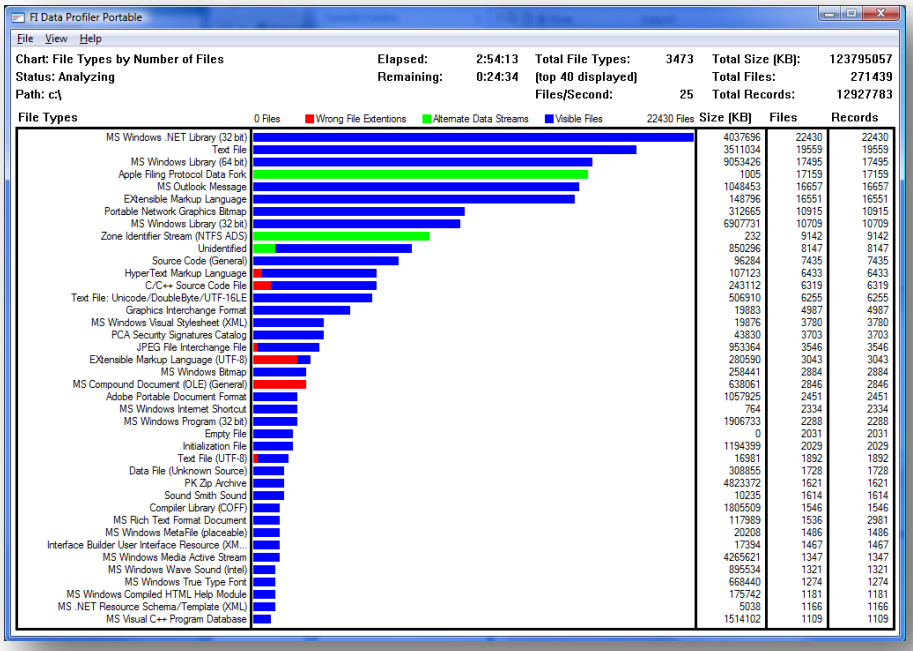


contains a graphic image. The garbled area above the image represents binary data that makes up the file's header. The multiple scrambled patterns, below the image, represent other objects in the file. These objects can be visually separated to help break down the structure of the file. Knowing where objects start and stop in a file is vital to being able to reverse engineer and interpret the file. The Byte Value Distribution (BVD) window provides a chart similar to an image histogram to further aid in identifying the file and/or each object in the file. This is how we identify files that cannot be identified with other products available today. For even deeper analysis, the Hexadecimal and Text view windows provide byte level access to the current object analysis.

First Responder Triage

Large powerful investigation platforms are not a good fit for the lean portable space of first responder triage. That is why we created Profiler Portable. Profiler Portable is a tool that runs from a thumb drive or CD in a forensically sound manner. While it shows a live chart of the progress it is making through a hard drive, it can be pre-programmed to log and collect files that set off red flags to represent anomalous behavior or might be pertinent to a particular investigation.

FI – DARK DATA DISCOVERY

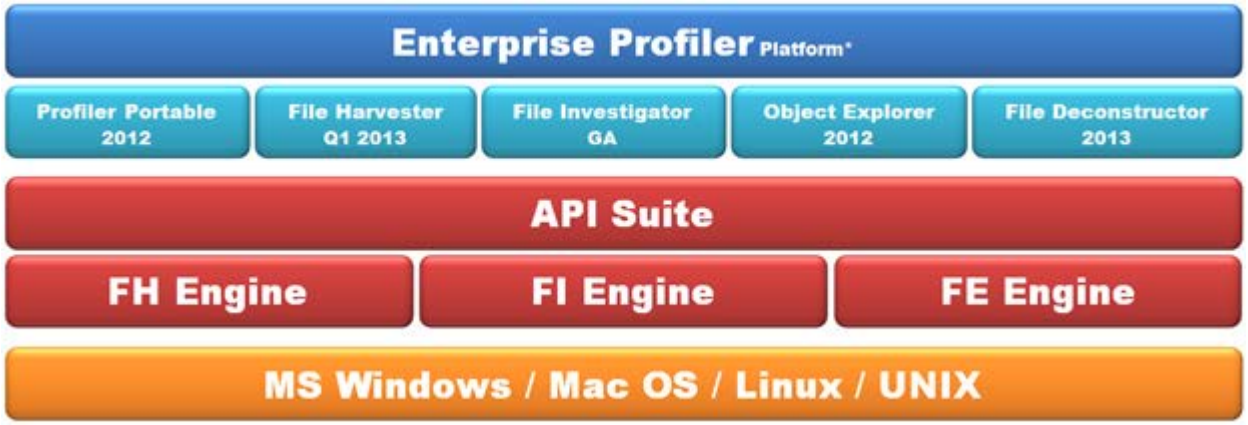


A first responder faces a time constraint that does not allow them to collect data from all of the computers found at the scene. With this tool, they can be collecting statistics on every computer simultaneously. After a short period, they will see which computers contain file types that may be pertinent to the case or are listed in the search warrant. The investigator can then focus on imaging or collecting from just the computers that are pertinent to the case at hand.

Logs and files collected with Profiler Portable are used by Enterprise Profiler, combined with the rest of the case files, and displayed along with all of the case findings.

Architecture

FID^3 built Enterprise Profiler on top of an existing proven architecture. Most software vendors start with an attractive user interface then gradually fill out the features in order to give their product functionality. They are rushed to get their product to market and skimp on features, figuring that they will catch up in functionality after the early adopters develop a revenue stream for the company. We chose a different more thoughtful route. Our underlying technologies were built first, and licensed to well-known discerning reputable companies. Over the years that followed, these companies helped to guide the development of these core technologies by requesting the features that would make them best in class.



The core development process has successfully completed the leading technology suite being used in Cyber Security, Electronic Discovery and Digital Forensics. Our Enterprise Profiler is built on top of this winning platform of technologies.

Conclusion

Today's data challenges span finding data that once existed but has since been deleted; whether it is simple easily recognizable file types or the more obscure, whether there has been attempts at concealing files within files or other means of obfuscation, and even whether or not the files were highly fragmented prior to deletion. The application of these findings has historically been focused on the eDiscovery or legal process related concerns. With FID^3, the use cases expand to the Information Assurance disciplines allowing the practitioner to understand when a user is collecting and storing sensitive data at their desktop; likely preparing to exfiltrate that data.

If your goal is to build the digital case to prosecute or to extend every effort to stop the loss of data altogether, FID^3 has solutions for your needs. In addition, they are built on a solid defined architecture that allows for the integration of the FID^3 capabilities into other popular platforms and ultimately integrating 3rd party and freeware capabilities into FID^3.

FI – DARK DATA DISCOVERY

Company Overview

FID3 has been delivering innovative software solutions since 2007. These solutions have uniquely solved complex technical challenges in the realm of computer forensics, electronic discovery and information assurance. The product's modular architecture provides for seamless bidirectional integration with existing technologies.

With the FID^3 building of a single platform for pre-case assessment and cutting edge Digital Forensics, we are not only automating Digital Forensics, we are bringing it into active eDiscovery usage in a seamless way. For example, when the user selects to see statistics on their data, we can provide the option of looking at visible files (as is typical of eDiscovery), hidden files (as is typical of Digital Forensics), or both combined. All this can start at a very high level and then drive down to departments, individual employees and individual files. We do not stop at finding just the files that the rest of the industry sees. Our technologies dig deeper to uncover the Dark Data. This is the data hiding in the unallocated file fragments, object slack fragments within files and less common file types that go unnoticed. Along with this level of intelligence comes the ability to profile company data down to individual user data, to provide another level of information assurance by automatically flagging anomalous behavior at the desktop. IT organizations and intelligence agencies care about these areas. Corporations are discovering that they need to be caring about these areas in their eDiscovery events as well.

Visit us on the web (www.FID3.com or info@FID3.com) or contact us directly (202-380-9449) to see what industry leaders and governments have already discovered in the dark.