

# CHILD PROTECTION:

## Streamlined Image Analysis Workflow with Nuix Investigator and NetClean Analyze DI

Child protection and law enforcement agencies must handle large and growing volumes of multimedia files in their efforts to catch perpetrators and identify and rescue child abuse victims. Through a partnership between software vendors NetClean and Nuix, agencies can now streamline the workflow for extracting, analyzing, categorizing, and investigating the sources of child abuse material, helping to identify and build a case against the perpetrators and their associates. This methodology eliminates duplication and manual steps in the investigative process, and works seamlessly with databases of known files to help child protection officers focus on the new and most relevant images and videos.

### Applying Technology to Defeat Child Abuse

To help police identify victims and perpetrators of child sexual abuse, law enforcement agencies on both sides of the Atlantic have used NetClean's Analyze Collaboration Server (CS) and the Hubstream Intelligence Server to create trusted databases of known child abuse multimedia files: Project VIC in the United States and the Child Abuse Image Database (CAID) in the United Kingdom.

These databases are useful because devices seized in evidence typically contain large numbers of known and unknown child abuse images and videos. Repeatedly reviewing known files is a wasteful and time-consuming ordeal for officers which is unlikely to help identify new perpetrators or victims.

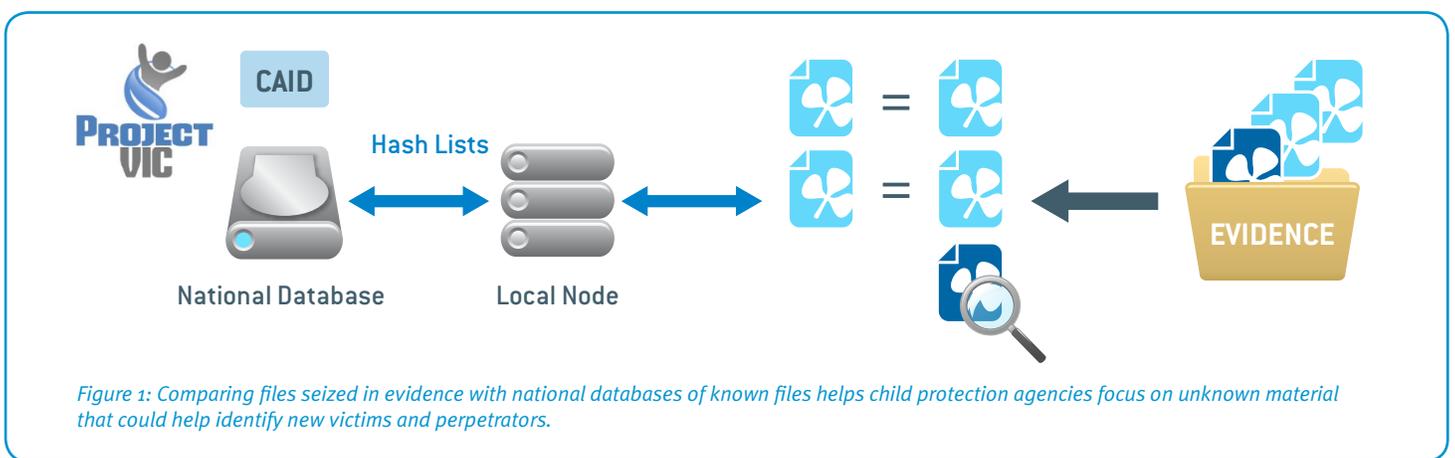
When investigators recover multimedia files from devices and online accounts they have seized from suspects, they can compare these with the known files in the Project VIC and CAID databases. This allows them to eliminate a large proportion of multimedia files they have seized and to focus on the remaining unknown digital media that may help them identify new victims and perpetrators. After identification, these files can be analyzed for new clues, classified, and added to the database for future reference.

The Project VIC and CAID databases hold large volumes of pertinent information including:

- The actual images and videos with their embedded exchangeable image file format (EXIF) metadata.
- Trusted MD5 and SHA1 cryptographic hashes of each item.
- Information about the cases and metadata concerning those files.
- Image fingerprints such as HYPERLINK "<http://blogs.microsoft.com/on-the-issues/2012/03/19/microsoft-photodna-technology-to-help-law-enforcement-fight-child-pornography/>" Microsoft PhotoDNA hashes, which analyze images based on their visual content, thus making it possible to match photos that have been resized or recolored.

Both databases use the victim-centric information exchange (VICS) protocol, based on the Open Data Protocol (OData), as their primary data exchange format.

The value of these databases is hard to question. As of early 2015, Project VIC has already helped identify and rescue more than 100 child abuse victims across the United States.



## NetClean Analyze DI

Analyze DI is specifically designed to simplify investigations involving digital media, eliminating duplicated efforts, reducing manual review to increase the impact of the investigations. By combining a rich toolset of technologies with automated processes to categorize and filter out non-pertinent material Analyze DI is able to reverse the trend of increasing caseload and turnaround times. This allows investigators to focus efforts on processing more cases, at a faster pace and with a greater outcome. The result is blazing fast investigations, without compromise.

## Nuix Investigator

Nuix's advanced digital forensics software is engineered to triage, process, analyze and bring to the surface critical evidence bridging entire data sets, regardless of the geographical location, repository, file type or size. Nuix Investigator can automatically identify and cross-reference key intelligence items such as email addresses and phone, social security and credit card numbers, eliminating the reliance on human memory and guesswork. Nuix Investigator is widely used in child protection investigations to improve efficiency and get fast results in a situation where every minute counts.

## ANALYZE DI AND TRADITIONAL FORENSIC TOOLS: A FRAGMENTED WORKFLOW

NetClean Analyze DI is a widely used tool for image and video analysis among law enforcement agencies globally. It provides a rich toolset for analyzing, reviewing, and managing visual data.

Analyze DI incorporates a wide range of image matching technologies, including Microsoft PhotoDNA. Microsoft PhotoDNA is a “robust hashing” technology used to identify images that are near duplicate, but not identical. It can determine if an image has been resized, recolored, compressed, slightly manipulated, or saved in a different format. Analyze DI incorporates other visual matching technologies to detect series of photos that have been taken in the same environment or that contain the same distinct attributes. Analyze DI can quickly process and analyze digital media stored as loose files or within forensic images.

However, to complicate matters, investigators regularly encounter evidence stored in smartphones, tablets, digital cameras, flash memory devices, and cloud email and storage services. Investigations into large organizations involve massive volumes of data in difficult-to-access formats such as file shares, email databases, email archives, and collaboration and document management systems.

To address this issue, investigators often start with forensic images of seized devices and run extraction scripts in forensic analysis tools to export multimedia material. Extracting images from Microsoft Outlook PST files, for example, can take hours or days of computing time for each evidence source using most forensic tools. Investigators then import the extracted files into Analyze DI for matching and analysis. They then work backwards and conduct further lengthy forensic analysis to identify the sources of the tagged files and conducting further lengthy forensic analyses to establish ownership, intent, distribution, and other aspects of the case.

As described, the traditional workflow is fragmented and long-winded, requiring multiple levels of processing in different tools. It also does not support sharing image tags and gathered or known intelligence from one tool to another.

Officers can use the advanced capabilities in Analyze DI to eliminate visually identical or very similar files, and then focus on analyzing and classifying the unknown multimedia files

## STREAMLINED PROCESS WITH NETCLEAN AND NUIX

The integration between NetClean Analyze DI and Nuix Investigator solves these problems and creates a much more powerful and streamlined investigative workflow. Nuix can extract multimedia files from a huge variety of data formats and sources, including:

- Hard drives from laptops, PCs, and servers
- Solid-state storage devices including USB sticks and flash memory cards
- Cellebrite, Micro Systemation XRY, and Oxygen Forensic mobile device forensic images
- Hard drive forensic images
- Personal email database files such as Microsoft Outlook .PST and .OST files
- Enterprise email database files such as Microsoft Exchange .EDB files
- Cloud email and storage services
- Corporate repositories such as Microsoft SharePoint, email archives, and file shares.

Nuix's patented parallel processing technology can index and extract terabytes of data per day on a single high-end server, while interrogating all metadata, text, and binary structure of documents to forensically identify relevant information. Nuix Investigator's flexible settings make it possible to focus on extracting multimedia files at maximum speed.

Having extracted these digital media, child protection officers next eliminate known images and videos by comparing the extracted files with the lists of hashes they have access to.

Officers can then export the remaining unknown multimedia files into Analyze DI. They can use the advanced capabilities in Analyze DI to further cull the data set by eliminating visually identical or very similar files, and then focus on analyzing and classifying the unknown multimedia files.

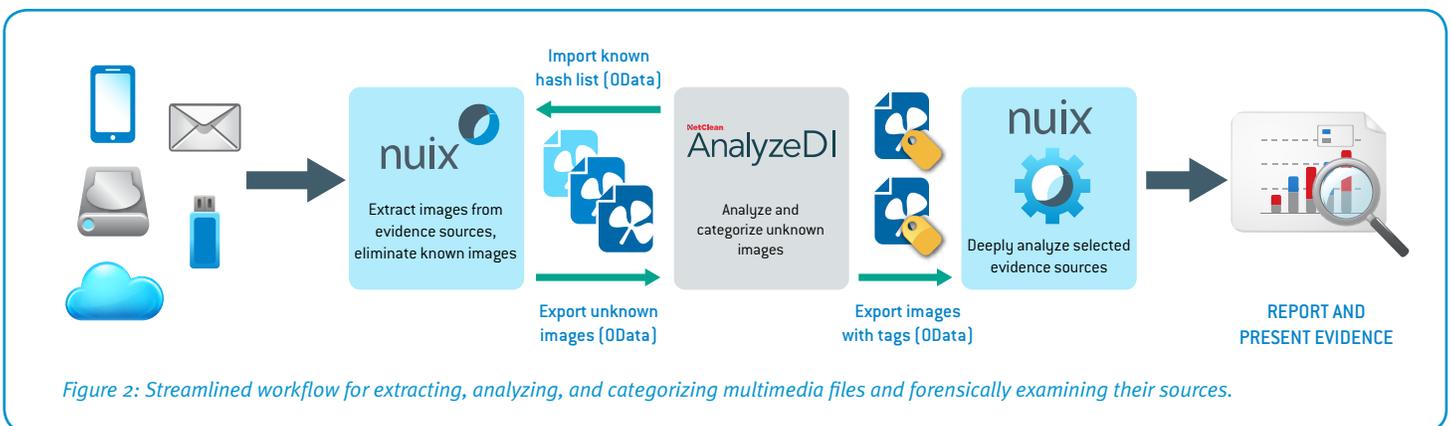
Once this process is complete, officers can export their classifications back into the Nuix Investigator case file. This will allow them to immediately:

- Identify the evidence sources that contain multimedia files of interest
- Reload these sources with greater forensic depth
- Analyze the text, communication patterns, browsing histories, and forensic artifacts for vital clues about the identity and location of victims and perpetrators.

Nuix's contextual analysis tools such as network diagrams can help demonstrate that a suspect distributed child abuse material, which is a more serious crime than possessing it.

## The Investigative Process at a Glance

- Extract multimedia files from seized evidence sources using Nuix Investigator
- Eliminate known files by comparing the extracted material to hash sets of existing digital media
- Export unknown files into NetClean Analyze DI, then analyze, categorize, and tag them
- Export the images back into Nuix Investigator along with their tags
- Deeply examine selected evidence sources for clues about the identity and location of victims and perpetrators
- Generate reports and present the evidence to relevant authorities



## ADVANTAGES OF NUIX

Nuix Investigator automatically extracts and cross-references intelligence such as names, email addresses and credit card numbers; and visually represents connections between suspects and events, places, and objects.

Visually analyzing data using Nuix's Web Review & Analytics application makes it much easier to detect trends and isolate outliers across massive volumes of evidence from multiple sources and formats. Nuix also helps to reduce repetitive tasks by automating workflows using scripts or the Nuix Director web application.

Analyzing text, communication patterns, browsing histories, and forensic artifacts can provide vital clues about the identity and location of victims and perpetrators

## FUTURE DEVELOPMENTS AND BROADER APPLICATIONS

Nuix Investigator already uses OData for its integration with NetClean Analyze DI. The software vendor is working with Project VIC and CAID to develop the ability to import their known image databases without third-party software.

While the focus of this integration has been on child exploitation material, investigators can apply a similar workflow to any matter that requires analysis and categorization of multimedia files.

---

### ABOUT NUIX

[nuix.com](http://nuix.com)

Nuix enables people to make fact-based decisions from unstructured data. The patented Nuix Engine makes small work of large and complex human-generated data sets. Organizations around the world turn to Nuix software when they need fast, accurate answers for digital investigation, cybersecurity, eDiscovery, information governance, email migration, privacy, and more.

### ABOUT NETCLEAN

[netclean.com](http://netclean.com)

NetClean provides intelligence solutions to detect, block and analyze digital media to create a safer society. We are the leading developer of technical solutions to fight child sexual abuse material. Our solutions are being used worldwide by multinational companies, government agencies, Internet service providers and law enforcement professionals.