INTELLIGENCE, COLLABORATION AND ANALYTICS FOR DIGITAL INVESTIGATIONS

By Paul Slater
Director of Forensic Solutions, Nuix

nuix
Simple. Powerful. Precise.

CONTENTS

Nuix has advocated for many years that investigators need to evolve beyond traditional forensic tools and workflows, so they can efficiently examine the contents of multiple evidence sources at once.

EXECUTIVE SUMMARY

Investigators must deal with large and growing volumes of digital evidence across an increasing number and variety of sources. Criminals and wrongdoers have grown skilled at using technology to conceal their activities. We would argue that some are more effective at covering their tracks than investigators are at applying technology to uncover them.

Nuix has advocated for many years that investigators need to evolve beyond traditional forensic tools and workflows, so they can efficiently examine the contents of multiple evidence sources at once. But just as the key facts may not be located within a single evidence source or connected to just one person, they may not even be in the same investigation, or the same agency, or the same country. As a result, efficient investigation must enable people to share intelligence, to collaborate across geographic and jurisdictional boundaries, and to find seemingly hidden connections across very large numbers of evidence sources.

Technology has stood in the way of these vital abilities. Investigative technologies have burrowed further and further down the rabbit hole of deeply examining single evidence sources. They can tell you everything you need to know about the binary structure of data on a hard drive, but nothing about how the instant message history stored in that data connects with a mobile phone seized in another investigation on the other side of the country.

This paper will examine technology-enabled processes for making those connections. It will discuss:

- Automatically extracting intelligence items such as email addresses and credit card numbers, correlating them across all available evidence sources, and sharing this information efficiently with other investigators.
- Providing a way for multiple investigators, subject matter experts and external agencies to review and collaborate on the evidence you have collected.
- Applying and combining data visualisations and analytics to progress rapidly from a bewildering array of information to highly relevant details.

In this way, investigators can apply technology where it is most suited, free themselves from tiresome menial work and make best use of their brainpower and intuition.

DIGITAL EVIDENCE DOESN'T SHARE ITS SECRETS EASILY

Investigators face many challenges when dealing with digital evidence. Modern telecommunications technologies make it easier for criminals to operate across jurisdictional and national borders, hide their activities and evade detection and prosecution. To combat them, law enforcement agencies need an efficient legal and technological framework to exchange intelligence.

Law enforcement agencies must also deal with large and growing volumes of data in an expanding number of devices. In addition to computers and mobile phones, potential evidence sources may include cloud email, social media, digital cameras, MP3 players, smart household appliances, GPS devices or even smart wearable devices such as eyeglasses or watches.

Large-scale investigations in areas such as counterterrorism and organised crime may involve data from multiple suspects, each with up to a dozen potential evidence sources.

As we have discussed, the traditional linear methodology of forensically examining each data source individually can never hope to keep up. The combination of slow forensic tools and case backlogs mean that by the time forensic technicians examine an evidence source, it may be months old. By this stage much of its intelligence value may be lost.

A lack of shared intelligence

For many years, Nuix has advocated an investigative methodology that makes it possible to examine and cross-reference multiple evidence sources at the same time. However, it is not uncommon for crucial information to reside outside the evidence gathered for a specific investigation. It may be in a previous or concurrent investigation conducted by the same personnel or someone else. It may be from a different agency, office, location or country.

According to the UK Association of Chief Police Officers, "Intelligence management involves linking information from a wide range of sources in order to build up a composite picture. It aims to highlight links between people, objects, locations and events that are essential in supporting [policing purposes]. Identifying links enables decisions to be made about priorities and resources needed to manage risk."

Many law enforcement agencies are acutely aware how important it is to share intelligence internally and externally. But these efforts often fail at a practical level. For example, more than a decade after the US Government's 9/11 Commission Report made a series of recommendations on sharing intelligence information between agencies, there are still many technical and procedural barriers to effective counterterrorism intelligence sharing in the United States.

Using traditional methods, investigators struggle to compare information between individual evidence sources in a single investigation. Sharing intelligence across multiple investigations and between agencies is an even more manual and labour-intensive process.

Technology tools an impediment to collaboration

Another major impediment to efficient investigations is the difficulty investigators face making digital evidence available for review to internal or external personnel. Investigative technology vendors have tried to solve this problem by bolting legal review platforms onto forensic investigation tools, or adding forensic processing and analysis capabilities to an existing review platform.

These tools are often cumbersome to set up. Because they come from a legal review background, where most evidence is stored in email and documents, they are very text-centric. They have poor abilities to examine multimedia such as photos and video, and social interactions including phone call logs, SMS and instant messages, Skype chat and call logs, and browser histories and caches.

Inability to find hidden connections

In investigative organisations, forensic technicians often work in isolation from case investigators and other stakeholders. The forensic specialists must make critical decisions about which data sources to examine, how to process them and what information to extract – often without knowing the broader details of the case. This approach can only be helpful if the case investigators and forensic technicians have a clear understanding of what to look for.

The connections between people, objects, locations and events can be critical in providing intent or collusion, but often they are not immediately obvious. It would take superhuman skill to mentally correlate connections from a single suspect's hard drives, mobile devices, instant messages, cloud email, cloud storage and social media interactions. Multiply this by the number of suspects in an investigation and technology becomes the only answer.

Under such circumstances, the ability to visually represent and analyse data can be a rapid shortcut to locating the key facts and connections of the case. But most investigative tools have limited abilities to visualise data, especially across multiple evidence sources.

USING TECHNOLOGY TO AUGMENT HUMAN BRAINPOWER

A common theme behind the problems we have discussed is that they require considerable human effort to conduct tasks that computers are much better at.

This situation has evolved because the developers of investigative tools have focused on increasing the forensic depth with which they can analyse individual evidence sources. However, solving crimes very often requires finding the connections across multiple individuals, places, events and evidence sources. Human intuition has its place in the process, no doubt. But much of the laborious work involves picking out and matching specific pieces of information from massive volumes of data.

Computers, when applied judiciously, have a natural advantage in intelligence sharing, collaboration and data visualisation. Here's how you can apply technology in the right places to assist human investigators.

Intelligence

Using the traditional digital investigation model, investigators must manually compare intelligence items across each evidence source. Even something as simple as proving person A and person B both used the same stolen credit card number is a complex matter or identifying credit card numbers in their various evidence sources, printing out lists of those numbers and then poring over those lists to find the matches. Even semi-automating this process, by exporting the data sets into a database and running comparisons, is time consuming and error prone.

Nuix's advanced investigative tools use a 'named entities' model to extract intelligence items that follow a particular pattern of letters and numbers. By default, this list includes:

- Names
- Email addresses
- IP addresses
- Company names
- Credit card numbers
- Bank account numbers
- Social security or identity numbers
- Amounts of money.

This list is not comprehensive. Investigators can easily define their own named entities, using regular expressions, to extract locally relevant or investigation-specific types of information such as passport, phone, vehicle identification or contract numbers.

Having identified relevant intelligence items, investigators can instantly see which suspects have those items in common, across all the evidence sources in the entire case. Typically they can also identify who shared what, with whom and when, using techniques such as timelines and network diagrams.

This ability to extract lists of relevant named entities makes it simple to compile and share libraries of intelligence related to investigations. It is an extremely rapid and automated way to uncover hidden connections between multiple evidence sources, people, locations and investigations.

This ability to share intelligence works in a number of ways:

- Agencies can compile lists of relevant names, email addresses or other intelligence items and search other evidence for those lists. For example, police investigating a suspected terror cell in Birmingham can provide a list of suspect phone numbers, email addresses, online aliases and bank account numbers to their colleagues in Manchester. The Manchester investigators can run this list through their existing case files to see if any connections emerge.
- Investigators can build lists of relevant words and phrases within case material such as documents or illegal or inflammatory propaganda material, and share these word lists with other agencies. Using near-duplicate analysis, they can identify identical and similar items within different evidence collections held in completely separate investigations.
- Agencies can share entire investigations by creating compound cases of separate investigation files. Investigators can then see the connections between intelligence items across multiple investigations.

These techniques also alleviate the issue of the delays in extracting intelligence using traditional methods. As soon as investigators uncover relevant and timely intelligence, they can run this across other potentially relevant cases and very rapidly find and delve into any matches.

USING TECHNOLOGY TO AUGMENT HUMAN BRAINPOWER cont

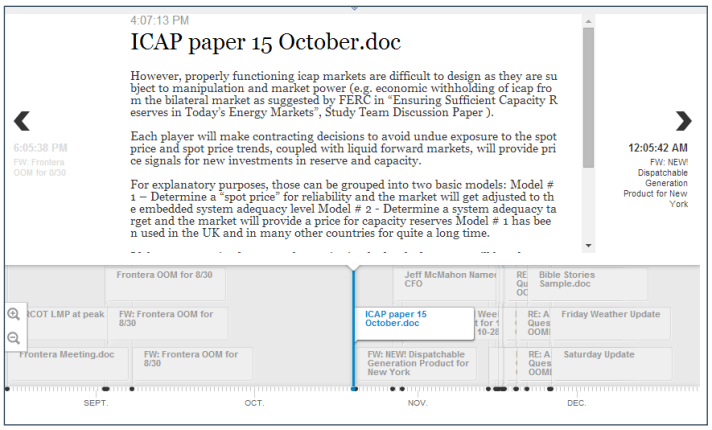

Figure 1: A timeline of events in Nuix Web Review & Analytics.

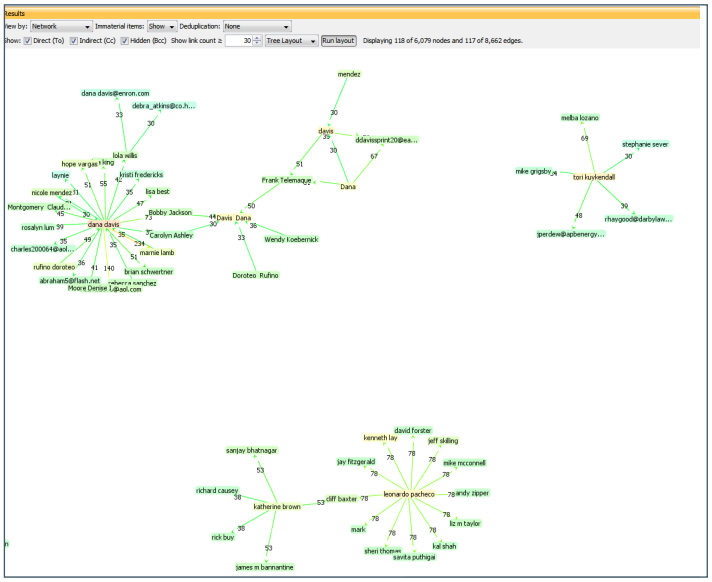

Figure 2: A network diagram showing the number of connections between suspects.

A growing number of Nuix customers are using our Web Review & Analytics solution to extend this collaborative model to tens or hundreds of investigators across multiple locations.

Collaboration

In our Investigative Lab white paper, Nuix discussed an operating model that enabled investigative teams to divide up digital evidence and spread the review workload between many people. At a basic level, this is a way to share work between multiple investigators to complete the task faster. They may choose to divide the evidence by date ranges, custodians, location, language or content.

It can also be a way to distribute different types of evidence to the people most qualified to understand it and its context. For example, investigators could pass on financial records to forensic accountants and internet activity to technical specialists. In an inappropriate images investigation, detectives could package potentially relevant pictures and videos for specialist child protection teams, while leaving other file types for their digital forensic investigators. In multi-jurisdictional investigations, investigative teams can produce evidence or intelligence packages for third parties to review, comment on and return.

A growing number of Nuix customers are using our Web Review & Analytics package to extend this collaborative model to tens or hundreds of investigators across multiple locations. Its simple interface enables people with minimal training or technology expertise to search, review, tag and analyse data from any web browser. Role-based security controls ensure you can make evidence easier to access while protecting confidential and sensitive information.

Larger law enforcement agencies, advisory firms and enterprises are using this model to set up centralised evidence processing facilities that can provide access to the results to any desktop across the organisation.

This model has considerable advantages for sharing intelligence. A centralised lab which stores case files related to all current investigations makes it easy to cross-reference intelligence items. It is also easy for one location or agency to provide in-depth access to their case data for colleagues in other locations.

Another useful technology for setting up a lab is Nuix's Director browser-based application, which enables organisations to define templates and workflows for processing data consistently and repeatably. Using Nuix Director, investigative organisations can ensure they handle evidence consistently, from one investigation to the next and across multiple locations. Adopting a templated workflow can help forensic laboratories show compliance with the quality assurance and testing accreditation schemes under which many operate.

USING TECHNOLOGY TO AUGMENT HUMAN BRAINPOWER cont

Analytics

As digital evidence becomes larger and more complex, investigators' greatest struggle is not a lack of information, but having too much to make sense of. Visually representing large volumes of data can be a fast way to locate the key facts and connections within the case. It enables people, even with limited technical knowledge, to follow a hunch or idea down to very specific details in a matter of seconds.

Common analytical techniques include:

- **Top types.** Quickly understanding the makeup of data sets by showing the most common file types as bar or pie charts.
- **Pivot**. Analysing the relationship between any two elements in a data set including custodians, file extensions, file types, languages, named entities, tags and word lists.
- **Date trending.** Visualising the frequency of data over the entire case or any filtered subset, then drilling down to year, month or day views.
- **Timeline.** Reviewing the content of emails, documents, phone calls or other communications from multiple sources or custodians in the order they happened.
- **Communication network.** Showing the interactions between persons of interest with an interactive network diagram that shows the number of connections for each link.
- **Link analysis.** Understanding the connections between people and intelligence items such as credit card numbers, IP addresses, organisations and sums of money.
- **Intersection.** Rapidly understanding how key elements in the data overlap and pinpointing the critical intersections between multiple result sets and data types.
- **Shingle and word lists.** Rapidly understanding the key words and phrases – and their context – in the case.

New grouping and filtering functionality in the latest release of Nuix Investigator instantly collates all available evidence of a particular type in the data set (see Figure 3). These types include browser history, cache and bookmarks, SMS messages, mobile calls, Skype calls and messages, USB devices, log files and Windows Registry entries.

Combining analytical techniques can help investigators progress from a bewildering array of information to highly relevant details very quickly. For example, you could filter an entire evidence set to just email messages within a relevant date range that contain credit card numbers. If that still returns too many results, you could use other techniques such as suspect names or keyword searches to further filter the evidence. Now you can use a network diagram to see who is emailing credit card numbers to whom.

Link analysis uses technology to replace the manual process of finding connections between suspects and evidence sources. It automatically tallies and displays connections between people and named entities such as credit card or phone numbers. When applied across a compound case containing multiple case files, link analysis has proven particularly effective in finding connections between seemingly unrelated people and events.

A timeline view, traditionally used for email messages, is also useful for SMS messages, mobile device call logs, instant messages, Skype chats and social media messages. In my experience, many people say things in instant messages that they would avoid in email. This may stem from the belief that these formats are not as rigorously logged as email. But from the investigator's perspective, advanced technologies make these communication formats just as permanent and searchable as email.

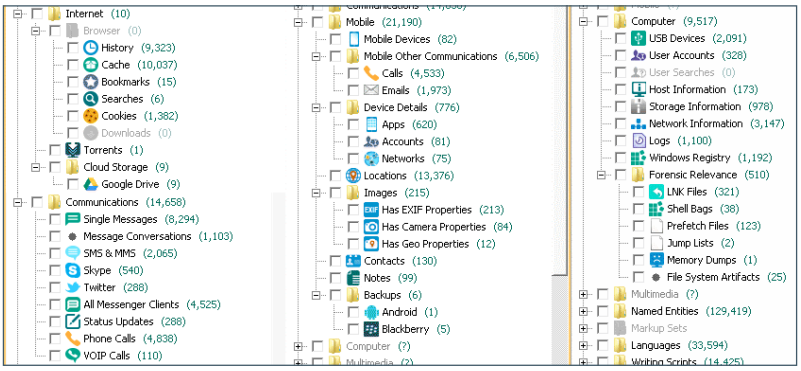

Figure 3: Filters in Nuix Investigator group and combine data across internet, communications, computer and mobile categories.

USING TECHNOLOGY TO AUGMENT HUMAN BRAINPOWER cont

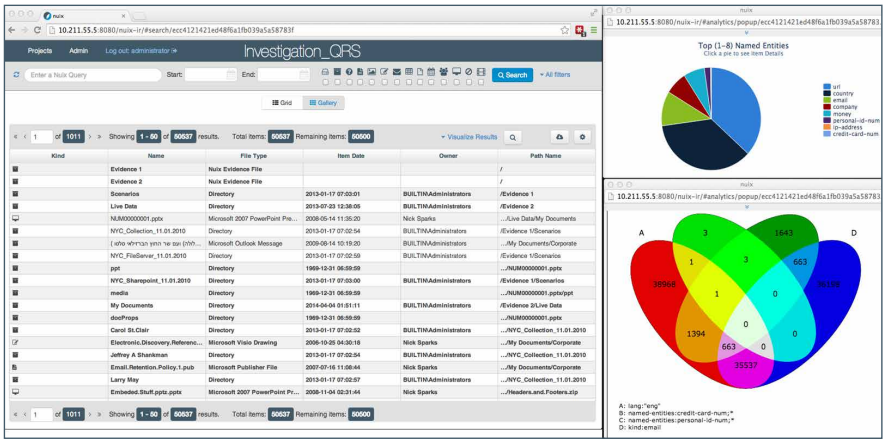

Figure 4: Using a combination of top types and intersection visualisations in Nuix Web Review & Analytics.

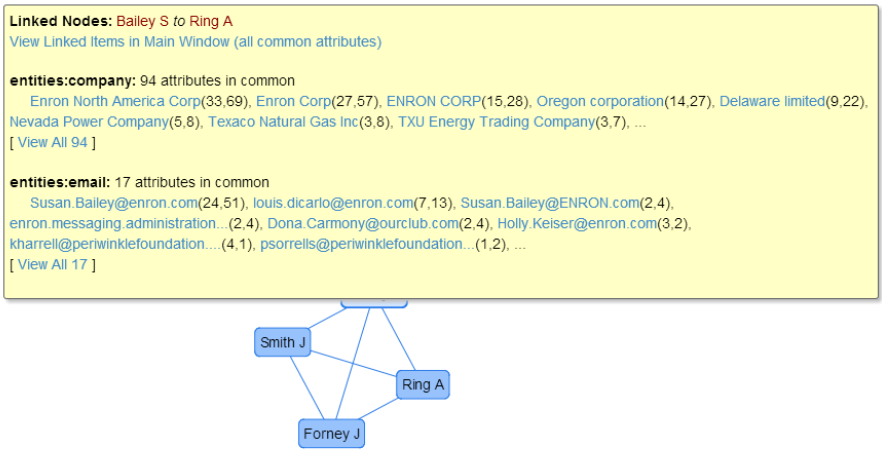

Figure 5: Link analysis shows the connections between individuals and named entities.

But what about forensics?

Investigators and forensic technicians reading this paper may be asking themselves, “But what about forensics – will any of this stand up in court?” The techniques in this paper do not eliminate the need for forensic analysis, particularly in the areas of provenance and authenticity.

However, as we have previously argued, the volume of evidence in most cases makes it too time-consuming to conduct deep forensic analysis on every data source. As a result, in-depth forensic analysis must become the exception rather than the rule.

The techniques in this paper are a much faster and more efficient way of identifying the evidence sources that contain the data required to prove or disprove the case. The investigative team can then pass a small number of evidence sources back to digital forensics specialists so they can conduct in-depth analysis that will satisfy courts and authorities.

It is also worth noting that a centralised lab using Nuix Investigator alongside Web Review & Analytics gives investigators a ‘single pane of glass’ view of all the data within an investigation. This set-up also removes the need to convert and move data between formats and tools during the investigation process. It is therefore much easier to maintain provenance and trace critical evidence identified during the investigation back to its original source.

Combining analytical techniques can help investigators progress from a bewildering array of information to highly relevant details very quickly.

ABOUT THE AUTHOR

Paul Slater
Director of Forensic Solutions, Nuix

Paul Slater has over 20 years' experience in investigations, digital forensics and eDiscovery as a police officer and consultant. Slater has an MSc in Computer Forensics and started his career in forensic technology as a computer forensic investigator in the UK's Greater Manchester Police. Slater has been a senior manager within PwC's and Deloitte's regional UK Forensic Technology teams and has served as interim head of the Digital Forensics Unit in the UK's Serious Fraud Office where he implemented workflows that enabled them to process 20 times more electronic evidence each year. Slater was also a member of the review board for the 2012 update of the UK Association of Chief Police Officers' Good Practice Guide for Digital Evidence.

To find out more about Nuix digital investigations visit

nuix.com/investigation

ABOUT NUIX

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. Organisations 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.

APAC
Australia: +61 2 9280 0699
» Email: sales@nuix.com

North America
USA: +1 877 470 6849
» Web: nuix.com

EMEA
UK: +44 207 877 0300
» Twitter: @nuix

nuix