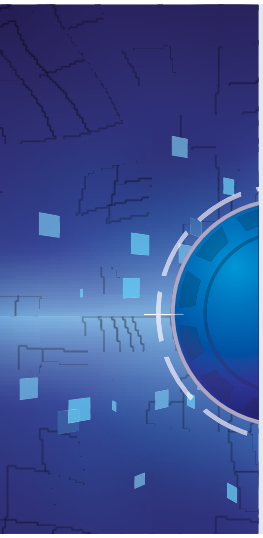


TRIAGEM FORENSE BASEADA EM CONTEÚDO: GERENCIANDO AS INVESTIGAÇÕES DIGITAIS NA ERA DO BIG DATA

O rápido crescimento, movimentação e alteração dos dados tornaram as ferramentas e metodologias da informática forense insustentáveis. Uma abordagem emergente que prioriza as evidências eletrônicas com base no seu conteúdo e contexto pode aumentar significativamente a eficiência para investigadores com tempo e recursos limitados.

nuix

Simple. Powerful. Precise.

CONTEÚDO

Sumário executivo 3
O crescimento de dados ultrapassando as capacidades de investigação 4
Desafios para os investigadores do setor público 4
Aumento do volume e variedade 4
A crescente complexidade dos dados corporativos 4
Pendências nos casos 4
Abordagem tradicional de investigação digital 5
Limitações desta abordagem 5
Estudo de caso: investigação de vazamento de dados 9
Triagem forense baseada em conteúdo – um método mais eficiente 6
Assimilar todos os dados 6
Realizar uma varredura leve de metadados 6
Analisar as relações básicas entre pessoas e evidências 7
Indexar profundamente as fontes de dados relevantes 7
Pesquisar e investigar 7
Inteligência de referência cruzada 7
Verificação forense apenas para as fontes de dados mais relevantes 7
Técnicas avançadas de análise de conteúdo 7
Recursos do Nuix – do Big Data até pequenos detalhes forenses 8
Estudo de caso: investigação de informações privilegiadas 9
Casos de uso comum 10
Estudo de caso: investigação de fraude multijurisdicional 10
Benefícios para os investigadores 11

SUMÁRIO EXECUTIVO

A era do Big Data é um ponto de crise para investigações e informática forense. O rápido crescimento dos dados tornou as ferramentas e metodologias da informática forense insustentáveis.

A lista de dispositivos que contêm armazenamento eletrônico e possíveis evidências em uma casa, cresce a cada ano. As investigações digitais regularmente encontram evidências armazenadas em smartphones, tablets, câmeras digitais, dispositivos de memória flash e serviços de armazenamento em nuvem. Quando as investigações envolvem grandes organizações, elas devem também lidar com grandes volumes de dados em formatos de difícil acesso.

Apesar destas mudanças, muitos investigadores seguem firmemente o método tradicional de analisar cada repositório de dados individualmente usando ferramentas forenses e, em seguida, manualmente, correlacionar as evidências que descobriram. Esta abordagem tornou-se imensamente demorada, levando a uma grande lista de pendências de casos não resolvidos.

Nos últimos anos, temos visto uma abordagem diferente por parte da lei e investigadores corporativos, que atinge os mesmos ou melhores resultados que os métodos tradicionais, mas muito mais rapidamente.

A triagem forense baseada em conteúdo envolve a coleta de todos os dados disponíveis em um único local de armazenamento e, em seguida, usando uma combinação de gerenciamento de dados, análise e técnicas forenses, ser possível entender o conteúdo e contexto das evidências digitais. Isto faz com que seja possível focar rapidamente nas as fontes mais importantes até que a evidência-chave surja.

Essencialmente, os investigadores precisam apenas usar a demorada análise de dados forenses se outros métodos, mais rápidos, falharam em descobrir as evidências que procuram. Isso funciona porque, na grande maioria dos casos, a evidência crítica está escondida à vista de todos, ao invés de em artefatos forenses, nos quais os investigadores dispendem tanto esforço analisando.

Apenas um aplicativo disponível atualmente pode gerenciar o processo de triagem forense baseada em conteúdo, do início ao fim. Isto torna o Nuix uma alternativa superior aos antigos aplicativos de informática forense.

O Nuix tem capacidades incomparáveis para lidar com grandes e complexos conjuntos de evidências, com poderosas ferramentas de pesquisa, análise e gerenciamento de dados. Ele pode destacar automaticamente itens de inteligência, tais como nomes, endereços de e-mail e números de telefone e cartão de crédito em várias fontes de dados, e os investigadores podem facilmente transferir listas destes itens entre investigações relacionadas.

O Nuix fornece contexto e análise mais profundos que uma busca de palavra-chave, permitindo que os investigadores comparem conjuntos de documentos relacionados e agrupamentos comuns de palavras. Ele também executa uma detalhada análise forense em arquivos apagados, espaço livre, registro do Windows e imagens de dispositivos móveis.

A triagem forense baseada em conteúdo economiza tempo e esforço consideráveis. Ela permite que os investigadores sobrecarregados e órgãos da lei com poucos recursos encontrem respostas mais rapidamente e consigam grande progresso em seu acúmulo de casos.

Na grande maioria dos casos, a evidência crítica
está escondida à vista de todos, ao invés de em artefatos forenses.

O CRESCIMENTO DE DADOS ULTRAPASSANDO AS CAPACIDADES DE INVESTIGAÇÃO

A era do Big Data impões dificuldades significativas para os investigadores nos setores privado, de aplicação da lei e de regulamentação.

De acordo com a empresa analista de tecnologia Gartner, o termo frequentemente usado "Big Data" não remete simplesmente aos grandes volumes de informação. Pelo contrário, significa "recursos de informação de alto volume, alta velocidade e variedade, que demandam formas inovadoras e rentáveis de processamento de informações para uma melhor percepção e tomada de decisão."[i]

Estes dados, sempre em crescimento, movimento, mudança e tornando-se mais complexos, têm forçado a maioria dos investigadores até sua capacidade máxima.

Desafios para os investigadores do setor público

Para os órgãos da lei e reguladores, as evidências eletrônicas apresentam uma série de desafios estratégicos, incluindo:

- A velocidade e o alcance da internet fazem com que seja mais fácil para os criminosos operar internacionalmente e fugir da regulamentação; para combatê-los, os órgãos da lei devem ter um quadro jurídico e tecnológico eficiente para a troca de informações e evidências
- Os órgãos de aplicação da lei devem lidar com grandes e crescentes volumes de dados, e em muitas jurisdições, os provedores de internet não são obrigados a conservar os dados para uso em investigações policiais
- O vasto volume de dados e os custos do seu armazenamento levarão os órgãos da lei a tomar decisões difíceis sobre quais informações manter e como gerenciá-las para garantir que elas permaneçam acessíveis.[ii]

Aumento do volume e variedade

Como regra geral, o número de aparelhos que contém dados envolvidos em uma investigação típica dobra a cada dois anos, e o volume de dados cresce ainda mais rapidamente.

A empresa de análise de tecnologia da informação IDC estima que, em 2012, o adulto médio no mundo desenvolvido gerou 1,8 terabytes de dados por ano.[iii] No entanto, o total de "informação ambiental" no universo digital sobre cada pessoa estende-se para 4,1 terabytes.

Quanto à variedade, o mais recente guia de boas práticas para evidências eletrônicas computadorizadas da Associação de Chefes de Polícia do Reino Unido recomenda que os policiais na cena do crime devem apreender os dispositivos, incluindo PCs ou laptops, discos rígidos externos, *dongles*, modems, placas de rede sem fio, roteadores, câmeras digitais, disquetes, fitas de backup, cartuchos Jaz/Zip, CD-ROMs, DVD-ROMs, placas PCMCIA, pen drives, cartões de memória e todos os dispositivos conectados através de USB ou FireWire.[iv]

Um número crescente de dispositivos ao redor da casa pode armazenar grandes quantidades de dados, incluindo smartphones, tablets, decodificadores de tv, reprodutores de DVD e mídia, e até mesmo eletrodomésticos, como geladeiras.

A crescente complexidade dos dados corporativos

Além de aumentar em volume e variedade, as evidências digitais estão se tornando mais complexas. Para investigadores e reguladores que trabalham em ambientes corporativos, as evidências podem ser armazenadas em compartilhamentos de arquivos, bancos de dados de e-mail, arquivos de e-mail, sistemas de colaboração e gerenciamento de documentos, entre outros.

Estes repositórios têm formas complexas de armazenamento e incorporação de vários níveis de profundidade de dados. Eles frequentemente usam formatos proprietários fechados que normalmente requerem uma interface de software fornecida pelo fabricante para a leitura da informação dentro deles.

O uso crescente de armazenamento e serviços em nuvem acrescenta outra camada de complexidade, especialmente porque os provedores de nuvem frequentemente movem os dados entre uma jurisdição e outra - mesmo que apenas por razões de eficiência.

Pendências nos casos

Encontrar a "pistola eletrônica fumegante" muitas vezes requer a busca e cruzamento de dados entre várias fontes de dados muito grandes com os quais os métodos de investigação tradicionais não conseguem lidar. Como resultado, os investigadores enfrentam uma crescente acumulação de casos e rotineiramente se afastam das potenciais evidências, pois eles não têm os recursos para processar os dados.[v]

Encontrar a "pistola eletrônica fumegante" muitas vezes requer a busca e cruzamento de dados entre várias fontes de dados muito grandes

ABORDAGEM TRADICIONAL DE INVESTIGAÇÃO DIGITAL

Ao manipular evidências eletrônicas, a maioria dos investigadores aplica metodologias e ferramentas forenses tradicionais.

Eles usam um aplicativo forense digital para examinar cada repositório de dados – tais como discos rígidos, dispositivos móveis e de memória flash. Para cada dispositivo, eles normalmente:

- Conectam o dispositivo a um bloqueador de gravação
- Geram uma imagem forense de todo o dispositivo
- Fazem uma cópia da imagem forense
- Analisam os dados armazenados na cópia da imagem forense
- Escrevem um relatório sobre os resultados desta análise.

Um investigador, em seguida, repetiria este processo para cada dispositivo relacionado com este caso. Tendo completado este processo para todos os dispositivos, os investigadores então usariam a inteligência humana para encontrar conexões e correlações entre as fontes de dados.

Limitações desta abordagem

Esta abordagem tem uma série de limitações, particularmente levando-se em conta o número crescente de dispositivos e volume de dados que os investigadores devem examinar. As ferramentas forenses tradicionais:

- Apenas podem analisar efetivamente um repositório por vez
- Não podem analisar cuidadosamente armazenamentos complexos de informação, como Lotus Notes, Microsoft Exchange, Microsoft SharePoint e arquivos de e-mail
- Têm dificuldade para processar grandes volumes de dados em um tempo razoável (os conjuntos de dados agora geralmente atingem centenas de gigabytes ou terabytes, que as ferramentas forenses podem não ser capazes de processar em absoluto)
- Não identificam automaticamente e organizam informações importantes, tais como nomes, endereços de e-mail, números de telefone e números de cartão de crédito - os investigadores têm de saber o que estão procurando.

Investigadores, apesar de brilhantes, não podem esperar de forma consistente e precisa cruzar e encontrar correlações entre milhões de pontos de dados. É fácil perder conexões, particularmente sem uma forma automatizada para identificar itens de inteligência.

Como resultado, a investigação de mídia digital de “pontos para provar” ou “elementos do crime” é a norma. Os investigadores raramente têm o luxo de uma análise de nível mais elevado.

ESTUDO DE CASO:
INVESTIGAÇÃO DE VAZAMENTO DE DADOS

Uma grande empresa identificou seu call center como a fonte de dados de cartões de crédito que vazaram, mas os investigadores não conseguiram localizar o ponto fraco do sistema. Usando o software Nuix para analisar os padrões de e-mail dos funcionários que tiveram acesso aos cartões de crédito, eles rapidamente identificaram um funcionário que enviou dezenas de imagens como cópia oculta (CCO) para um endereço de e-mail externo, no final de cada dia. Estas imagens eram fotocópias digitalizadas de cartões de crédito e documentos de identidade, que a empresa exigia dos clientes para enviar para verificação. O funcionário estava enviando estas informações para um associado, que, então, comprometeria os cartões.

Quando os investigadores identificaram este associado, o Nuix tornou fácil reutilizar a inteligência que reuniram a partir da do caso original e demonstrar as correlações entre os dois infratores.

É fácil perder conexões, particularmente,
sem uma maneira automatizada para identificar itens de inteligência

TRIAGEM FORENSE BASEADA EM CONTEÚDO – UM MÉTODO MAIS EFICIENTE

É proibitivamente demorado e custoso analisar conjuntos de dados de vários terabytes usando métodos tradicionais. Os investigadores devem procurar outras opções.

Nos últimos anos, muitos pesquisadores publicaram trabalhos recomendando os procedimentos de triagem forense baseada em conceitos, tais como, amostragem estatística[vi], aprendizado de máquinas[vii] e apreensão seletiva de evidências.[viii] Alguns órgãos de aplicação da lei pesam fatores como a ameaça representada pelo infrator, a gravidade do crime e o risco para a vítima, ao escolher quais dispositivos analisar.[ix]

Através de nossa experiência de trabalho com as autoridades policiais e investigadores corporativos ao longo de muitos anos, o Nuix testemunhou uma emergente abordagem baseada em conteúdo para a triagem forense. Isto vem da experiência de outras disciplinas como descoberta legal e governança da informação, onde a análise do conteúdo de grandes volumes de dados é a norma.

Nossa metodologia de triagem forense envolve a coleta de todos os dados disponíveis em um único local de armazenamento, em seguida, usando uma combinação de gerenciamento de dados, análise e técnicas forenses para se concentrar nas fontes mais importantes até que a evidência-chave surja.

Ele alcança os mesmos ou melhores resultados que os métodos forenses tradicionais, mas mais rapidamente e com mais eficiência.

Esta abordagem funciona porque, na grande maioria dos casos, a evidência crítica está escondida à vista de todos, em vez de em artefatos forenses que os investigadores passam tanto tempo analisando.

O processo de triagem forense baseada em conteúdo segue uma série de passos lógicos.

Assimilar todos os dados

A primeira etapa deste processo requer assimilar todas as fontes de dados em um único repositório. Essas fontes de dados podem incluir:

- Unidades de disco rígido de laptops, PCs ou servidores
- Dispositivos de armazenamento de estado sólido, incluindo unidades USB e cartões de memória flash
- Arquivos de banco de dados de e-mail pessoal como o Microsoft Outlook .PST
- Arquivos de banco de dados de e-mail corporativo como o Microsoft Exchange.EDB
- Serviços de e-mail ou armazenamento em nuvem
- Repositórios corporativos como o Microsoft SharePoint, arquivos de e-mail e compartilhamentos de arquivos
- Imagens forenses de discos rígidos
- Imagens forenses de dispositivos móveis.

Realizar uma varredura leve de metadados

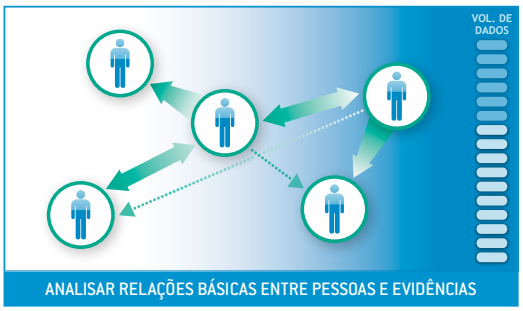
Uma varredura leve de metadados tabula informações, tais como, o proprietário ou remetente, tamanho, formato, nome do arquivo ou linha de assunto e datas relevantes para cada arquivo, mensagem de e-mail e anexos no armazenamento de evidências. Ela cria um valor de *hash* criptográfico para cada item. Uma verificação leve de metadados não extrai o texto completo de cada item, mas é muito mais rápida que a indexação completa.

Usando os metadados e valores de *hash*, é possível ocultar ou remover itens duplicados. Embora duplicatas às vezes possam ser relevantes para uma investigação – por exemplo, para mostrar quem recebeu um documento particular – elas muitas vezes apenas inflam o conjunto de dados. Normalmente, mais de um quarto de todos os dados em um conjunto de evidências é duplicado.

A triagem forense baseada em conteúdo alcança os mesmos ou melhores resultados que os métodos tradicionais, mas de forma mais rápida e eficiente

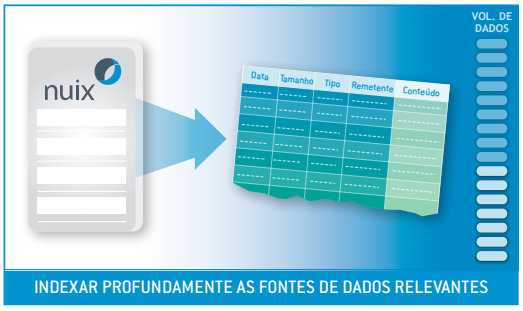
Analisar as relações básicas entre pessoas e evidências

Utilizando técnicas como diagramas de rede e linhas do tempo, os investigadores podem ver conexões e fluxos de informação entre os suspeitos ou custodiantes. Isso pode ajudar a rapidamente restringir datas, fontes de dados e pessoas para examinar em maior profundidade. Alternativamente, podem ser reveladas lacunas de informação que justifiquem uma investigação mais aprofundada.



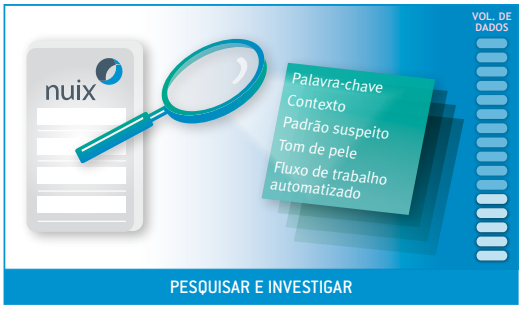
Indexar profundamente as fontes de dados relevantes

Uma vez identificadas as fontes mais prováveis de evidência crítica, os investigadores extraem o texto completo e metadados a partir dessas fontes. Nesta fase, a velocidade e rigor da ferramenta de indexação são fundamentais. O software de indexação deve ser capaz de extrair rapidamente o texto e um espectro completo de metadados a partir de tantos formatos de arquivo e dispositivos quanto possível.



Pesquisar e investigar

Um índice completo de dados e metadados permite que investigadores conduzam seus fluxos de trabalho padrão para busca de evidências em todas as fontes ao mesmo tempo. Ele também possibilita sofisticadas técnicas de pesquisa e análise, algumas das quais tiveram origem em áreas complementares como a descoberta legal e governança da informação (veja “Técnicas avançadas de análise de conteúdo”).

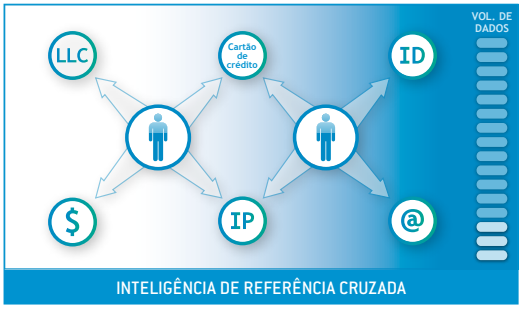


Inteligência de referência cruzada

Ferramentas avançadas de investigação podem extrair automaticamente e destacar itens de inteligência, incluindo:

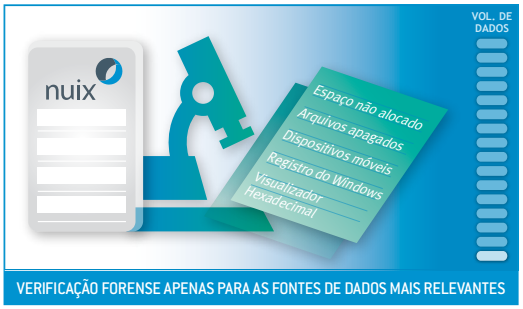
- Nomes
- Endereços de e-mail
- Endereços IP
- Nomes de empresas
- Números de cartões de crédito
- Números de contas bancárias
- Números de seguro social ou identidade
- Somas de dinheiro.

A referência cruzada dessa inteligência em todas as evidências disponíveis pode revelar rapidamente as relações entre pessoas e entidades, entregar pontos para provar e também oferecer inteligência mais ampla. Ela traz à luz as conexões que os investigadores humanos podem perder. Ela também permite aos investigadores construir bibliotecas de inteligência que podem ser usadas em vários casos.



Verificação forense apenas para as fontes de dados mais relevantes

Na maioria dos casos, esse processo já terá localizado evidência crítica. Se não, é quase certo que terá fornecido pistas sobre onde tal evidência está oculta. Os investigadores podem então usar suas habilidades forenses digitais para escavar profundamente nas fontes de dados mais prováveis. Desta forma, evitam passar incontáveis horas de análise forense de material irrelevante.



TÉCNICAS AVANÇADAS DE ANÁLISE DE CONTEÚDO

Combinando análise forense digital tradicional com técnicas de *eDiscovery* e governança da informação, os investigadores podem:

- Rápida e simplesmente identificar padrões suspeitos em e-mails como mensagens enviadas fora do horário comercial; mensagens enviadas de contas corporativas para endereços pessoais, meios de comunicação ou empresas concorrentes; e mensagens que contêm arquivos zip criptografados como anexos.
- Automatizar pesquisas de listas de palavras que contêm termos comumente usados em fraudes e uso de análise de "triângulo da fraude" para identificar ligações entre pressão, racionalização e incentivo para cometer fraudes.
- Representar graficamente dados através de diagramas de rede, linhas do tempo e gráficos de tendências por data para tornar os padrões mais evidentes.
- Usar geoposicionamento de endereço IP ou dados GPS incorporados em fotos para identificar localizações em um mapa.
- Reconstruir conversas de e-mail de várias fontes para que um investigador possa lê-las na ordem que foram enviadas entre os indivíduos.
- Aplicar análise de tom de pele para rapidamente identificar imagens inapropriadas e rastrear sua origem.
- Identificar documentos duplicados e quase duplicados para ver quais suspeitos têm recebido ou enviado e-mails, documentos ou anexos chave; para analisar como os documentos mudaram ao longo do tempo; ou para encontrar documentos que usem linguagem semelhante.
- Utilizar a tecnologia de quase duplicatas para acelerar o processo de identificação de evidências relevantes em clusters não alocados através da ligação dos dados recuperados aos conteúdos semelhantes em arquivos alocados.
- Realizar as mesmas pesquisas e análises em vários dispositivos, sistemas de arquivos, plataformas, em uma única etapa ou criando fluxos de trabalho e investigação automatizados e reprodutíveis.
- Visualizar palavras-chave em "clusters" contextuais em torno de palavras, reduzindo a incidência de falsos positivos nos resultados da pesquisa.
- Combinar a análise de quase duplicados e contexto de palavras para rapidamente identificar e eliminar grandes quantidades de dados irrelevantes. Organizando os dados em grupos de conteúdo semelhante, os investigadores podem agir de forma mais inteligente, deixando-os de lado ou direcionando-os para uma análise mais profunda.

O Nuix oferece uma alternativa completa e superior aos antigos aplicativos de informática forense, com capacidades incomparáveis para lidar com conjuntos de dados grandes e complexos

RECURSOS DO NUIX – DO BIG DATA ATÉ PEQUENOS DETALHES FORENSES

O Nuix oferece uma alternativa completa e superior aos antigos aplicativos de informática forense, com capacidades incomparáveis para lidar com conjuntos de dados grandes e complexos.

Além disso, o Nuix é o único aplicativo atualmente disponível que torna possível o método de triagem forense baseada em conteúdo já discutido neste artigo.

O software de investigação eletrônica Nuix pode:

- Em um único servidor high-end, indexar completamente mais de 2.5 terabytes de dados por dia e realizar uma varredura leve de metadados de mais de 10 terabytes por dia.
- Obter todos os metadados, textos e estruturas binárias de documentos para identificação forense de informações relevantes.
- Manter intacta a cadeia de custódia de cada documento com auditoria total em níveis de itens e de caso, aderindo a rigorosas diretrizes de manuseio de evidências, como as da Associação de Chefes de Polícia do Reino Unido.
- Extrair e tornar pesquisáveis arquivos armazenados em até 100 níveis de profundidade.
- Ler todos os formatos mais comuns de e-mail, estruturas de arquivo, idiomas e conjuntos de caracteres.
- Criar sumários e sintetizar as evidências em relatórios flexíveis e análises visuais dinâmicas, que os investigadores podem usar para comunicar complexos resultados para cima na cadeia de comando. O Nuix proporciona controle total sobre os perfis de metadados para permitir a criação de relatórios relevantes e consistentes em nível de itens através de fontes de dados díspares.
- Agregar evidências e inteligência de vários casos para entender o contexto amplo de um caso em particular e quem são as pessoas envolvidas.
- Gerenciar fluxos de trabalho de forma inteligente, dividindo um caso em subconjuntos e distribuindo tarefas de análise entre uma equipe de investigadores, fornecendo monitoramento, auditoria e visibilidade detalhados do progresso no decorrer do caso.

Para investigações que envolvem sistemas empresariais, o Nuix oferece poderosas capacidades para indexar formatos de dados não estruturados, incluindo arquivos de e-mail, compartilhamentos de arquivos, bancos de dados de e-mail do Lotus Notes, bancos de dados de e-mail do Microsoft Exchange, outros formatos de encapsulamento de e-mail Microsoft e sistemas de colaboração Microsoft SharePoint.

Na maioria dos casos, o Nuix concentra-se nestes dados em nível binário e de sistema de arquivos. Isso ignora as interfaces de programação dos fornecedores de software, que não são concebidas para indexação de alto desempenho e podem não retornar consistentemente todos os dados e metadados. Além disso, esta abordagem evita a divisão dos conjuntos de dados em componentes menores ou a conversão dos mesmos em um formato mais legível. Isso garante que o Nuix preserve totalmente os metadados e cadeia de custódia das informações, o que pode ser essencial para fins legais.

Embora o Nuix seja mais conhecido por suas habilidades de processar grandes volumes de dados, ele também pode executar uma detalhada análise forense, incluindo:

- Indexação e visualização do registro do Microsoft Windows
- Recuperação total e parcial de arquivos apagados e espaço livre
- Prospecção em imagens forenses e verificação de artefatos forenses com um visualizador hexadecimal
- Carregamento de imagens móveis Cellebrite e Micro Systemation XRY para analisar padrões de comunicação
- Análise profunda dos sistemas de arquivos Mac OS HFS+ e HFSX e Linux.

Embora o Nuix seja mais conhecido por suas habilidades de processar grandes volumes de dados, ele também pode executar uma detalhada análise forense

CASOS DE USO COMUM

Os clientes do Nuix, mais comumente, usam o software para investigar:

- Fraude
- Dados privados (PII) e de cartões de crédito (PCI) armazenados sem segurança adequada
- Imagens inapropriadas
- Tráfico de drogas ou de pessoas
- Compromisso ou violação de rede
- Luta contra o terrorismo e inteligência
- Roubo de propriedade intelectual, após um alerta de prevenção de perda de dados
- Práticas corruptas no exterior, violação da lei ou pagamentos indevidos
- Falsificação de documentos
- Problemas com recursos humanos ou de emprego.

ESTUDO DE CASO: INVESTIGAÇÃO DE INFORMAÇÕES PRIVILEGIADAS

Um regulador corporativo apresentou um caso contra um ex-diretor de um banco comercial de abuso de informações privilegiadas. O regulador acreditava que ele havia escrito as negociações em seu BlackBerry, mas não podia provar isso usando ferramentas forenses padrão. Usando o software Nuix, o regulador analisou os metadados completos do Servidor Corporativo BlackBerry do banco, que incluía códigos que identificam as mensagens de e-mail que o diretor enviou para seu corretor solicitando as negociações ilegais. Isso permitiu que o regulador provasse rapidamente seu caso e apresentasse acusações bem sucedidas.

ESTUDO DE CASO: INVESTIGAÇÃO DE FRAUDE MULTIJURISDICIONAL

Quando um órgão do governo começou a investigar uma empresa que vendia de forma fraudulenta aeronaves que não existiam, reconheceu que seria necessária uma equipe de até 20 investigadores para examinar os dados disponíveis usando métodos tradicionais. O órgão havia apreendido cerca de 40 dispositivos, incluindo computadores de mesa, laptops e smartphones. Investigar sequencialmente cada dispositivo teria tornado impossível localizar ligações entre diferentes custodiantes e compras.

Em vez disso, o órgão assimilou todos os dados disponíveis em um único local de armazenamento, e depois usou o software Nuix para indexá-los e criar referências cruzadas. As ferramentas do Nuix ajudaram um único investigador a identificar rapidamente a evidência mais crítica, permitindo ao órgão fazer acusações.

Além disso, usando as funções de quase duplicata do Nuix para encontrar documentos semelhantes, o investigador trouxe à luz uma série de empresas associadas, desconhecidas pelo órgão, que estavam realizando operações fraudulentas de peças de aviões, barcos e outros produtos de alto valor. O Nuix tornou fácil para o órgão transferir a inteligência do processo inicial para investigações das empresas relacionadas, o que levou a novas acusações.

Os investigadores podem acessar dados armazenados em complexos repositórios corporativos e serviços baseados em nuvem, bem como pesquisar até os menores detalhes forenses

BENEFÍCIOS PARA OS INVESTIGADORES

Utilizar uma metodologia de triagem forense baseada em conteúdo para investigar todas as fontes de dados ao mesmo tempo é muito mais eficiente e sofisticado que os métodos tradicionais de extração de pontos para provar ou de inteligência em grandes volumes de dados.

Os benefícios desta abordagem incluem:

- Os investigadores podem acessar dados armazenados em complexos repositórios corporativos e serviços baseados em nuvem, bem como pesquisar até os menores detalhes forenses quando necessário
- As ferramentas do Nuix extraem automaticamente e criam referências cruzadas de inteligência, tais como nomes, endereços de e-mail e números de cartões de crédito; isto evidencia ligações que podem não ser imediatamente óbvias
- Utilizar a análise de conteúdo para a triagem de fontes de evidências ajuda os investigadores a rapidamente chegar ao cerne da questão, evitando uma demorada e desnecessária análise forense
- Examinar os resultados de pesquisa no contexto, usando técnicas como grupos de palavras e documentos quase duplicados, proporciona resultados mais relevantes e menos falsos positivos que pesquisas básicas de palavras-chave
- Analisar visualmente os dados torna muito mais fácil detectar tendências e isolar os outliers através grandes volumes de evidências de várias fontes e formatos
- Os investigadores podem reduzir tarefas repetitivas ao automatizar fluxos de trabalho e transferência de informações úteis, grupos de palavras e padrões de pesquisa através de múltiplos casos
- Ele economiza inteligência e horas de pessoal, proporcionando uma poderosa forma de resolver pendências de investigação e casos não resolvidos.

REFERÊNCIAS

i Mark A. Beyer, Douglas Laney, “The Importance of “Big Data”: A Definition”, Gartner Inc., Junho de 2012

ii Ver, por exemplo, Australia New Zealand Policing Advisory Agency, submission to Australian Government Joint Select Committee on Cyber-Safety, 20 April 2011, http://www.aph.gov.au/Parliamentary_Business/Committees/House_of_Representatives_Committees?url=/jscc/subs/sub_151.pdf

iii IDC, “The Digital Universe in 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East”, Dezembro de 2012, http://www.emc.com/leadership/digital- universe/index.htm

iv Association of Chief Police Officers, “Good Practice Guide for Computer-Based Electronic Evidence”, http://www.7safe.com/electronic_evidence/ACPO_guidelines_computer_evidence.pdf

v Ver, por exemplo, Royal Canadian Mounted Police, “Audit of Technological Crime Program”, Fevereiro de 2012, http://www.rcmp-grc.gc.ca/aud-ver/reports-rapports/tech- crime-crimin-tech-eng.htm and National Policing Improvement Agency, “Success of NPIA eForensics pilot set to help forces bring more offenders to justice quicker”, Agosto de 2012, http://npia.pressofficeadmin.com/component/content/article/38-press-releases/513

vi John Clayton, “Investigation into a digital forensics triage tool using sampling, hashes and bloom filters – SHAFT”, Edinburgh Napier University, Setembro de 2012, http://researchrepository.napier.ac.uk/id/eprint/5668

vii Fabio Marturanaa, Simone Tacconib, “A Machine Learning-based Triage methodology for automated categorization of digital media”, Digital Investigation, Fevereiro de 2013, http://www.sciencedirect.com/science/article/pii/S1742287613000029

viii Ilyoung Honga, Hyeon Yua, Sangjin Leea, Kyungho Lee, “A new triage model conforming to the needs of selective search and seizure of electronic evidence”, Digital Investigation, Fevereiro de 2013, http://www.sciencedirect.com/science/article/pii/S1742287613000042

ix National Policing Improvement Agency, op. cit.

PARA SABER MAIS SOBRE O SOFTWARE DE INVESTIGAÇÃO ELETRÔNICA NUIX, VISITE

nuix.com/investigation

SOBRE O NUIX

O Nuix capacita as pessoas a tomar decisões baseadas em fatos a partir de dados não estruturados. O mecanismo patenteado do Nuix faz pequenos trabalhos de grandes e complexos conjuntos de dados gerados por pessoas. Organizações em todo o mundo recorrem ao software Nuix quando precisam de respostas rápidas e precisas para investigação digital, segurança cibernética, eDiscovery, governança de informações, migração de e-mail, privacidade e muito mais.

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

América do Norte
USA: +1 877 470 6849
» Web: nuix.com

EMEA
Reino Unido: +44 207 877 0300
» Twitter: @nuix