El uso de los métodos deductivo e inductivo para aumentar la eficiencia del procesamiento de adquisición de evidencias digitales*

Use of Deductive and Inductive Methods to Increase the Efficiency in the Acquisition and Processing of Digital Evidence

O uso dos métodos dedutivo e indutivo para aumentar a eficiência do processamento de aquisição de evidencias digitais

Bayron José Prieto Castellanos [a]
Pontificia Universidad Javeriana, Colombia
bprieto@ediligence.co
ORCID: http://orcid.org/0000-0002-9780-2100

DOI: https://doi.org/10.11144/Javeriana.cc18-46.umdi

Fecha de recepción: 21 Abril 2017
Fecha de aprobación: 15 Junio 2017
Fecha de publicación: 15 Diciembre 2017

Resumen:

El presente artículo tiene como propósito mostrar la aplicación de los métodos de investigación deductivo e inductivo en el ámbito de la optimización del procesamiento de evidencias digitales. Dicha pesquisa tiene como base un esquema de adquisición estructurada de la información, y se desarrolla con la finalidad de aumentar la eficiencia de los procesos descritos, en el contexto de la práctica informática forense. Según lo anterior, el autor define dichos métodos filosóficos, en tanto que son útiles en otras áreas del conocimiento, y presenta la aplicación de ambos mecanismos de acuerdo con el marco general de obtención y análisis de evidencias digitales E-Discovery, así como con el proceso de adquisición, procesamiento, puesta a disposición e investigación (APPI), que es definido y desarrollado también en el presente estudio. Asimismo, el autor presenta todos estos conceptos bajo el lente de los lineamientos que constan en las leyes 527 de 1999 —Ley de Comercio Electrónico— y 906 de 2004, al igual que bajo el rasero del Decreto Reglamentario 1747 de 2000 y la norma ISO 27001.
Código JEL: O32
Palabras clave: E-Discovery, proceso APPI, métodos de investigación inductivo y deductivo, mensajes de datos, evidencias digitales, elemento material probatorio.

Abstract:

This article aims to show how to apply the inductive and deductive research methods in the field of optimization of digital evidence processing. The inquiry is based on a scheme of structured information acquisition and is developed in order to increase the efficiency of the described processes in the context of the computer forensics practice. According to this, the author defines the abovementioned philosophical methods inasmuch they are useful in other fields of knolwedge, and then shows how to apply both mechanisms under the general frame E-Discovery for obtaining and analyzing digital evidence and in keeping with the APMaR process (Acquisition, Processing, Making available & Research), which will be also defined and developed in this study. Additionally, the author presents all those concepts through the lens of the regulations provided in the acts 527 from 1999 —the e-Commerce Act— and 906 from 2004 and the benchmark provided by the Regulatory Decree 1747 from 2000 and the standard ISO 27001.
Keywords: E-Discovery, APMaR process, inductive and deductive research methods, data messages, digital evidence, evidentiary materials.

Resumo:

O presente artigo tem como propósito mostrar a aplicação dos métodos de pesquisa dedutiva e indutiva no âmbito da otimização do processamento de evidencias digitais. Tal pesquisa tem como base um esquema de aquisição estruturada de informação e é desenvolvida com a finalidade de aumentar a eficiência dos processos descritos, no contexto da prática informática forense. De acordo com o exposto, o autor define tais métodos filosóficos, em tanto que são úteis em outras áreas do conhecimento, e apresenta

Notas de autor

[a] Autor de correspondencia. Correo electrónico: bprieto@ediligence.co

a aplicação de ambos os mecanismos de acordo com o marco geral de obtenção e análise de evidencia digital E-Discovery, bem como com o processo de aquisição, processamento, posta a disposição e pesquisa (APPI, pelas suas iniciais em espanhol), que é definido e desenvolvido também no presente estudo. Mesmo assim, o autor apresenta todos esses conceitos sob o lente das diretrizes que constam nas leis 527 de 1999 —Lei de Comércio Eletrônico— e 906 de 2004, igual como sob o raseiro do Decreto Regulamentário 1747 de 2000 e a norma ISO 27001.
Palavras-chave: E-Discovery, processo APPI, métodos de pesquisa indutiva e dedutiva, mensagens de dados, evidencias digitais, elemento material probatório.

Introducción

En el mundo de las evidencias que hoy en día dan soporte a múltiples procesos legales, existe una división entre lo análogo o físico, y lo digital. Por una parte, en la dimensión análoga las evidencias son elementos físicos que dan cuenta de un comportamiento, tales como los objetos que fueron utilizados para perpetrar un crimen —como los casos de armas o elementos punzocortantes—, o documentos que respaldan transacciones —como cartas, libros contables o facturas—. Por otro lado, en el contexto digital las evidencias son elementos que están almacenados en un formato digital y que cumplen el mismo propósito que sus análogas de formato físico.

Dada la naturaleza de las evidencias digitales, que acarrea su fácil modificación y volatilidad, se deben seguir procedimientos rigurosos para el aseguramiento de su veracidad, con el fin de que sean válidas en un proceso legal. Estos procedimientos deben incluir protocolos que las aseguren y mantengan su integridad; a su vez, deben ser procedimientos eficientes que permitan el uso óptimo del tiempo y de los recursos necesarios para la investigación que se organiza en torno a ellas. Con base en lo anterior, este artículo plantea un procedimiento general de *adquisición, procesamiento, puesta a disposición* e *investigación* de evidencias digitales (APPI); metodología que surge a partir del proceso E-Discovery. En tanto, se propone también la aplicación de los métodos de investigación deductivo e inductivo con el fin de optimizar el proceso propuesto.

De acuerdo con la macroestructura temática formulada, el presente artículo presenta en primera medida el marco general de obtención y análisis de evidencias digitales E-Discovery. En segundo lugar, se presenta el modelo APPI, el cual que se origina a partir del proceso referido anteriormente. A continuación, el estudio presenta los métodos de investigación deductivo e inductivo, junto con su aplicación en el modelo. Posteriormente, se muestran los resultados obtenidos tras la aplicación de estos métodos en las etapas de adquisición y procesamiento de evidencias digitales.

Marco general de obtención y análisis de evidencias digitales: E-Discovery

E-Discovery —que en su versión extendida corresponde a *Electronic Discovery*, lo cual en este contexto se traduce como el procedimiento de *Investigación Electrónica*— es el proceso que enmarca aquellas investigaciones digitales que involucran mensajes de datos (Ley 527, 1999). Según lo anterior, su fin es conservarlos para que sean utilizados como elementos materiales probatorios válidos ante autoridades judiciales y administrativas, y en general idóneo para su uso en ámbitos legales en los cuales sea necesaria la inclusión de pruebas de este tipo (Prieto Castellanos, 2014b).

En concordancia con lo hasta tanto descrito, se formula un contexto en el cual uno de los objetivos de la auditoría forense es la detección y prevención de delitos económicos y financieros. Asimismo, surge un entorno en el que los hallazgos del trabajo auditor deben ser puestos a disposición de las autoridades legales, para el análisis de dicha información y la emisión de una decisión. En tal escenario, el E-Discovery permite la materialización y la preservación de esos hallazgos o pruebas, de tal forma que sean tenidos en cuenta como elementos materiales probatorios de carácter digital.

Cabe añadir que la auditoría forense se enfoca en delitos económicos y financieros. En contraste, la informática forense es una disciplina más general que en todo caso se encarga de analizar sistemas de información con el propósito de encontrar mensajes de datos, independientemente del campo de acción en el cual fue cometida la conducta que contraviene la ley. De acuerdo con lo anterior, se ocupa tanto de aspectos como los económicos, financieros e informáticos, como de los penales y civiles. En suma, lo anterior se resume en la siguiente tabla:

TABLA 1
Comparación entre auditoría forense e informática forense

	Marco de actuación	Campo de acción
Auditoría forense	Detección y prevención	Temas económicos y financieros
Informática forense	Detección de conductas	Cualquiera que se pueda evidenciar por medio de mensajes de datos

Fuente: elaboración propia

En tanto, de acuerdo con el diagrama denominado Electronic Discovery Reference Model (EDRM, 2005), el proceso para la obtención de evidencias digitales se puede visualizar en el siguiente esquema:

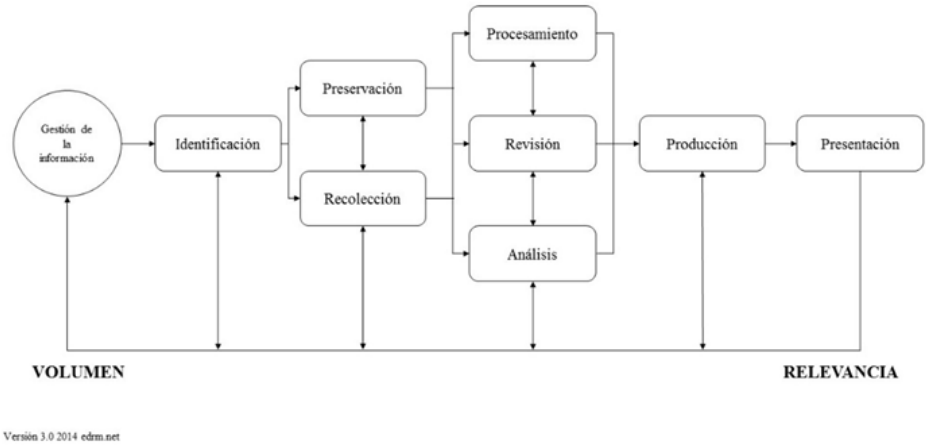

FIGURA 1
Electronic Discovery Reference Model: modelo de referencia del E-Discovery
Fuente: EDRM (2005)

En ese mismo plano, los siete acápites subsiguientes constituyen un desglosamiento conceptual del proceso de E-Discovery con sus diferentes etapas.

Identificación

En esta etapa del proceso, se identifican las posibles fuentes de información almacenada electrónicamente, ESI —siglas que equivalen en inglés a *electronically stored information*— o *mensajes de datos*, en términos de la ley colombiana (Ley 527, 1999) relevantes para un proceso legal. Dicho proceso se desarrolla en la tanto que se ha determinado la amplitud, el alcance y la profundidad de dichos mensajes de datos (EDRM, 2005). Cabe añadir que en ese contexto la *profundidad* puede ser comprendida como la relevancia que tienen los mensajes de datos para que puedan ser contemplados como evidencias digitales; concepto que es definido más ampliamente en el acápite denominado "Adquisición".

Preservación

En esta etapa del proceso, se protegen los mensajes de datos o evidencias digitales relevantes para temas legales, garantizando su integridad, confidencialidad y disponibilidad. Cabe añadir que la norma ISO 27001 define estos tres conceptos así: (a) la *integridad* hace referencia a la preservación de la información, así como a la vigilancia de su exactitud y completitud, al igual que de sus métodos de tratamiento (Prieto Castellanos, 2014a); (b) la *confidencialidad* se refiere al acceso a la información, que solamente puede ser otorgado al personal autorizado (Prieto Castellanos, 2014a); y, finalmente, (c) la *disponibilidad* de la información corresponde al acceso y la utilización de esta por parte del personal autorizado en el momento requerido (Prieto Castellanos, 2014a).

Recolección

Esta etapa es aquella en la que se reúnen todos los mensajes de datos relevantes para el proceso legal, que posteriormente serán procesados, revisados y analizados en las fases restantes del proceso E-Discovery —que incluyen el procesamiento y la revisión, así como los demás momentos de la sucesión—. Este ciclo se ejecuta simultáneamente con el de la preservación.

Procesamiento

En esta fase, se reduce el volumen de la información adquirida, que a su vez es identificada detalladamente para su posterior revisión y análisis. Cabe aclarar que los mensajes de datos pueden llegar a esta etapa en estados que suscitan algún tipo de restauración, como por ejemplo cuando se presentan en cintas o *backups*. Por otro lado, hay archivos digitales que pueden estar almacenados en formatos compuestos, [1] como por ejemplo los archivos con extensión *.pst* (formato de correo electrónico de Outlook), *.zip* o *.rar*. En consecuencia, esta identificación consiste en la depuración e indexación de los mensajes de datos y la preservación de sus metadatos asociados. Adicionalmente, en esta etapa se pueden recuperar mensajes de datos que fueron borrados de la evidencia original.

Revisión

Esta es la etapa en la cual se revisan los mensajes de datos procesados, su contenido y sus metadatos, con el fin de establecer su relevancia dentro del proceso. La revisión descrita permite organizar los mensajes de datos en grupos de acuerdo con sus características, con el fin de mejorar la eficiencia y la efectividad de la etapa de análisis.

Análisis

Luego de la revisión y organización de los mensajes de datos, se procede a analizar y evaluar su contenido y contexto de acuerdo con patrones claves tales como los temas y las personas que son recurrentes en ellos. En concordancia, el objetivo del análisis es proporcionar al equipo legal datos verificados que sirvan de soporte para las decisiones que se vayan a tomar en el caso.

Producción

En la fase de producción, los resultados de las tres etapas anteriores son organizados para su presentación. Esta organización incluye exponer esos resultados mediante informes y presentaciones que contengan información clara y entendible para cualquier persona.

Presentación

En esta etapa los mensajes de datos son presentados antes de la audiencia, litigio o juicio. Estos asumen su forma nativa con los fines de proveer más información, validar hechos o posiciones existentes, y verificar que todos los mensajes de datos relevantes para el caso fueron procesados y analizados.

Cabe señalar que el E-Discovery es un proceso cíclico y en cualquiera de sus etapas se puede retomar la fase inmediatamente anterior tantas veces como sea necesario. Por otro lado, en este proceso se consideran dos factores importantes: (a) el volumen y (b) la relevancia.

Por una parte, el factor del *volumen* está directamente relacionado con la cantidad de mensajes de datos que se identifican, preservan y recolectan (tres primeras etapas del E-Discovery); y, por otra, la *relevancia* se define de acuerdo con los resultados obtenidos a lo largo del procesamiento, la revisión y el análisis (tres fases subsiguientes del E-Discovery). Finalmente, se proporcionan los lineamientos para la producción y la presentación de los mensajes de datos o elementos materiales probatorios (que corresponden a las dos etapas restantes del E-Discovery).

Modelo APPI para el tratamiento de evidencias digitales

De acuerdo con los parámetros del E-Discovery que devienen de su estudio, y según su aplicación en la práctica profesional, este proceso se puede adaptar, y su estructura se puede modificar, sin que la herramienta pierda su esencia ni su propósito. En ese sentido, es importante anotar que los cambios descritos se formulan de acuerdo con el objeto y las reglas del negocio al cual se aplican.

Por esta razón, la puesta en práctica del proceso APPI es pertinente. Cabe recordar que la abreviatura se desglosa en cuatro acciones esenciales: *adquisición, procesamiento, puesta a disposición* e *investigación de evidencias digitales*. En consecuencia, es posible afirmar que este proceso es una adaptación más simplificada del E-Discovery que reúne todas sus características. A continuación, se dedican cuatro acápites a definir cada una de sus etapas:

Adquisición

Esta fase remite a la recolección de mensajes de datos como elementos de material probatorio idóneo para su postulación ante autoridades legales, o para la ejecución de controles propios por parte de la organización. Como parte de la explicación de esta primera etapa, se hace necesario aclarar las diferencias existentes entre mensajes de datos, evidencia digital y elemento material probatorio.

1. Un *mensaje de datos* es cualquier información que se genera, recibe, almacena o envía por medios electrónicos, según lo menciona el literal a) del artículo 2 de la Ley 527 de 1999 (Ley 527, 1999).
2. La *evidencia digital* es aquel archivo o conjunto de archivos electrónicos contenidos en un dispositivo físico que se pueden encontrar en el campo sujeto a análisis, o que se localizan remotamente, como es el caso de la información almacenada en internet. Esta evidencia se

compone de un cúmulo de mensajes de datos que pueden estar en cualquier formato (Ramos Álvarez y Ribagorda Garnacho, 2004); tales son los casos de archivos en formatos de texto —verbigracia, con extensiones *.txt*, *.doc* o *.docx*—, imágenes —por ejemplo, con extensiones *.gif*, *.jpg* o *.png*—, o bases de datos —con extensiones como *.mdf*, *.db*. o *accdb*.—, entre otros. Esta evidencia es recuperada con muestras de la totalidad o parte de los mensajes de datos que se encuentran en el dispositivo o elemento digital examinado. En esencia, una evidencia digital relaciona una conducta con su autor; según eso, existen dos tipos de evidencias digitales:

a) El primer tipo de evidencia digital es generado por mecanismos sistemáticamente programados. Tal es el caso del registro de entradas a una base de datos, o un archivo de texto en el que se describen los cambios sufridos por la información que se encuentra contenida en un sistema.
b) La segunda clase de evidencia digital es la generada por una persona y almacenada en un dispositivo. En consecuencia, ejemplo de ello es un archivo de Word que describe el procedimiento implementado para robar números de tarjetas de créditos de un banco específico.

3. Un *elemento material probatorio* es en términos generales cualquier objeto que demuestre una conducta que atenta contra la ley. Según el literal g) del artículo 275 de la Ley 906 de 2004, un mensaje de datos puede ser considerado un elemento material probatorio una vez que haya sido aportado a un proceso legal; asimismo debe estar protegido, de modo que garantice su integridad, confidencialidad y disponibilidad. En suma, lo anterior significa que el mensaje de datos recolectado en el campo de investigación es el mismo mensaje de datos que se presenta ante una autoridad legal. Adicionalmente, debe haber un registro en el cual se evidencia quién ha sido responsable de custodiar y transportar el mensaje de datos o el contenedor en el que este se encuentra y, asimismo, quién o quiénes han sido los investigadores y han tenido contacto con dicho elemento.

Como parte del proceso expuesto, es esencial tener en cuenta las siguientes consideraciones:

- Un mensaje de datos está contenido en una evidencia digital que ha sido adquirida mediante una o varias muestras.
- Una evidencia digital puede contener uno o más mensajes de datos.
- Para que un mensaje de datos sea considerado un elemento de materia probatoria válido ante autoridades legales, debe estar contenido en una evidencia digital que garantice su integridad, confidencialidad y disponibilidad.

En concordancia, la figura 2 muestra las relaciones que existen entre la evidencia digital, los mensajes de datos y los elementos materiales probatorios:

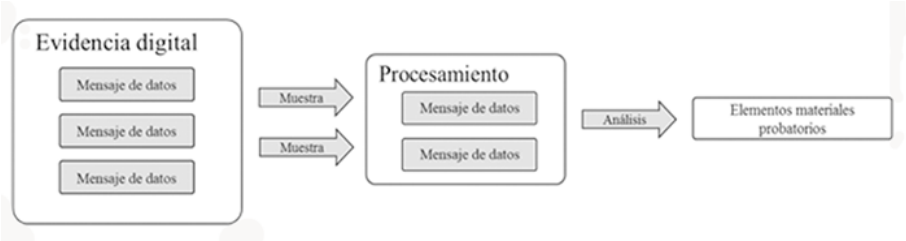

FIGURA 2
Relación existente entre evidencia digital, mensajes de datos y elementos materiales probatorios
Fuente: elaboración propia

Por su parte, el Decreto 1747 de 2000 define al iniciador como la persona que, actuando por su cuenta, o en cuyo nombre se haya actuado, envíe o genere un mensaje de datos (Decreto 1747, 2000).

De esta manera, en la etapa de adquisición se definen dos tipos de iniciadores:

1. El *iniciador del mensaje de datos*. Es aquel que crea el mensaje de datos: el autor. Buen ejemplo de esta categoría es el usuario que crea un nuevo documento de texto, realiza las modificaciones que considera necesarias y guarda el archivo.
2. El *iniciador de la evidencia digital*. Es el que agrupa los mensajes de datos y crea la evidencia digital. Ejemplo de ello es el analista forense que está en el campo de trabajo, tomando muestras de mensajes de datos que considera pertinentes para el proceso legal.

Procesamiento

Esta fase hace referencia al procesamiento de las evidencias digitales adquiridas en la etapa anterior. En consecuencia, el procesamiento incluye el traspaso de la información recolectada en los dispositivos de destino que harán parte del expediente o soporte documental que se lleve del caso, al igual que la extracción y/o recuperación de datos contenidos adentro de las evidencias digitales, entre otros. El procesamiento está clasificado en tres tipos:

1. El *procesamiento corto*. Esta forma de procesamiento solamente incluye la extracción de los mensajes de datos contenidos en las muestras tomadas, con la utilización de herramientas de software que se limitan a identificar los archivos digitales contenidos en ellas. En consecuencia, supone el desarrollo del proceso sin la puesta en marcha de acciones de recuperación propias de procesamientos más avanzados.
2. El *procesamiento completo*. A diferencia del anterior, este además incluye la extracción de mensajes de datos contenidos en las muestras tomadas. Asimismo, también involucra la realización de otras acciones como la recuperación de datos borrados o perdidos, y la organización o indexación del volumen de mensajes de datos adquiridos.
3. El *procesamiento alterno*. Este tipo de procesamiento tiene las mismas características del procesamiento mencionado en el numeral anterior, pero supone la interacción con otras herramientas de software. Tal procedimiento se desarrolla con el fin de agilizar y personalizar los procesamientos de muestras tomadas, de acuerdo con las necesidades del caso.

Puesta a disposición

Esta etapa alude a la puesta a disposición de los diferentes casos procesados para que los expertos en la materia que está sujeta a investigación —quienes cuentan con formación profesional diferente a la de un analista forense— puedan analizar, revisar o efectuar cualquier otra acción que consideren pertinente adentro de esta fase de investigación. Según lo expuesto, en esta etapa es esencial que las herramientas usadas para la puesta a disposición sean intuitivas en su uso para personal no técnico, con el fin de identificar el mayor número de elementos materiales probatorios. Cabe destacar que esta etapa funciona en las investigaciones de tipo interdisciplinar.

Investigación

Esta etapa alude al análisis y la revisión de mensajes de datos de acuerdo con el propósito del proceso legal. Dependiendo del análisis e inspección de los mensajes de datos, se puede definir si estos pueden ser considerados elementos materiales probatorios o no.

La figura 3 expone el proceso comprendido por las tres etapas que han sido expuestas previamente:

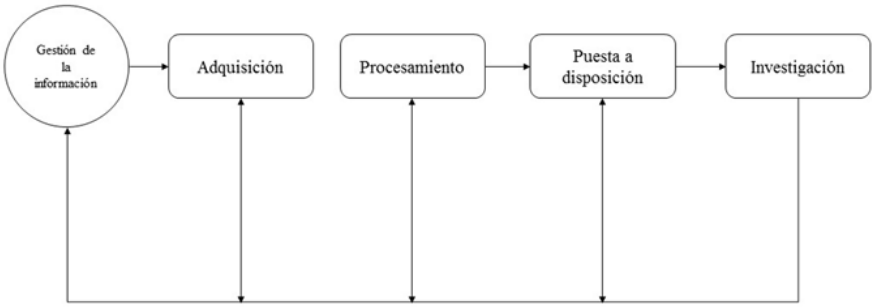

FIGURA 3
Esquema de representación conceptual APPI
Fuente: elaboración propia

En tanto, a continuación se presenta la equivalencia entre el E-Discovery y el APPI:

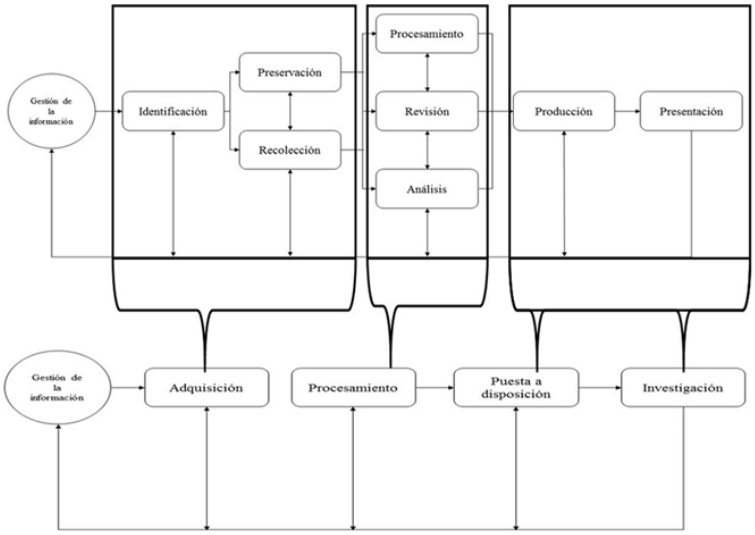

FIGURA 4
E-Discovery vs. APPI
Fuente: EDRM (2005) con modificaciones propias

Cabe señalar que, al igual que el E-Discovery, el APPI es un proceso cíclico, y bajo ese modelo también se pueden retomar etapas anteriores, de ser pertinente. Asimismo, es fundamental hacer notar que de todas las etapas del proceso APPI, la adquisición y el procesamiento son las que más influyen en la eficiencia de todo el proceso. Lo anterior obedece a que según los resultados de estas dos se puede llevar a cabo con éxito las etapas de puesta a disposición e investigación, o estas serán equívocas, anulando todo el proceso. Por esta razón, es necesario utilizar mecanismos idóneos para optimizar el procesamiento de evidencias digitales. Con el foco que constituye dicha optimización, los factores del volumen y la relevancia cobran una importancia notoria, ya que en la medida en que haya un mayor volumen de información, hay mayores demoras en el

procesamiento; asimismo, en la medida en que los mensajes de datos procesados y puestos a disposición sean menos relevantes, menores serán los avances en la investigación, y la finalización del proceso legal será más difícil. En tal escenario, el volumen y la relevancia ideales de todo el proceso APPI tienen el siguiente comportamiento:

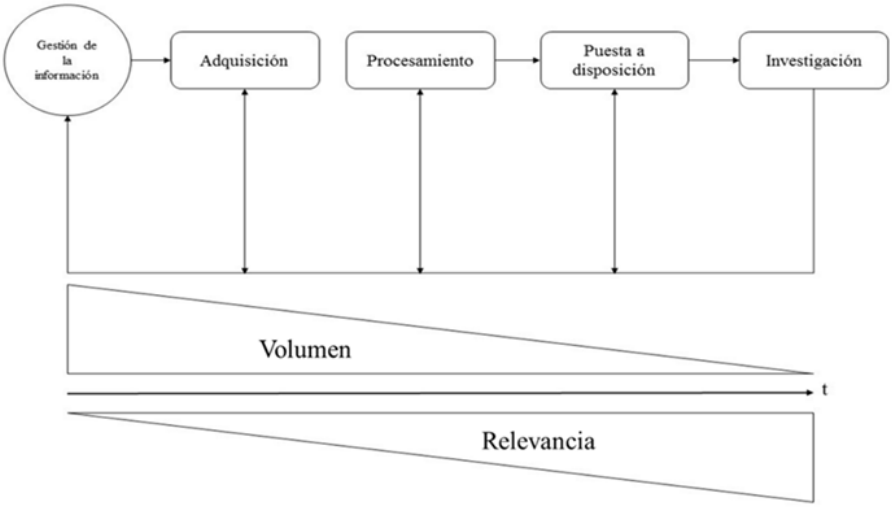

FIGURA 5
Comportamiento de los factores del volumen y la relevancia en el transcurso del proceso APPI
Fuente: EDRM (2005) con modificaciones propias

Ante dicho panorama practico y conceptual, el presente estudio argumenta que un mecanismo idóneo para la optimización del procesamiento de evidencias digitales corresponde a la aplicación de los métodos de investigación deductivo e inductivo. Con base en ello, se pueden implementar dichos modelos filosóficos, y así disminuir el volumen mientras se aumenta la relevancia de la segunda etapa del modelo APPI. Cabe señalar que lo anterior supone además la adquisición de muchas muestras en la primera etapa; es decir, involucra el aumento del volumen y la disminución de la relevancia.

Los métodos de investigación inductivo y deductivo: su aplicación en la informática forense

El término *método* —concepto de origen griego que significa el *camino hacia*— es descrito como un conjunto de actividades que se formulan de acuerdo con una serie de pasos específicos, postulados con el fin de llegar a un resultado determinado (Hurtado León y Toro Garrido, 2007). En ese plano, de acuerdo con Richard McKeon (2010) existen tres métodos filosóficos:

1. El *dialítico*. Este método es definido y desarrollado principalmente por filósofos como Platón y Hegel. Consiste en eliminar las contradicciones del proceso. Asimismo, incluye mecanismos para transformarlas en parte de un conjunto en el que todo está relacionado con todo.
2. El *logístico*. Como método ha sido definido y desarrollado principalmente por filósofos como Demócrito y Descartes. Se basa en la necesidad de aceptar que existen primicias con respecto al objeto de investigación, y a partir de ellas se llega a deducir gran parte del proceso investigado.
3. El *de indagación*. Ha sido definido y desarrollado principalmente por filósofos como Aristóteles y Francis Bacon. Este método consiste en formular convergencias de métodos de varios tipos, con el fin de desarrollar conocimiento o darle solución a un problema (Martínez Chánez, 1998).

Asimismo, en la actualidad estos métodos se dividen en dos grupos de acuerdo a su finalidad o propósito: (a) los que sirven para enseñar, y (b) los que son utilizados para investigar (Martínez Chánez, 1998). Por una parte, los métodos idóneos para enseñar son aquellos que aplica un maestro o profesor para transmitirle conocimiento a una persona, y permitir que aprenda. Por otro lado, los métodos utilizados para investigar son aquellos que se usan para desarrollar y adquirir conocimiento. Estos últimos generalmente son aplicados en las investigaciones que se desarrollan en las diferentes áreas del conocimiento que existen en la actualidad, tales como las matemáticas, la biología, la sociología, la economía, la ecología, la informática y las ciencias forenses.

Según la ciencia o disciplina en la cual se esté investigando, los objetivos propuestos, el tipo de investigación, y los resultados que se esperan de ella, se pueden aplicar uno o varios métodos de investigación. De acuerdo con la propuesta de Bernal Torres (2006), entre los principales métodos de investigación se encuentran:

1. El *galileano*, el cual se cimenta en la experimentación.
2. El *cartesiano*, que se basa en el análisis.

Asimismo, como métodos de investigación complementarios, Bernal Torres (2006) se fija en los modelos formulados a continuación:

1. Inductivo
2. Deductivo
3. Inductivo-deductivo
4. Hipotético-deductivo
5. Analítico
6. Sintético

El método inductivo

El modelo inductivo —que etimológicamente se deriva de la *conducción a* o *hacía*— es un método basado en el razonamiento, el cual "permite pasar de hechos particulares a los principios generales" (Hurtado León y Toro Garrido, 2007, p. 84). Fundamentalmente consiste en estudiar u observar hechos o experiencias particulares con el fin de llegar a conclusiones que puedan inducir, o permitir derivar de ello los fundamentos de una teoría (Bernal Torres, 2006).

Sin embargo, uno de los problemas de este tipo de método es que solo puede ser aplicado a objetos de cierta clase, cuyas partes deben ser identificables durante el estudio. Cabe destacar que la anterior condición se formula con el fin de encontrar todos los elementos propios del análisis. En ese sentido, la inducción científica no podrá ser completa en el contexto del conocimiento buscado, ya que es casi imposible observar todos los elementos que influyen en la investigación (Hurtado León y Toro Garrido, 2007). En suma, para probar que una teoría es cierta o correcta se usan las estadísticas que permitirían confirmar o desvirtuar la postura de que determinada teoría en estudio es en efecto correcta.

En tal contexto, el siguiente es un ejemplo del método inductivo en el que el *enunciado universal* es válido únicamente si —y solo si— la suma de los *enunciados singulares* constituye la totalidad del universo:

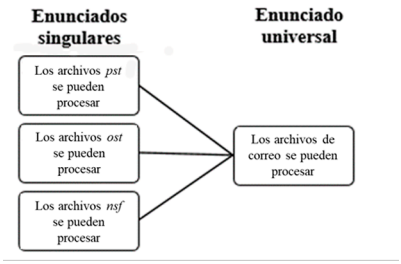

FIGURA 6
Ejemplo general del método inductivo
Fuente: elaboración propia

El método deductivo

Por su parte, el método deductivo —que en términos de sus raíces lingüísticas significa *conducir* o *extraer*— está basado en el razonamiento, al igual que el inductivo. Sin embargo, su aplicación es totalmente diferente, ya que en este caso la deducción intrínseca del ser humano permite pasar de principios generales a hechos particulares. Lo anterior se traduce esencialmente en el análisis de los principios generales de un tema específico: una vez comprobado y verificado que determinado principio es válido, se procede a aplicarlo a contextos particulares (Bernal Torres, 2006).

A pesar de ser reconocido como el primer método científico, pues fue utilizado por los antiguos griegos, y tuvo auge notable durante la Edad Media y Edad Moderna —durante la cual se comenzó a poner a prueba la veracidad de las Sagradas Escrituras—, uno de sus principales inconvenientes es que "otorga validez formal al contenido del pensamiento racional, pero no veracidad a su contenido" (Hurtado León y Toro Garrido, 2007, p. 83). Lo anterior significa que, aunque un hecho particular sea racional en el sentido estricto de la palabra, eso no supone que sea verídico o correcto en la realidad.

Para solucionar este dilema, los científicos comprueban sus hipótesis o proposiciones deductivas, por medio de la realización de experimentos en los cuales tanto el principio general como los hechos particulares y el resultado de los experimentos están totalmente alineados; en suma, todos son consistentes con la realidad (Hurtado León y Toro Garrido, 2007). En consecuencia, el siguiente es un ejemplo del método deductivo en el que a partir de un *enunciado universal* se puede inferir un *enunciado singular*:

FIGURA 7
Ejemplo general del método deductivo
Fuente: elaboración propia

En resumen, el proceso inductivo es lo contrario al deductivo. Según lo anterior, el método inductivo se desarrolla con base en hechos o prácticas particulares, para llegar a organizar fundamentos teóricos. En contraste, el método deductivo basa sus cimientos en determinados fundamentos teóricos, hasta llegar a configurar hechos o prácticas particulares. De acuerdo con ello, la combinación de ambos métodos se resume en el siguiente cuadro conceptual:

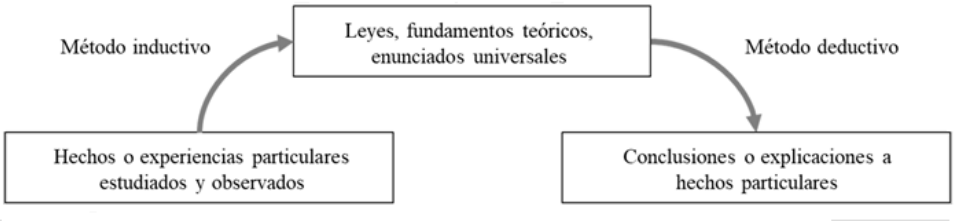

FIGURA 8
Métodos de investigación deductivo e inductivo
Fuente: elaboración propia

Con la finalidad de entender la aplicación de estos métodos en la adquisición y procesamiento de evidencias digitales, antes se deben comprender algunos conceptos claves, tales como *informática forense, hash* y *cadena de custodia*. La siguiente sección desarrolla dichos conceptos.

Informática forense

La informática forense es un área de desarrollo que hace parte de las ciencias de la computación y las ciencias forenses. Mediante diferentes metodologías, esta apoya las investigaciones legales; es definida como "la aplicación de técnicas científicas y analíticas especializadas a infraestructuras y dispositivos tecnológicos que permitan identificar, preservar, analizar y presentar datos que sean válidos dentro de un proceso legal" (Tomás Morales, 2015, p. 116). Lo anterior se traduce en que recolecta mensajes de datos mediante la aplicación de diferentes técnicas, dependiendo del contenedor donde estos se almacenen. Asimismo, una vez recolectados los mensajes de datos y por medio de diferentes metodologías, se identifica la información relevante dentro de ellos. A continuación, se preservan de tal manera que no se modifique ninguno de sus atributos; se analizan para identificar cualquier conducta que vaya en contra de la ley, y se presentan como elementos de materia probatoria para su estudio por parte de las autoridades.

Cabe señalar que la informática forense se puede aplicar en diferentes ámbitos, por lo que se divide en cinco grandes áreas que abarcan de manera general las ciencias de la computación. Estas serán explicadas con base en fuentes teóricas de índole anglosajona, por lo que los encabezados subsecuentes irán en inglés:

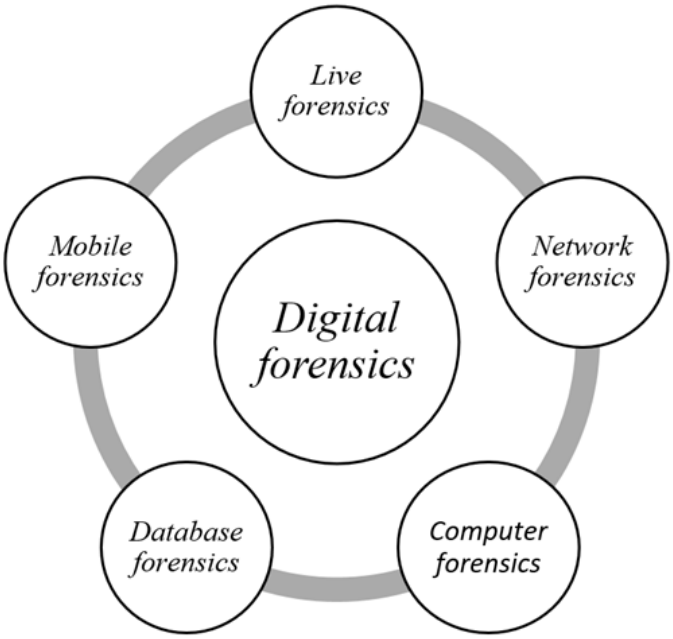

FIGURA 9
Áreas de la informática forense
Fuente: elaboración propia

1. *Live forensics*. El trabajo desarrollado en esta área consiste en la adquisición de la información en el momento en que se está cometiendo una infracción (Vacca y Rudolph, 2011). Lo anterior significa que la detección tiene lugar cuando el sistema o dispositivo se encuentra bajo ataque, o se desarrolla con la transmisión de información relevante para el caso. Un ejemplo es el que se expone a continuación:

    - Cuando se implementan sistemas de *honeynets*[2] y *honeypots*[3] para que, por medio de programas maliciosos, los *hackers*[4] intenten acceder abusivamente a un sistema de información. En ese contexto, el rastro que deje el *hacker* en la *honeynet* o los *honeypots* es evidencia util para adelantar un proceso legal.

2. *Network forensics*. Como punto de partida, cabe señalar que la red o network es esencial para la comunicación entre sistemas. En ese escenario, tal como lo suscita su nombre, el enfoque de la *network forensics* consiste en capturar y analizar eventos de que tienen lugar en la red (Lai, Gu, Jin, Wang y Li, 2010), con la finalidad de detectar información relevante para una investigación. Con el objetivo de dar más claridad sobre este objetivo, cabe señalar que esta área de la informática forense analiza los datos que se encuentran en una red privada, al igual que en la internet; ejemplo de ello es cuando:

    - Se identifica la dirección IP desde la que fue enviado un correo electrónico en el cual se encuentran las instrucciones para duplicar una tarjeta de crédito.
      A su vez, este grupo se puede dividir principalmente en dos ramas que hacen necesario un estudio de fondo de los mensajes de datos encontrados. Usualmente, estas dos ramas solo son usadas cuando hay seguridad de que la información contenida es relevante:

        a) *Email forensics*. Esta rama hace referencia a los gestores de correo electrónico, que se encuentran en internet. En consonancia, su objetivo es preservar todos los correos electrónicos de una persona que es relevante para el proceso legal que se esté adelantando. Por ejemplo:

- La preservación de los mensajes de correo electrónico del usuario *autores@autor.com*; sitio web al cual se puede acceder mediante la dirección URL *mail.autor.com*. Estos mensajes de correo electrónico poseen información que incluye números de cuentas utilizadas para el lavado de dinero.

b) *Web forensics*. Esta rama concierne a un sitio web al cual se puede ingresar por medio de cualquier navegador. En ese contexto, su objetivo es preservar los mensajes de datos que componen un sitio web que contiene información relevante para el proceso legal que se está adelantando. Por ejemplo:

- La preservación de la información del sitio web *www.autor.com* en el cual se encuentra evidencia de un delito financiero.

3. *Database forensics*. En su mayoría, los sistemas de información necesitan un 'lugar' en donde guardar los datos: este 'espacio' de almacenamiento es la base de datos (*database*, en inglés). En esta se encuentra la totalidad de información a la cual se puede acceder mediante una consulta (Easton, 2015). Esta rama de la informática forense se enfoca en obtener dos elementos principales: (a) la información de la base de datos y (b) los procedimientos idóneos para el aseguramiento de la información contenida en ella. Todo lo anterior obedece a la finalidad de determinar una conducta que contradice la ley. Ejemplo de lo anterior es:

   - El acceso a una base de datos bancaria para poder obtener datos sobre transacciones bancarias ilegales.

   En el plano hasta tanto discutido, cabe destacar que la pericia informática forense idónea para la obtención de esta información depende directamente del tipo de sistema de información o de base de datos que se maneje. En estos casos, el investigador no necesariamente debe aprender o tener conocimiento de dichas herramientas —aunque sería lo ideal—; basta con que tenga claro el enfoque de lo que se desea obtener, o del vector de investigación que atañe en el contexto de la arquitectura de datos del sistema de información. Con base en ello, apoyándose en los desarrolladores, el investigador puede consolidar una salida óptima ante adquisiciones de este tipo. Por otra parte, en el plano de investigaciones del tipo descrito, los archivos de datos pueden ser vistos propiamente como mensajes de datos; en consecuencia, su adquisición es muy similar a las que son propias del campo de la *computer forensics*.

4. *Computer forensics*. Esta división de la informática forense se encarga de la obtención de los dispositivos de almacenamiento que hacen parte de los computadores; lo anterior incluye a los computadores de escritorio o servidores. Cabe señalar que algunos autores denominan esta rama de la informática forense *hardware forensics* (Easton, 2015).

   Asimismo, hoy en día la mayoría de personas tienen contacto con un computador. Por ejemplo, al trabajar se utiliza un computador para almacenar información esencial para el desarrollo de las tareas diarias. Asimismo, se puede tener un computador en casa para actividades de entretenimiento tales como ver una película o escuchar música, y todas esas actividades dejan un rastro: una huella que es importante muchas veces para el proceso legal que se esté adelantando. De esta importancia se deriva esta rama, que se enfoca en la información que puede ser adquirida en diferentes dispositivos electrónicos como los ya nombrados computadores. También se tienen en cuenta dispositivos tales como impresoras, escáneres, copiadoras, memorias USB, discos duros internos y externos, al igual que en general cualquier contenedor de mensajes de datos, con la finalidad de procesarlos y analizarlos (Maras, 2015). Dichos exámenes podrían permitir la

obtención de información relevante para procesos legales. Un ejemplo sería el a continuación transcrito:

- Un documento de texto en el cual se encuentren números de tarjeta de crédito utilizados para perpetrar robos.

5. *Mobile forensics*. Esta rama se especializa en obtener diferentes tipos de mensajes de datos provenientes de toda clase de dispositivos móviles. Al respecto, cabe señalar que los dispositivos móviles han ganado popularidad en los últimos veinte años. Los teléfonos inteligentes y tabletas son parte de la vida diaria de cualquier persona, y por eso es muy probable que en esta clase de dispositivos se encuentre mucha información sobre las conductas de un individuo.
   En ese sentido, las fuentes más comunes de mensajes de datos que probablemente podrían servir para el proceso legal que se está adelantando son las llamadas y las conversaciones que constan en la mensajería de texto. En lo anterior se abarcan fuentes tales como WhatsApp, Line, Snapchat y Facebook Messenger. Además, se pueden obtener fotografías, imágenes o mensajes de correo electrónico almacenados en los dispositivos. Por ejemplo:

   - Una fotografía de cheques falsificados da la certeza de que la persona tuvo contacto con los cheques y, en consecuencia, puede estar implicada en el intento de cobro de estos.

De acuerdo con las características de las cinco áreas de informática forense exploradas, así como según los factores de relevancia y volumen previamente abordados, es pertinente afirmar que el comportamiento de estos en las áreas de *network, database, computer* y *mobile forensics* está expresado en la figura 5. En tanto, estos factores se comportan del siguiente modo en el área de *live forensics*:

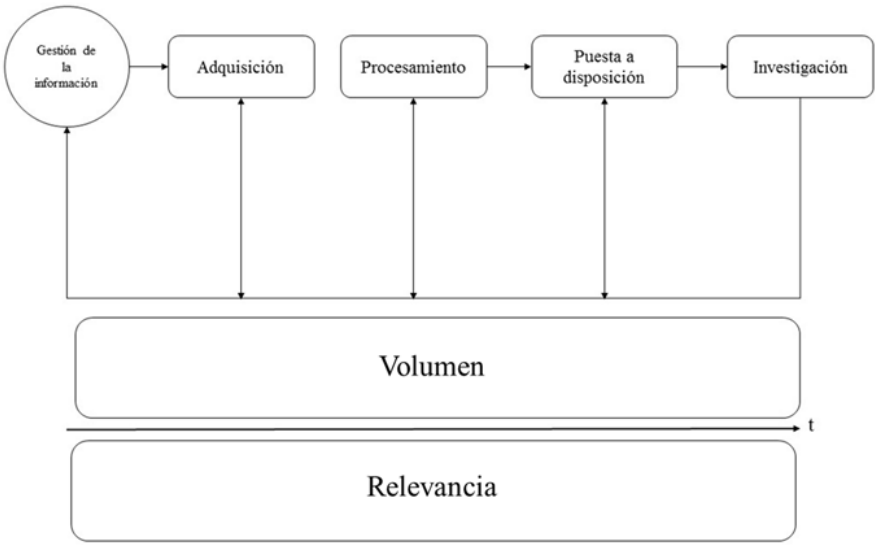

FIGURA 10
Comportamiento de los factores de volumen y relevancia en el transcurso del proceso APPI aplicado al área de *live forensics*
Fuente: elaboración propia

Como se nota en la figura 10, en el área de *live forensics* el volumen se mantiene constante, al igual que la relevancia, dados el procesamiento —que se da en tiempo real— y la naturaleza de la metodología de adquisición de evidencias; en consecuencia, la mayoría de estas pueden ser relevantes desde el inicio hasta el final del proceso. En ese contexto, es importante mencionar que el modelo APPI propuesto está basado en los grupos de *network, database, computer* y *mobile forensics*.

Conceptos claves de la informática forense

Una vez han sido explicados los campos de aplicación de la informática forense, es fundamental abordar algunos conceptos claves que son esenciales en esta disciplina. A continuación, se explicarán diversos aspectos esenciales de la temática aludida.

1. *Imagen forense*. Es una copia espejo de un mensaje de datos adquirido dentro de una investigación. Lo anterior se traduce en que es una copia bit a bit de la información que se encuentra en un dispositivo. Sus dos principales características son (i) que tiene una función *hash* única que garantiza su integridad desde la creación de la imagen, y (ii) que no modifica la información o los mensajes de datos a los cuales se les realiza este procedimiento. En términos tecnológicos, una imagen forense es una evidencia digital que contiene mensajes de datos los cuales pueden ser considerados —aunque no son necesariamente— elementos de materia probatoria: un aspecto el cual depende del análisis que se haga de dichos componentes durante la etapa de investigación en el APPI. En ese sentido, existen dos tipos de imágenes forenses:

a. *Imagen forense completa*. Es una copia espejo de todos los sectores de un dispositivo o equipo. Cabe especificar que estas imágenes pueden ser (i) físicas; es decir que corresponden a la superficie del dispositivo, y (ii) lógicas, es decir que corresponden a las particiones lógicas de este.

b. *Imagen forense parcial*. Es una copia espejo de solo uno o varios mensajes de datos que están contenidos adentro del dispositivo o equipo, y que se ubican en un sector particular. Ejemplo de lo descrito es una imagen forense del archivo de correo electrónico de Outlook (*.pst*) que está ubicado en el escritorio del sistema operativo.

2. *Función o algoritmo* hash. Es un sistema criptográfico compuesto de algoritmos matemáticos cuya entrada es un archivo digital y cuya salida es un valor único para dicho archivo en números hexadecimales. Asimismo, la función *hash* tiene tres características principales:

a. Es una función unidireccional, lo cual quiere decir que solo se puede calcular el valor de salida con base en el valor o archivo de entrada. Asimismo, nunca se podrá llegar al archivo de entrada a partir de su función *hash*.

b. Para su cálculo se puede tener una entrada de N caracteres de extensión. Un ejemplo en tal caso es un archivo de texto con la palabra "hola". No obstante, el resultado siempre tendrá un numero constante de caracteres de salida, según el tipo de algoritmo. Verbigracia: "*a218b6a3b166bc378cfba03ba5015a02*" (32 caracteres). [5]

c. En el marco de ese estándar, se reduce temporalmente la posibilidad de colisiones. Las colisiones ocurren cuando dos archivos totalmente diferentes que tienen valores diferentes tienen el mismo *hash*. En ese escenario, si existieran dos archivos con el mismo *hash*, no se podría tener certeza de cuál de los dos es el archivo al cual se le realizó la imagen forense (Easton, 2015).

Cabe añadir que una función hash puede componerse de dos sistemas criptográficos: MD5 y SHA1. El primero es un sistema hash de 128-bit desarrollado en 1991 por Ron Rivest, notable pues reemplazó al MD4; en tanto, el segundo es un algoritmo de 160-bit desarrollado por la NSA (Easton, 2015). En las distintas etapas del proceso APPI, se utilizan ambos sistemas criptográficos, ya que los dos sistemas presentan colisiones si se procesan individualmente. En suma, se ha comprobado que, si dos archivos tienen colisión en el algoritmo MD5, estos dos mismos archivos no tienen colisión en el sistema SHA1. Lo anterior significa que no hay colisiones para estos dos algoritmos si son implementados conjuntamente.

3. *Contenedor de evidencia digital*. Es el elemento o dispositivo en el que se guardan las evidencias digitales y los mensajes de datos. Existe variedad importante de contenedores entre los que se encuentran el CD, el DVD, el Blu-ray, el USB y el disco duro. En tanto, la cadena de custodia registra la ubicación y los procedimientos hechos a la evidencia que se encuentra en estos contenedores.

4. *Cadena de custodia*. Es un concepto propio de las ciencias forenses. Comprende los procedimientos documentados en los cuales se detalla el estado de la evidencia en cada momento del tiempo, como consta en una bitácora. Lo anterior significa que esta documentación parte del momento inicial, cuando se adquiere la evidencia, y se aboca hasta el momento final, cuando se presenta ante un juez. Si un evento no se registrara en la cadena de custodia, la evidencia no sería fiable (Easton, 2015); por tanto, esta no podría ser utilizada como elemento material probatorio. Existen dos tipos de cadena de custodia:

a. *Cadena de custodia histórica*. Hace referencia a la cadena de custodia de un contenedor de evidencia digital u objeto de custodia física que se encuentra cerrado o finalizado, debido a procedimientos propios del manejo de evidencias.

b. *Cadena de custodia activa*. Alude al estado actual de un contenedor de evidencia digital u objeto de custodia física, y a la información sobre su iniciador, al igual que a otros de sus rasgos pertinentes.

Es importante mencionar que una cadena de custodia activa puede convertirse en una cadena de custodia histórica. Sin embargo, una cadena de custodia histórica no se puede transformar en una cadena de custodia activa.

5. *Objeto de custodia física*. Es el elemento del que se lleva registro del estado de la cadena de custodia. Lo anterior significa que, si la cadena de custodia recae sobre una evidencia digital, esta será el objeto de custodia. Sin embargo, si la cadena de custodia recae sobre un contenedor de evidencia digital, este será el objeto de custodia. En ese sentido, existen tres modelos o casos en concreto para identificar el objeto de custodia física:

a. *Primer caso*: de muchos a uno. Se trasponen las evidencias de muchos contenedores de adquisición a un solo contenedor de evidencia digital. En este caso, las evidencias digitales adquiridas en el campo de análisis son guardadas inicialmente en varios contenedores de evidencia. De acuerdo con el procedimiento de manejo de evidencias, al pasar estas a un solo contenedor de evidencia digital permanente, este será el objeto de custodia física.

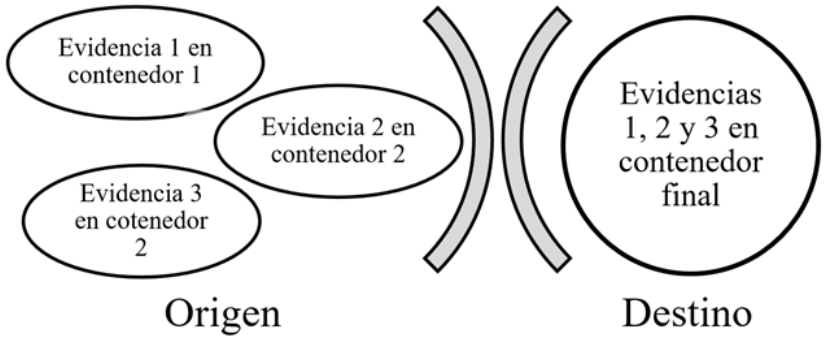

FIGURA 11
Objeto de custodia física: *de muchos a uno*
Fuente: elaboración propia

b. *Segundo caso*: de uno a muchos. Se trasponen las evidencias de un contenedor a muchos contenedores de evidencia digital. En este caso, la evidencia adquirida en el campo de análisis es guardada en varios contenedores de evidencia. Un ejemplo pertinente es que una sola evidencia sea traspasada a varios DVD. De acuerdo con esto, el objeto de custodia física sería la evidencia digital; es decir, la totalidad de discos de DVD.

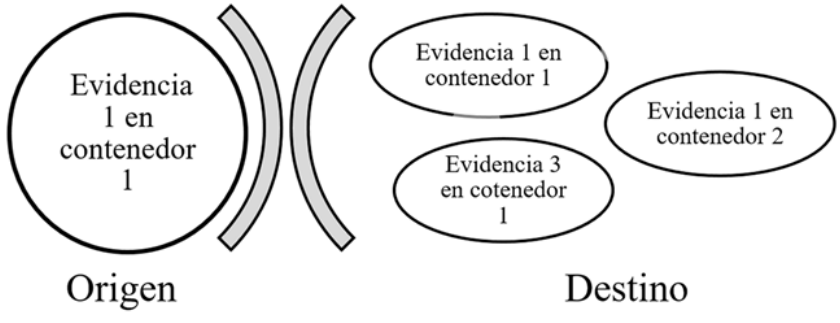

FIGURA 12
Objeto de custodia física: *de uno a muchos*
Fuente: elaboración propia

c. *Tercer caso*: de muchos a muchos. Se trasponen las evidencias de muchos contenedores a muchos contenedores de evidencia digital. En este caso, las evidencias adquiridas en el campo de análisis son guardadas en varios contenedores de evidencia digital. Asimismo, al ser traspasadas quedan almacenadas en varios contenedores de evidencia digital final. De acuerdo con esta situación, el objeto de custodia física serían los varios contenedores finales.

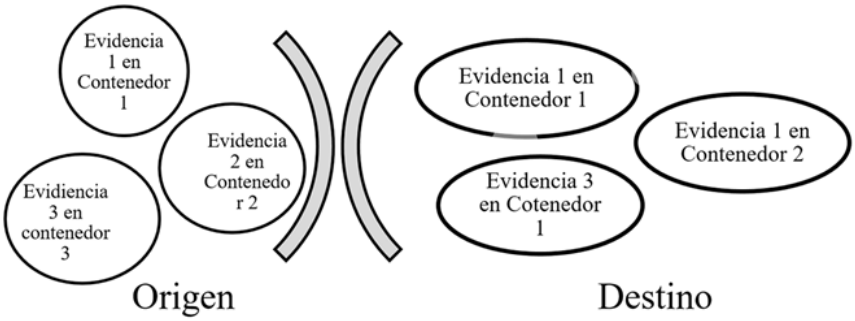

FIGURA 13
Objeto de custodia física: *de muchos a muchos*
Fuente: elaboración propia

Métodos de investigación deductivo e inductivo para la adquisición y el desenvolvimiento del proceso APPI en una evidencia digital

Con la finalidad de relacionar los métodos deductivo e inductivo con las etapas de adquisición y procesamiento de evidencias digitales del modelo APPI —respectivamente—, hay que entender al detalle el proceso de estas dos primeras etapas.

Método de investigación deductivo en las etapas de adquisición

Se puede dar una visión general de la aplicación del método de investigación deductivo a todo el proceso APPI del material aludido. Lo anterior es posible en la medida en que se contempla el volumen de evidencias digitales de la etapa de adquisición como el enunciado universal, así como la relevancia de estas en la etapa de investigación:

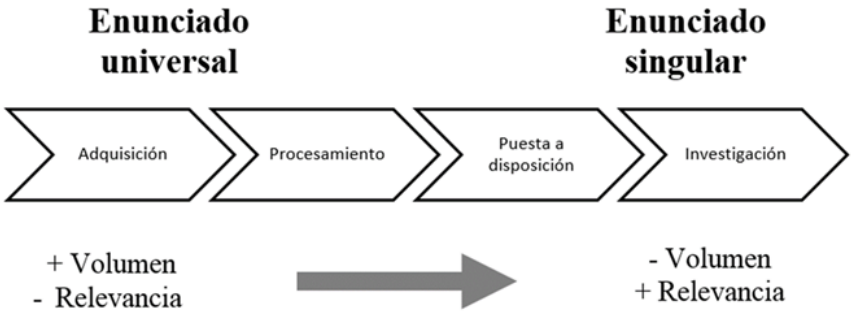

FIGURA 14
Aplicación del método deductivo en el proceso APPI: volumen y relevancia de evidencias digitales
Fuente: elaboración propia

Cabe recordar que el propósito de este artículo es aplicar los dos métodos de investigación hasta tanto aludidos, solamente en las etapas de adquisición y procesamiento. En ese escenario, la etapa de adquisición consta de cuatro muestras que se deben tomar de una evidencia, aunque el número de evidencias adquiridas

varía dependiendo del caso. Por otra parte, el tipo de muestra varía de acuerdo con los vectores de investigación establecidos.

A modo de ejemplo práctico, se aplicará a continuación el modelo implementado con sistemas de información de comunicaciones que corresponden al correo electrónico. Se seleccionaron los sistemas de esta clase, ya que constituyen el tipo de evidencia digital más común:

- *Muestra 1. Adquisición de una imagen forense física completa*. Se realiza una imagen forense de cada sector del disco duro: esta imagen se toma con el dispositivo apagado. En el caso de los computadores, se tiene que extraer el disco duro y conectarlo a otro mediante un bloqueador de escritura.[6]

  Es importante tener en cuenta que, en caso de que el disco duro esté cifrado con alguna clase de software tal como Bitlocker, no será posible acceder a los mensajes de datos en este tipo de imagen. Por lo anterior, resulta útil la adquisición de una segunda muestra según lo indica la siguiente sección. Estas imágenes suelen tener un peso similar a la capacidad del disco duro contenedor, es decir: si el disco duro tiene una capacidad de 500 GB, la imagen forense tendrá un tamaño igual o superior a ese valor.
- *Muestra 2. Adquisición de una imagen forense lógica completa*. En esta etapa se realiza una imagen forense de todas las particiones del disco duro. Al tratarse de una partición lógica, esta permitirá ver los mensajes de datos que aquel contenga, independientemente de si el dispositivo se encuentra protegido por un software de cifrado o no. Con todo, la desventaja de esta muestra es que no incluye otras particiones lógicas del disco físico; en consecuencia, el resultado se complementa con la primera muestra. En cualquier caso, se recomienda la adquisición de las imágenes lógicas de todas las particiones que tenga el dispositivo electrónico. Al igual que en la muestra anterior, en este caso la imagen tendrá un peso igual o superior a la capacidad de la partición del contenedor del dispositivo adquirido.
- *Muestra 3. Adquisición de una imagen forense parcial de un sistema de información*. En este caso, si se quiere únicamente el archivo de correo electrónico de Outlook (*.pst*), se procede a buscar solamente ese archivo mediante el explorador del equipo, con el propósito de realizarle la imagen forense respectiva.
- *Muestra 4. Adquisición de una imagen forense parcial de los mensajes de datos que son posibles elementos de materia probatoria*. En esta etapa, se realiza una imagen forense únicamente al mensaje de datos específico dentro del sistema de información. Lo anterior significa que esta acción se aplica únicamente a los mensajes de correo electrónico (*.msg*) que contienen los *.pst* que son relevantes para la investigación. En consecuencia, se les realiza una imagen forense, con las finalidades de preservarlos y de mantener su integridad, confidencialidad y disponibilidad.

A la sazón del proceso descrito, la siguiente figura resume los tipos de imágenes tomadas en cada muestra de adquisición de evidencias digitales mediante el método deductivo:

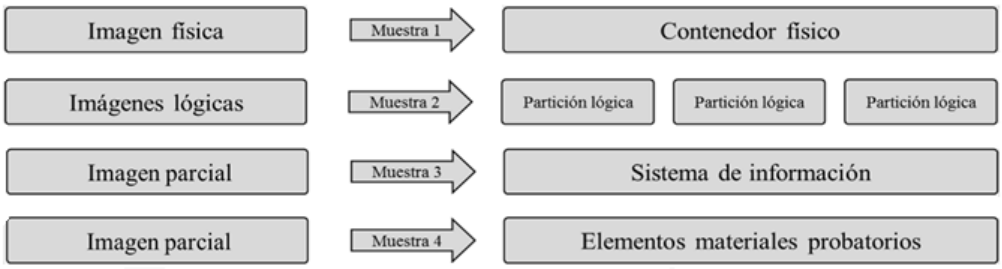

FIGURA 15
Tipo de imagen tomada mediante la muestra de una evidencia digital
Fuente: elaboración propia

Con la finalidad de comprender el papel del proceso más claramente, es pertinente añadir que el *método deductivo* se usa en la *etapa de adquisición*. En esta, se tienen grandes volúmenes de información —que corresponden al enunciado universal—, y, conforme avanza la adquisición de muestras, se obtienen el mensaje o los mensajes de datos que sirven como elementos materiales probatorios —que coinciden con el enunciado singular—:

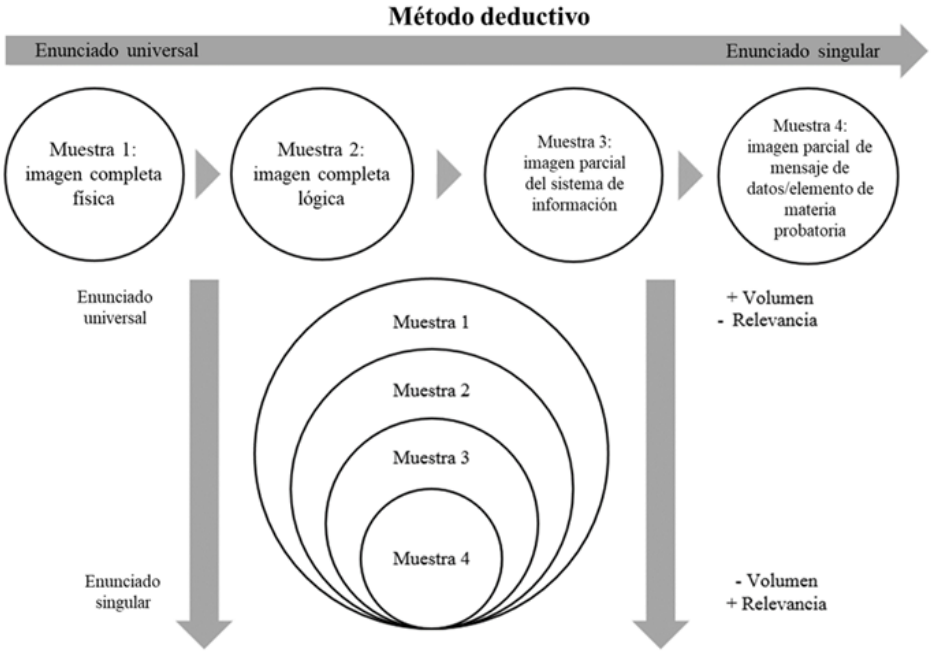

FIGURA 16
Método deductivo utilizado en la etapa de adquisición
Fuente: elaboración propia

Método de investigación inductivo en las etapas de procesamiento

Por su parte, la fase de procesamiento consta de tres etapas. En ese plano, el número de evidencias procesadas varía dependiendo del número de evidencias adquiridas. Cabe añadir que, como en la explicación anterior, la situación en la que está basado el ejemplo corresponde a un caso de complejidad alta:

- *Etapa 1. Procesamiento de la muestra 3*. Como lo evidencia la figura 17, la muestra 3 es la imagen forense parcial que contiene los mensajes de datos del sistema de información —en el caso anterior, el

archivo de correo electrónico de Outlook *.pst*—. En este caso, el análisis se ejecuta con la finalidad de identificar todos los mensajes de correo que contiene. Esta etapa no suele durar más de dos o tres días.

- *Etapa 2. Procesamiento de la muestra 2*. Como consta en la figura 15, la muestra 2 es la imagen forense lógica completa que contiene toda la información del disco duro. En esta etapa, se pueden recobrar mensajes de datos borrados: por ejemplo, es posible recuperar otros archivos de correo electrónico *.pst* borrados. Cabe aclarar que el objetivo de esta fase no es solamente recuperar los archivos borrados, sino identificar todos los mensajes de datos pertinentes. Asimismo, esta etapa de procesamiento solamente suele aplicarse si en la primera no se logró identificar elementos de materia probatoria.
- *Etapa 3. Procesamiento de la muestra 3*. Como lo evidencia la figura 15, la muestra 3 es la imagen forense física completa que contiene los mensajes de datos del disco duro al cual se le hizo la copia espejo. En esta etapa, al igual que en la anterior, se puede recuperar información borrada, identificar mensajes de datos relevantes y establecer cuáles archivos están dañados y cuáles sectores no pueden ser leídos. Esta etapa de procesamiento solamente suele aplicarse si en la segunda muestra no se logra identificar elementos de materia probatoria; ejemplos de ello son aquellos casos en los que el sistema operativo ha sido formateado.

Con el objetivo de comprender más claramente su rol en el proceso descrito, es pertinente recordar que el *método inductivo* se usa en la *etapa de procesamiento.* En esta, se tienen pequeños volúmenes de información —que coinciden con el enunciado singular— y, conforme avanzan las etapas de procesamiento, se realiza una búsqueda más extensa del elemento material probatorio —que corresponde al enunciado universal—

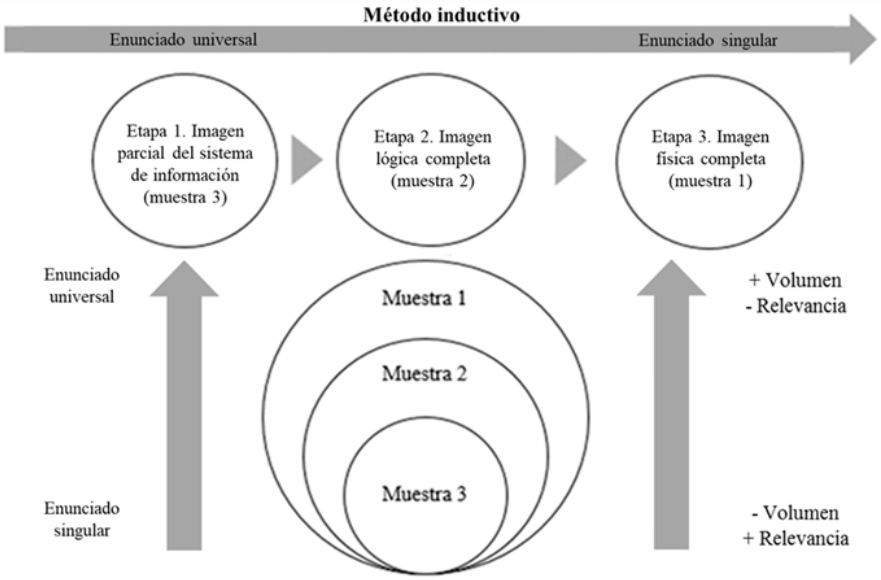

FIGURA 17
Método inductivo utilizado en la etapa de procesamiento
Fuente: elaboración propia

Cabe aclarar que la muestra 4 no se incluye en la figura 17 porque a esta se le aplica un procesamiento corto, ya que no es necesaria la recuperación de información. Además, no es relevante el tiempo de procesamiento de este proceso, dado que es prácticamente instantáneo. En tanto, los resultados de ambos métodos de investigación son presentados en la siguiente sección.

Resultados obtenidos en la práctica: cifras y estadísticas

En la práctica profesional y el trabajo de campo, se han podido obtener resultados que refieren al volumen de información recolectada en la etapa de adquisición, y al volumen de información procesada en la etapa de procesamiento. Aclaradas las dimensiones esenciales del estudio, es posible adentrarse en la propuesta que sustenta la presente exposición.

Con el propósito de observar los efectos prácticos de la aplicación de los métodos inductivo y deductivo en el método APPI —específicamente en las etapas de adquisición y procesamiento—, se tomarán como base cuatro evidencias que fueron adquiridas con protocolos académicos en equipos de cómputo aleatorios. Dichas recolecciones fueron llevadas a cabo con sistemas de información de administración de correos electrónicos. En tanto, la metodología se desarrolló asegurando la obtención de resultados mediante el método científico clásico. Con ello, se formularon comparaciones de las diferentes muestras tomadas de acuerdo con el acápite denominado "Método de investigación deductivo en las etapas de adquisición" en el presente documento.

Etapa de adquisición y método deductivo en la práctica

El objetivo de esta sección es mostrar el volumen de mensajes de datos recolectados en la etapa de adquisición de cuatro evidencias distintas tomadas en muestras de equipos de cómputo aleatorios. Tal pesquisa se postula con la finalidad de reconocer el método inductivo en ese proceso.

Para esta etapa, las principales variables de valoración son (a) el tamaño de la muestra y (b) el tiempo que requirió su adquisición. En ese sentido, hay que tener en cuenta que los tiempos de adquisición varían de acuerdo con el dispositivo al que se le haga la muestra. Asimismo, entre los principales elementos que influyen en estos tiempos se encuentran el procesador, la memoria volátil, el tipo de disco duro y la versión de puerto USB. En seguida figura una tabla con esas cifras:

TABLA 2
Volumen de mensajes de datos recolectados de cuatro evidencias distintas

Tipo de muestra	Evidencia 1		Evidencia 2		Evidencia 3		Evidencia 4	
	Tamaño	Tiempo	Tamaño	Tiempo	Tamaño	Tiempo	Tamaño	Tiempo
Muestra 1. Imagen forense física completa	140,67 GB	5 h 31 min 26 s	152,03 GB	5 h 18 min 43 s	71,00 GB	5 h 51 min 27 s	256,69 GB	4 h 25 min 3 s
Muestra 2. Imagen forense lógica completa	140,61 GB	9 h 49 min 47 s	152,02 GB	19 h 49 min 49 s	70,09 GB	16 h 1 min 0 s	256,69 GB	6 h 9 min 16 s
Muestra 3. Imagen forense parcial a sistema de información	24,93 GB	1 h 46 min 20 s	14,15 GB	0 h 34 min 27 s	4,25 GB	0 h 15 min 17 s	40,60 GB	1 h 40 min 25 s
Muestra 4. Imagen forense parcial a elemento material probatorio (EMP)	187,10 KB	0 h 0 min 13 s	4,95 MB	0 h 0 min 5 s	980,9 KB	0 h 0 min 1 s	412,75 KB	0 h 0 min 1 s
Total	**306,21 GB**	**17 h 7 min 43 s**	**318,20 GB**	**25 h 43 min 4 s**	**145,34 GB**	**22 h 07 min 45**	**553,98 GB**	**12 h 14 min 45 s**

Fuente: elaboración propia

Con base en la tabla 2, se dan las siguientes conclusiones, organizadas según los ejes temáticos correspondientes:

- *Tiempo de adquisición*. El tiempo total adquisición de las cuatro muestras es mucho mayor que el lapso que tomaría la adquisición de cualquiera de las muestras si estas hubieran sido tomadas individualmente.
- *Tamaño de la muestra*. Se nota la disminución del tamaño de las muestras durante el proceso de adquisición de evidencias, el cual se desarrolló de acuerdo con el método deductivo y lo explicado en el acápite previo titulado "Métodos de investigación deductivo e inductivo para la adquisición y el desenvolvimiento del proceso APPI en una evidencia digital".

Las cifras comparativas reflejan datos de gran relevancia. En la primera etapa, el volumen de información adquirida es mucho mayor, en comparación con el resultado del hallazgo de los elementos materiales probatorios que se nota en la última etapa del proceso. A continuación, se puede apreciar la proporción del tamaño de los mensajes de datos relevantes, con respecto al total de información que se adquirió en cada muestra:

TABLA 3
Porcentaje de mensajes de datos relevantes por cada muestra de cuatro evidencias distintas

Evidencia No.	Porcentaje de mensaje de datos relevantes		
	Muestra 1	Muestra 2	Muestra 3
1	0,00013300633%	0,00013306308%	0,00075050140%
2	0,00325593633%	0,00325615051%	0,03498233216%
3	0,00138154930%	0,00139948637%	0,02308000000%
4	0,00016079707%	0,00016079707%	0,00101662562%

Fuente: elaboración propia

De acuerdo con la tabla 3, en la mayoría de evidencias y las diferentes muestras tomadas, el porcentaje de elementos de materia probatorio no supera al 0,03% en el mejor de los escenarios. En tanto, en la peor de las situaciones este no es mayor al 0,0001% de toda la información adquirida. Según lo expuesto, es claro que al ser la muestra 3 la más pequeña, la probabilidad de encontrar un elemento material probatorio en ella es mucho mayor.

Etapa de procesamiento y método deductivo en la práctica

En la etapa de procesamiento, la variable principal es el tiempo de procesamiento —al igual que en la sección anterior—. En otras palabras, esta coincide con cuánto tiempo se demora el procesamiento de cada una de las muestras de las cuatro evidencias anteriormente mencionadas. Asimismo, al igual que en la etapa de adquisición, los tiempos de procesamiento varían de acuerdo con dos situaciones:

- *Modalidad general de procesamiento*. Una modalidad de procesamiento se desglosa como la combinación de múltiples opciones de procesamiento. En ese sentido, entre las principales opciones se pueden encontrar el cálculo de *hash*, la expansión de archivos compuestos, la inclusión de archivos eliminados y la identificación de extensiones erróneas, entre otros aspectos. En tanto, en cuanto compete a los tiempos de procesamiento, estos son dados a continuación en modalidad de recuperaciones profundas de datos; en suma, se encuentran dos opciones —de procesamiento— esenciales en ese contexto: (a) la restauración de datos (*data carve*, en inglés) y (b) la restauración de metadatos (*meta carve*, en inglés). Asimismo, cabe añadir que estas dos opciones exigen tecnología de punta para generar resultados en tiempos óptimos para el área de la informática forense.
- *Características técnicas del equipo de procesamiento*. Es reseñable que los tiempos de procesamiento dependen completamente del hardware y el software disponibles. Actualmente, el sistema de procesamiento de evidencias digitales puede procesar un gigabyte de información en treinta y nueve minutos, en una modalidad de procesamiento de recuperación de datos profunda.

A continuación, se presentan los tiempos de procesamiento de acuerdo con los tamaños de las evidencias anteriormente citadas:

TABLA 4
Proyección de tiempos de procesamiento por volumen
de mensajes de datos de cuatro evidencias distintas

Tipo de Muestras	Evidencia 1		Evidencia 2		Evidencia 3		Evidencia 4	
	Tamaño	Tiempo	Tamaño	Tiempo	Tamaño	Tiempo	Tamaño	Tiempo
Muestra 1. Imagen forense física completa	140,67 GB	91 h 43 min 55 s	152,03 GB	99 h 22 min 35 s	71,0 GB	46 h 15 min 0 s	256,69 GB	167 h 45 min 45 s
Muestra 2. Imagen forense lógica completa	140,61 GB	91 h 40 min 5 s	152,02 GB	99 h 22 min 30 s	70,09 GB	45 h 56 min 25 s	256,69 GB	167 h 45 min 45 s
Muestra 3. Imagen forense parcial de sistema de información	24,93 GB	16 h 20 min 45 s	14,15 GB	9 h 20 min 15 s	4,25 GB	3 h 16 min 25 s	40,60 GB	26 h 39 min 0 s
Muestra 4. Imagen forense parcial a elemento material probatorio (EMP)	187,10 KB	Procesamiento corto	4,95 MB	Procesamiento corto	980,9 KB	Procesamiento corto	412,75 KB	Procesamiento corto
Total	**306,21 GB**	**199 h 44 min 45 s 8,3 días**	**318,2 GB**	**208 h 5 min 20 s 8,6 días**	**145,34 GB**	**95 h 57 min 50 s 4 días**	**553,98 GB**	**362 h 10 min 30 s 15,1 días**

Fuente: elaboración propia

Con base en la anterior tabla, se puede inferir que la eficiencia de la depuración basada en elementos materiales probatorios se obtiene a partir de un procesamiento corto de la muestra 4. Ahora bien, si no se llegase a encontrar la totalidad de elementos materiales probatorios, sería necesario hacer un procesamiento completo a la muestra 3, cuyo tamaño es menor en comparación con el de las muestras 1 y 2. Cabe señalar que lo anterior haría eficiente tal procesamiento, ya que en el caso expuesto no se tiene el mismo volumen de datos que las imágenes físicas y lógicas adquiridas inicialmente.

A modo de conclusión

De acuerdo con los resultados que han sido expuestos, es posible reflexionar en torno a los siguientes puntos:

1. Aplicar el método deductivo en la etapa de adquisición del proceso APPI permite tener copias de respaldo de las muestras sucesivas. Estas no solo son elementos tecnológicos, sino que son evidencias de respaldo para los casos en los que dichas muestras sean presentadas a un juez o autoridad.
2. La aplicación del método inductivo en la etapa de procesamiento del procedimiento APPI permite optimizar los tiempos de depuración, toda vez que las muestras adquiridas son procesadas en orden ascendente. Lo anterior significa que estas se organizan de la de menor tamaño a la de mayor tamaño.
3. Es importante la identificación de los vectores de investigación que se desea adquirir. Esto permite planear las muestras y los demás elementos propios de cada adquisición y procesamiento de evidencias.
4. El procesamiento se induce en el orden inverso al de la adquisición de las muestras. Asimismo, se aumenta el nivel en la medida en que los vectores de investigación no sean encontrados. Además, si ya se han hallado suficientes vectores de investigación, los niveles superiores de adquisición no serán necesarios; no obstante, estos sí brindarán soporte a las muestras más pequeñas.
5. Resulta idóneo estandarizar los mecanismos que abarcan la adquisición y el procesamiento de evidencias digitales dados por medio del modelo APPI, e implementar mecanismos basados en los métodos de investigación y la teoría informática forenses, así como los relativos a la seguridad de la información. Lo anterior permite robustecer los resultados obtenidos, además de optimizar la obtención de dichos resultados y la utilización de recursos para tal propósito.
6. El modelo de adquisición y procesamiento —deductivo e inductivo, respectivamente— puede variar en investigaciones de delitos financieros en las que no se incluyan comunicaciones de correo electrónico como vectores de investigación; se alude, en este caso, a situaciones que se asientan directamente en sistemas de información contable.

Referencias

Bernal Torres, C. A. (2006). *Metodología de la investigación: para la administración, economía, humanidades y ciencias sociales*. Ciudad de México: Pearson Educación.

Decreto número 1747 de 2000, *Por el cual se reglamenta parcialmente la ley 527 de 1999*, Presidencia de la Répública de Colombia § Presidencia de la República de Colombia, Bogotá, 11 de septiembre de 2000.

Duquenoy, J. S. (2008). *Ethical, legal and professional issues in computing*. London: Cengage Learing EMEA.

Easton, C. (2015). *CCFP Certified Cyber Forensics Professional Certification. Exam Guide*. United States: Mc Graw HIll.

Electronic Discovery Reference Model. (2005). *Creating practical resources to improve E-Discovery & information governance*. Recuperado de http://www.edrm.net/resources/edrm-stages-explained

Hurtado León, I. y Toro Garrido, J. (2007). *Paradigmas y métodos de investigación en tiempos de cambio*. Caracas: CEC.

Joshi, S. y Sardana, A. (2011). *Honeypots: A new paradigm to information security*. Boca Raton: CRC.

Lai, X., Gu, D., Jin, B., Wang, Y. y Li, H. (2010). An efficient searchable encryption scheme and its application in network forensics. En *Forensics in telecomunications, information, and multimedia* (pp. 66-78). Shanghai: Institute for Computer Sciences, Social Informats and Telecomunications Engeneering.

Ley 527 de 1999, *Por medio de la cual se define y reglamenta el acceso y uso de los mensajes de datos*, Presidencia de la República de Colombia § Presidencia de la República de Colombia, Bogotá, 18 de agosto de 1999.

Maras, M. H. (2015). *Computer forensics: cybercriminals, laws and evidence*. United States: Jones & Bartlett Learning.

Martínez Chánez, V. M. (1998). *Fundamentos teóricos para el proceso del diseño de un protocolo en una investigación*. Ciudad de México: Plaza y Valdes.

McKeon, R. (2010). *Richard McKeon*. Recuperado de http://www.richardmckeon.org/

Prieto Castellanos, B. (2014a). *La piratería de películas: cómo combatirla*. Bogotá: Praci.

Prieto Castellanos, B. (2014b, diciembre 13). Valor probatorio de un mensaje de datos. *Asuntos Legales*. Recuperado de https://www.asuntoslegales.com.co/analisis/bayron-prieto-513031/el-valor-probatorio-de-un-mensaje-de-datos-2201771

Ramos Álvarez, B. y Ribagorda Garnacho, A. (2004). *Avances en criptología y seguridad de la información*. Madrid: Díaz de Santos.

Tomás Morales, S. D. (2015). *Retos del derecho ante las nuevas amenazas*. Madrid: Dykinson.

Vacca, J. y Rudolph, K. (2011). *Systems forensics, investigation, and response*. Ontario: Jones & Bartlett Learning.

Notas

1 Los formatos de archivos digitales compuestos son aquellos que tienen información estructurada por niveles en su interior cuya forma es la de otros archivos digitales.
2 La *honeynet* es una red física compuesta de *honeypots* (Joshi y Sardana, 2011).
3 Los *honeypots* configuran un sistema de información de recursos cuyo valor radica en el uso no autorizado o ilícito de tal recurso (Joshi y Sardana, 2011).
4 Un *hacker* es considerado un programador experto, en tanto que es capaz de reescribir un código para personalizarlo y mejorarlo (Duquenoy, 2008).
5 Hash MD5 adquirido de un mensaje de datos.
6 Dispositivo que permite acceder a la información sin modificar ninguno de sus bytes.
* Artículo de investigación.

Para citar este artículo: Prieto Castellanos, B. J. (2017). El uso de los métodos deductivo e inductivo para aumentar la eficiencia del procesamiento de adquisición de evidencias digitales. *Cuadernos de Contabilidad, 18*(46). 1-27. https://doi.org/10.11144/Javeriana.cc18-46.umdi