UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Escuela Técnica Superior de Ingeniería Informática

Definición de una metodología para el análisis informático forense en entornos IoT.

Trabajo Fin de Máster

Máster Universitario en Ciberseguridad y Ciberinteligencia

AUTOR/A: Sorribas Segura, Adrián

Tutor/a: Nachiondo Farinós, Teresa del Montecarmelo

CURSO ACADÉMICO: 2021/2022

Resum

En l'actualitat, el Internet de les Coses (IoT, *Internet of Things*) està cada vegada més present al nostre voltant, tant en l'àmbit laboral com en el personal. IoT presenta la capacitat de recollir gran quantitat de dades del nostre entorn, de la nostra forma de viure o del nostre cos, sent esta tecnologia testimoni del nostre dia a dia. Esta característica pot ser molt valuosa en una investigació informàtica forense, no obstant, comporta importants desafiaments.

En primer lloc, es te que tindre en consideració la naturalesa heterogènia de l'entorn IoT, en el que podem identificar-hi tres fonts principals d'evidències: el núvol, la xarxa y la diversitat de dispositius IoT. La utilització del núvol y la xarxa en el sistemes IoT pot comportar que els investigadors forenses poques vegades puguin realitzar l'adquisició d'evidències per ells mateixos, depenent normalment dels proveïdors de servicis y administradors o proveïdors de la xarxa, dificultant així el poder comprovar la integritat de les evidències recollides.

Es per tant, la gran quantitat y varietat de dades generades pels sistemes IoT, junt amb un emmagatzemament distribuït entre les diferents ubicacions, afegeixen un repte en l'àmbit legal ja que, en la cerca d'evidències, es te que complir tant les lleis relatives a la privacitat del propietari de les dades, com les lleis específiques respecte al país d'on es troben emmagatzemades les mateixes. En conseqüència de les problemàtiques citades anteriorment, es proposa com a treball final de màster la definició d'una metodologia per a la realització d'una investigació informàtica-forense en entorns IoT, remarcant sobre tot els aspectes relatius a la integritat de les evidències y els aspectes legals que impliquen.

Paraules clau: Informàtica Forense, Sistemes IoT, Integritat de les evidències, Legalitat

Resumen

En la actualidad, el Internet de las Cosas (IoT, *Internet of Things*) está cada vez más presente en nuestro entorno, tanto laboral como personal. IoT presenta la capacidad de recabar grandes cantidades de datos de nuestro entorno, de nuestra forma de vida o de nuestro cuerpo, siendo esta tecnología testigo de nuestro día a día. Esta característica puede ser muy valiosa en una investigación informática forense, sin embargo, entraña importantes desafíos.

En primer lugar, debe considerarse la naturaleza heterogénea del entorno IoT, en el que podemos identificar principalmente tres fuentes de evidencias: la nube, la red y la diversidad de dispositivos IoT. La utilización de la nube y la red en los sistemas IoT puede conllevar que los investigadores forenses rara vez puedan realizar la adquisición de evidencias por sí mismos, dependiendo normalmente de los proveedores de servicios y administradores o proveedores de red, dificultándose así el poder probar la integridad de las evidencias recabadas.

Del mismo modo, el gran volumen y variedad de datos generados en los sistemas IoT, junto a un almacenamiento distribuido entre diferentes ubicaciones, añaden un reto también a nivel legal, ya que, en la búsqueda de evidencias, debe cumplirse tanto las leyes relativas a la privacidad del propietario de los datos, como las leyes específicas al respecto del país donde se encuentren almacenados los datos. Por todo lo anteriormente indicado, se propone como trabajo fin de máster la definición de una metodología para

la realización de una investigación informática forense en entornos IoT, haciendo hincapié en los aspectos relativos a la integridad de las evidencias y los aspectos legales que impliquen.

Palabras clave: Informática Forense, Sistemas IoT, Integridad de las evidencias, Legalidad

Abstract

Currently, the Internet of Things (IoT) is increasingly present in our environment, both at work and personally. IoT presents the ability to collect large amounts of data from our environment, our way of life or our body and for this reason, this technology is becoming essential to our daily chores. This feature can be invaluable in a computer forensic investigation, however, it comes with significant challenges.

First of all, the heterogeneous nature of the IoT environment must be considered, in which we can mainly identify three sources of evidence: the cloud, the network and the diversity of IoT devices. The use of the cloud and the network in IoT systems are responsible that forensic investigators are rarely able to acquire evidence by themselves, usually depending on service providers and administrators or network providers, thus making it difficult to prove the integrity of the evidence collected.

In the same way, the vast volume and variety of data generated in IoT systems, together with distributed storage between different locations, also add a challenge at a legal level, since in the search for evidence, both the laws related to the privacy of the owner of the data, such as the specific laws regarding the country where the data is stored must be considered. For all the above, the definition of a methodology for conducting a computer forensic investigation in IoT environments is proposed as a master's thesis, emphasizing the aspects related to the integrity of the evidence and the legal aspects that they involve.

Key words: Digital Forensic, IoT systems, Evidence Integrity, Law

Índice general

Índice de figuras

Índice de tablas

CAPÍTULO 1
Introducción

Relojes inteligentes, *Smart TV*, asistentes virtuales, sensores de temperatura, etc. son dispositivos que están día a día influyendo directamente en nuestra forma de percibir la realidad. Estos dispositivos forman parte de la vida cotidiana ya que facilitan enormemente y simplifican los quehaceres humanos.

Disponemos de sensores y actuadores que permiten mejorar la eficiencia temporal y energética en un modelo de sociedad que cada vez exige progresar hacia un mundo más sostenible y eficaz.

Estos dispositivos pertenecen a la familia del IoT *(Internet of Things)* y en los últimos años se ha visto incrementado de forma exponencial su uso. El desarrollo y los avances en la tecnología 5G[1] y el despliegue masivo de sensores para fomentar las ciudades inteligentes auguran un aumento todavía más pronunciado al existente actualmente.

Sin embargo, todo este despliegue de dispositivos interconectados también muestra ciertas carencias y vulnerabilidades. Se está construyendo un modelo de sociedad cada vez más interconectado y que también representa varias carencias en el ámbito de la ciberseguridad. Se está desarrollando el despliegue de esta tecnología de una forma muy veloz con lo que conlleva que los ciberdelincuentes también se aprovechen de este despliegue para buscar vulnerabilidades en una tecnología emergente y tan diversa.

Durante el primer semestre del año 2021, ya se observó que se duplicaron respecto al año 2020 los ataques hacia infraestructuras y dispositivos IoT [10]. Es por este motivo, por lo que el análisis forense en estas infraestructuras es cada vez más habitual, siendo los dispositivos IoT una fuente más de evidencias en un incidente.

Sin embargo, es de reseñar que aunque existen diversas metodologías propuestas para la realización del análisis forense, éstas no siempre se adecuan a las características de los entornos IoT. Es por ello, que este proyecto tiene como objetivo identificar las carencias de las metodologías ya propuestas y presentar una nueva metodología específica para entornos IoT, que en la medida de lo posible mejore las carencias detectadas.

1.1 Motivación

La motivación de este proyecto viene dada por varios factores relativos al sector del IoT. En primer lugar, en los últimos años se ha producido un incremento exponencial en la instalación y uso de dispositivos IoT tanto en la industria como en la vida cotidiana. Además, se espera que este crecimiento sea aún mayor en los próximos años llegando

[1] Tecnología de comunicación que permite aumentar la velocidad de conexión de forma exponencial, reducir latencias y la interconexión de mayor nombre de dispositivos.

a tener interconectados más de 25 mil millones de dispositivos en el año 2030 como se muestra en la Figura 1.1.

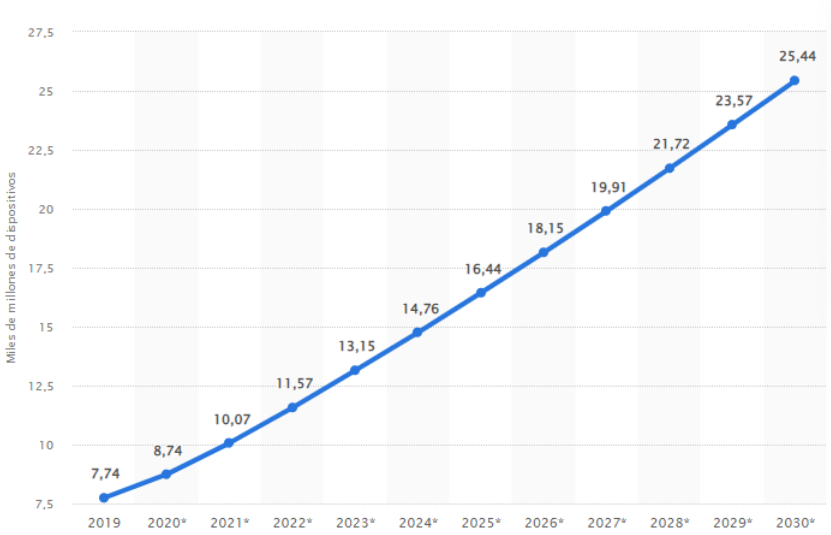

Figura 1.1: Incremento de la tecnología IoT. [35]

Este incremento de dispositivos interconectados también se verá afectado por el incremento exponencial de ciberamenazas que se está produciendo en un mundo cada día más interconectado. Este incremento ya se vio reflejado en 2019 y 2020, cuando el porcentaje de ataques a dispositivos IoT aumentó un 218 % y un 66 % respectivamente [34]. La Figura 1.2, muestra la comparativa de amenazas de los años 2020 y 2021. Como puede observarse, las amenazas han seguido aumentando un 6 % en 2021, con una cifra récord de 60,1 millones de incidentes de seguridad.

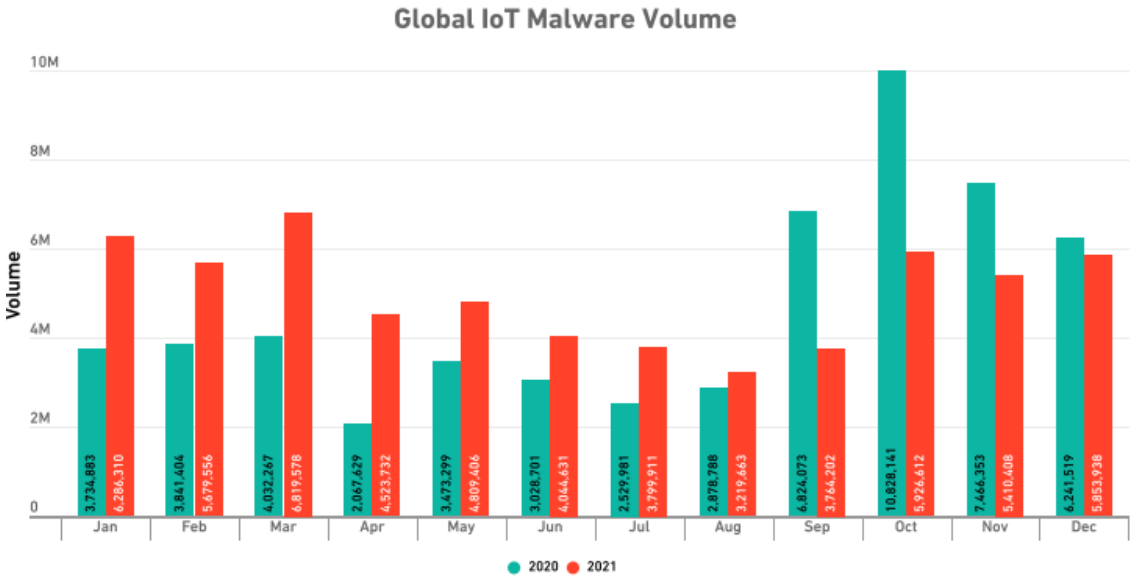

Figura 1.2: Comparativa de ataques IoT entre el año 2020 y 2021. [34]

Si desglosamos el total de ataques que afecta a la tecnología IoT, se puede observar en la Figura 1.3 cómo los sectores objetivo son mayoritariamente el sector público y educativo.

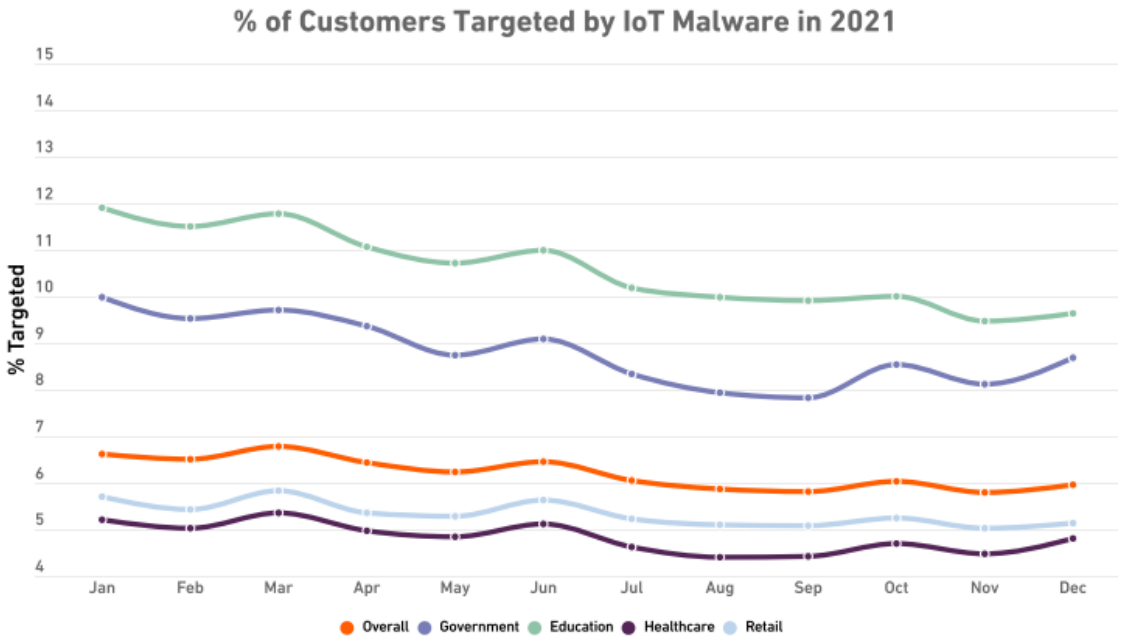

Figura 1.3: Ataques IoT por sectores. [34]

Dado el incremento de ciberamenazas en entornos IoT, cada vez son más las investigaciones forenses en las que se ven envueltos estos dispositivos.

De este modo, surge la necesidad de definir una nueva metodología en la que poder fundamentar el análisis forense en dispositivos IoT.

1.2 Objetivos

El principal objetivo de este proyecto es la definición de una metodología para la realización de una investigación informática forense en entornos IoT, haciendo hincapié en los aspectos relativos a la integridad de las evidencias y los aspectos legales que impliquen.

El objetivo principal se puede desglosar en los siguientes objetivos parciales:

- Investigar y analizar las tres principales fuentes de evidencias en IoT: la nube, la red y la diversidad de dispositivos.
- Abordar y analizar la dificultad de poder probar la integridad de las evidencias recabadas y asegurar la cadena de custodia.
- Analizar la normativa legal de acceso a los datos en IoT debido a un almacenamiento distribuido entre diferentes ubicaciones.
- Identificar, definir y comparar las principales metodologías para la realización del análisis forense.

- Definir un nuevo *Framework*[2] para la realización del análisis forense que asegure y subsane la mayoría de las problemáticas existentes en el análisis forense en IoT.

El principal resultado esperado tras finalizar este proyecto es conseguir definir una nueva metodología que subsane o mejore la mayoría de carencias detectadas en las metodologías ya propuestas para el análisis forense en dispositivos IoT.

1.3 Estructura de la memoria

La estructura de la memoria presente en este documento va a ser distribuida en diferentes capítulos cuyo contenido se detalla a continuación:

En el capítulo 2 se muestra en profundidad el estado del arte del proyecto junto con una puesta en contexto de la situación actual y de las diversas alternativas existentes. También se enumeran y desarrollan las tecnologías utilizadas y se realiza un análisis en profundidad de las principales carencias que existen en el análisis forense en el ámbito de dispositivos IoT.

En el capítulo 3 se detallan y analizan las principales metodologías para el análisis forense existentes en la actualidad, y se muestran las diferencias y similitudes entre ellas.

En el capítulo 4 se muestra el análisis y diseño para la definición de una metodología para la realización del análisis forense en el que se subsane las carencias detectadas en los marcos actuales.

En el capítulo 5 se aúnan las conclusiones obtenidas al finalizar el proyecto.

En el capítulo 6 se detallan los trabajos futuros para la continuación y mejora del proyecto.

Para concluir, también se adjuntan los anexos A, B, C y D con información complementaria a la puesta en marcha del proyecto.

[2]Marco de trabajo que compone una serie de estándares en un campo.

CAPÍTULO 2
Contexto y descripción del proyecto

El estado del arte se puede entender desde dos puntos de vista en este proyecto: Los inicios de la tecnología IoT y los inicios de la informática forense.

En este capítulo se tratará de esclarecer e investigar los fundamentos sobre los que se basa el presente proyecto.

Además, también se describirá el proyecto y se analizará la situación actual de la tecnología, describiendo y analizando las carencias existentes.

2.1 Tecnología IoT

2.1.1. Orígenes y evolución del IoT

La breve historia del IoT [26] se debe entender respecto a tres grandes hitos en la historia de la informática: el desarrollo de la comunicación inalámbrica, los MEMS (Sistemas Microelectromecánicos) y el desarrollo de los microservicios de Internet. Es pues, la breve historia de la evolución de una primera conexión M2M (Máquina a máquina) a finalmente una red de miles de dispositivos interconectados entre sí sin interacción humana.

Debemos remontarnos hasta el año 1874, para encontrar el que es considerado el primer experimento de telemetría de la historia el cual se realizó en el Mont Blanc[1] y que consistía en la obtención de datos meteorológicos desde unos dispositivos situados en el citado lugar y que transmitían los datos a París a través de un enlace de radio.

En los años 90 llegó Internet, un gran hito en el que Berners-Lee[2] estableció exitosamente una primera comunicación entre un cliente HTTP *(Hypertext Transfer Protocol)* y un servidor a través de Internet. Esta comunicación fue el nacimiento del WWW *(World Wide Web)*.

Gracias a la gran revolución del nacimiento de Internet, y la creciente conexión de clientes a la red, fue en 1990 cuando se considera que se conectó el primer dispositivo IoT de la historia. [19] Este dispositivo fue conectado por John Romkey[3] y se denominó como "tostadora conectada". Este dispositivo conectado a la red permitía que cualquier

[1] Montaña situada en los Alpes Franceses y uno de los puntos más elevados de la Unión Europea.
[2] Científico nacido en Londres el 8 de junio, es considerado el padre fundador de la Word Wide Web.
[3] Desarrolló la primera pila TCP/IP en la industria mientras estaba en el Instituto de Tecnología de Massachusetts.

usuario desde cualquier ordenador del mundo pudiera encender, apagar y ver el tiempo de tostado: había nacido el IoT.

Posteriormente fueron apareciendo más proyectos IoT como el XCoffe desarrollado por estudiantes de la Universidad de Cambridge, que conectaron la primera cámara web controlando la máquina de café del departamento para verificar el estado del café. Estos dispositivos utilizaban el protocolo TCP/IP y se controlaban a través de SNMP[4].

Poco a poco se fue popularizando esta tecnología y fue finalmente en la era 2000, cuando aparecieron las primeras redes inalámbricas y cuando creció exponencialmente el número de dispositivos IoT conectados a Internet.

En 2008, se superó un nuevo hito en la historia del IoT, pues los dispositivos conectados a Internet ya superaban al número de personas conectadas.

El término IoT se popularizó entre la población en el año 2009 cuando el profesor Kevin Ashton[5] lo introdujo al público y del que citó textualmente lo siguiente:

«El Internet de las Cosas tiene el potencial de cambiar el mundo, tal como lo hizo el Internet. Tal vez incluso más.», Kevin Ashton

A partir de esta etapa se populariza todavía más el uso de esta tecnología y se empiezan a diseñar coches, electrodomésticos e incluso realizar trasplantes que incorporan dicha tecnología.

En 2011 se introduce el estándar IPv6, un nuevo estándar que entre otras mejoras respecto a IPv4, proporciona alrededor de 1.028 veces más direcciones que IPv4, lo cual beneficiaba en gran medida al Internet de las cosas. También en ese mismo año, el término IoT aparece por primera vez en el *Hype Cycle* de Gartner[6] como se muestra en la Figura 2.1, una gran noticia para el futuro de esta tecnología.

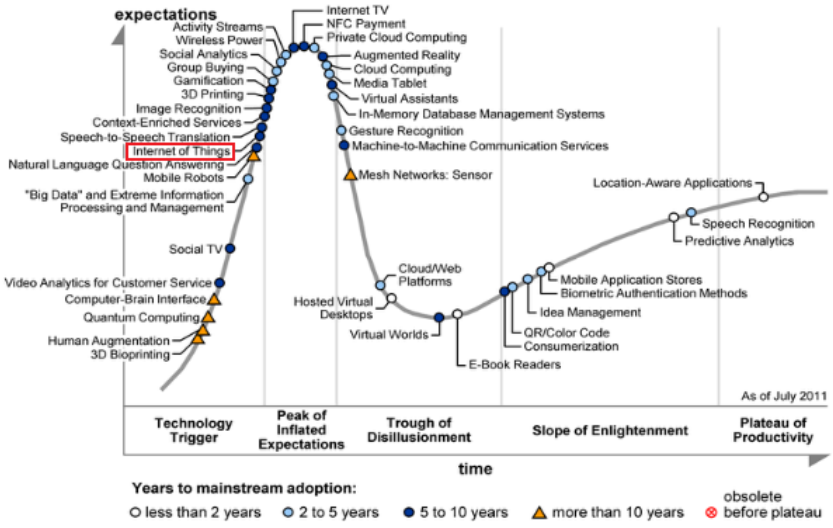

Figura 2.1: Primera aparición del IoT en el Hype Cycle. [7]

[4] *Simple Network Management Protocol*. Protocolo en la capa de aplicación y que basa su funcionamiento en IP para intercambiar datos.

[5] Nacido en Reino Unido en 1968, se considera uno de los padres del IoT.

[6] Informe anual en el que se muestra una gráfica donde están representadas la madurez, adopción y aplicaciones de nuevas tecnologías con mayor potencial y expectativa.

Posteriormente, en el año 2014, grandes empresas tecnológicas del mundo se unen para crear la iniciativa IoT-GSI Slobal Standars[7] con el fin de crear un estándar en esta tecnología a nivel mundial y poder compartir información.

Todo eran buenas noticias para esta tecnología emergente hasta 2016, cuando aparece el primer *Malware* dedicado exclusivamente a dispositivos IoT llamado MIRAI, cuyo objetivo se basa en infectar routers y cámaras IP. Este *Malware* basa su estrategia en recopilar las contraseñas por defecto de los dispositivos IoT, que los usuarios no cambian, con el fin de crear una botnet[8] para generar ataques DoS (Denegación de Servicio) contra servicios de terceros.

En la actualidad, nos encontramos en un momento de completo auge en el mundo IoT y en el que las amenazas de los ciberataques no cesan.

En la Figura 2.2, se puede observar el incremento exponencial que está teniendo el uso de la tecnología IoT y las previsiones de crecimiento que les esperan respecto a las tecnologías convencionales.

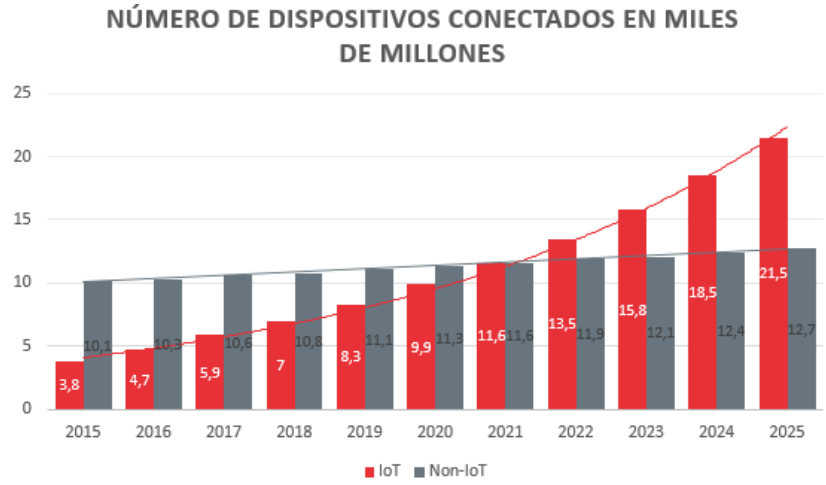

Figura 2.2: Número de dispositivos conectados en miles de millones. [14]

Es por este motivo por el que, debido al gran crecimiento del uso de esta tecnología y la consecuente generación masiva de datos, se espera que en un ámbito tan diverso como es el del IoT, se incremente masivamente el nombre de ciberataques contra los dispositivos mencionados.

Con el fin de realizar este proyecto, se toma este punto como referencia: una tecnología en auge amenazada por un mundo cada vez más hostil en referencia a los ciberataques y con unos fabricantes de dispositivos IoT, y usuarios, poco concienciados para hacer frente a estos ataques. Es en este punto cuando crece la importancia del término análisis forense en IoT.

2.1.2. Funcionamiento del IoT

Como se ha mencionado en el punto anterior, el IoT se trata de una tecnología en la que su uso ha aumentado exponencialmente en la última década y que permite hacer accesibles desde cualquier punto del planeta a cualquier dispositivo que esté conectado.

[7] Internet of Things Global Standards Initiative.
[8] Conjunto de ordenadores o dispositivos infectados y controlados por el atacante de forma remota.

En el siguiente apartado, se va a explicar el funcionamiento de esta tecnología así como también se enumerarán y explicarán las aplicaciones más populares en las que se hace uso.

Además, en la Figura 2.3, se puede observar un resumen sobre la taxonomía del IoT, incluyendo los aspectos clave que se van a desarrollar en los siguientes apartados. En esta figura, se resumen todos los puntos decisivos que afectan al IoT y que se deben tener en cuenta al realizar cualquier análisis forense sobre estos dispositivos.

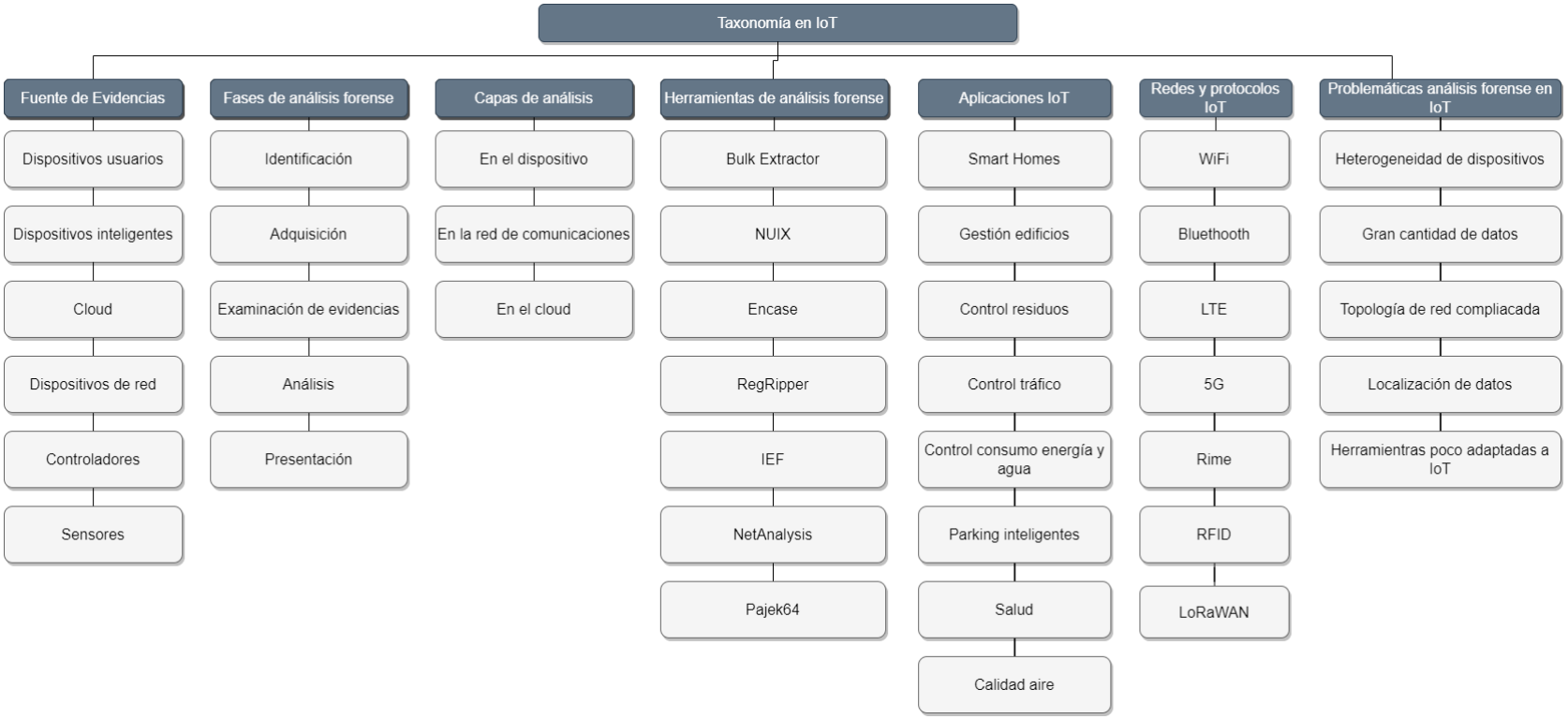

Figura 2.3: Esquema global tecnología IoT. Fuente propia.

Ciclo de vida del IoT

En el ciclo de vida del IoT se observa la forma en la que fluye la información en todas las etapas del proceso, desde su origen hasta la obtención de resultados como se observa en la Figura 2.4.

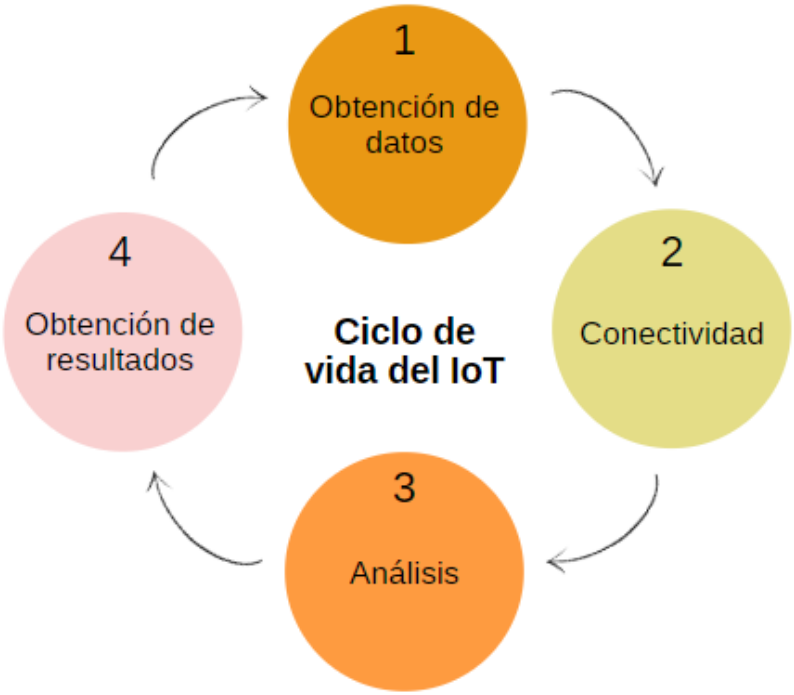

Figura 2.4: Ciclo de vida del IoT. Fuente propia.

De este modo, observamos las cuatro etapas más relevantes en la tecnología IoT: La obtención de datos a través de los sensores, la conectividad a través de la que fluyen los datos, el análisis y procesamiento de la información y, finalmente, la obtención de resultados.

- **Obtención de datos:** Se trata del primer paso en el IoT y consiste en la obtención de datos generalmente a través de sensores como, por ejemplo, sensores de temperatura, movimiento, acelerómetros, etc. Así pues, se genera la información del entorno para posteriormente ser procesada.
- **Conectividad:** Una vez generados los datos de interés, se procede al envío de los mismos a la nube o a algún servidor de la red a través de diferentes medios de conectividad, ya sean inalámbricos como redes Wifi, Bluetooth, LTE, LoRaWAN, 5G, etc. o a través de redes cableadas como Ethernet. En este caso, los dispositivos IoT buscan en sus comunicaciones una compensación entre consumo de energía, alcance y ancho de banda.
- **Análisis:** Una vez ha llegado la información generada en el primer paso, ésta es analizada y procesada para verificar si la información obtenida cumple ciertos requisitos y si es compatible con los valores estimados.

- **Obtención de resultados:** Una vez se ha realizado el análisis de los datos, se obtienen los resultados que pueden determinar una actuación automatizada, como por ejemplo la activación del sistema de refrigeración cuando la temperatura observada supera un límite superior o, simplemente mostrarse a través de la interfaz de usuario si no requiere ninguna actuación.

Aplicaciones del IoT

La tecnología IoT tiene un gran abanico de aplicaciones, motivo por el cual se ha convertido en una de las tecnologías más punteras en la actualidad. En este apartado se van a mencionar unas de las más relevantes actualmente y en las que ha supuesto una completa revolución:

- **Casas inteligentes:** En la actualidad se trata de uno de los campos en los que más se ha aumentado el uso del IoT. Interconectividad de dispositivos, asistentes personales y la domótica ha hecho que cada vez el IoT se haya adentrado más en nuestros hogares hasta finalmente poder controlar la totalidad de los dispositivos del hogar desde un simple teléfono inteligente.

- **Gestión de edificios:** En la gestión de edificios también ha supuesto un gran avance ya que permite mejorar y monitorizar de forma más efectiva la conservación de los mismos evitando así deficiencias estructurales, polución, sistemas eléctricos y canalizaciones.

- **Control de residuos:** Mediante esta aplicación se permite monitorizar de forma eficiente la ocupación de los contenedores de basura, así como la temperatura en estos en los meses de calor intenso. De esta forma se puede mejorar la logística de recolección de residuos haciéndola más eficiente.

- **Congestión del tráfico y logística:** El tráfico en las grandes urbes ya supone un grave problema en la actualidad. Es pues, un buen control de las congestiones permite agilizar la economía y mejorar la logística haciéndola más eficiente.

- **Consumo de energía y agua:** Evitar el desperdicio en el consumo de energía y agua es fundamental para evitar sobrecostes en las ciudades. De este modo, gestionar de forma eficiente el uso de éstas, permite evitar desperdiciar, por ejemplo, el consumo de agua innecesario en una fuga no localizada o incluir sensores en los sistemas de iluminación de las ciudades para que se enciendan únicamente cuando sea necesario.

- **Párkings inteligentes:** Dado que cada vez el uso de vehículos utilitarios es mayor, una gestión eficiente de las plazas de aparcamiento es vital. De esta forma, se puede monitorizar las plazas libres de un párking notificando al usuario con anterioridad para que sepa si es necesario desplazarse o evitar ir a ese lugar.

- **Salud:** La monitorización y utilización de dispositivos IoT es fundamental en este ámbito. Desde el uso de marcapasos hasta sistemas de monitorización en los diabéticos, permite mejorar notoriamente la vida de los pacientes que padecen enfermedades crónicas.

- **Calidad del aire:** El control de la calidad del aire es fundamental en un mundo en el que cada vez existe más contaminación. Monitorizar estos parámetros permite conocer las zonas donde existe más polución y de esta forma poder tomar medidas.

2.2 Análisis forense digital

2.2.1. Orígenes y evolución del análisis informático forense

Los inicios de la informática forense se remontan a los inicios de los años 80, cuando los ordenadores personales se popularizaron al poder ser adquiridos por los consumidores. Esta popularización del uso de los dispositivos de cómputo personales junto con la aparición de Internet, hizo que cobrara importancia la aparición de la figura del analista informático-forense.

Más tarde, en el año 1984, [15] el FBI[9] creó un programa llamado Programa de Medios Magnéticos, y que más tarde, en 1991, pasaría a llamarse CART *(Computer Analysis and Response Teams)* [36], un equipo de respuesta y de análisis de evidencias informáticas.

Otro hito importante fue en 1990 cuando se fundó la compañía *New Technologies* por Michael Anderson. *New Technologies* fue una de las primeras empresas dedicada exclusivamente al análisis forense de evidencias informáticas.

Más tarde, en el año 1995, se estableció la IOCE *(International Organization on Computer Evidence)*, una organización compuesta de varias agencias gubernamentales involucradas en investigaciones informático forenses y que trataba de ser un foro de intercambio de información sobre las investigaciones de delitos relacionados con la informática. [11]

El auge de esta disciplina vino de la mano de la primera década del siglo XXI donde cada vez las agencias como el CART tenían que analizar mayor cantidad de datos debido al aumento sustancial de equipos informáticos.

Finalmente, en la actualidad, el gran auge de los dispositivos IoT ha hecho que se precisen de nuevas técnicas y metodologías en el análisis forense-informático ya que las reglas han cambiado desde los orígenes.

Actualmente existen miles de millones de dispositivos interconectados entre sí, analizando datos en tiempo real y geolocalizados en distintas partes del mundo. Además, se añade una nueva dificultad, las evidencias pueden encontrarse en el *cloud*[10] y encontrarse en una localización distinta al dispositivo que las genera, incluso en un país con una legalidad distinta. Estos son los retos del futuro: la legalidad, la localización y la integridad de las evidencias, tres aspectos en los que se va a centrar y focalizar en el presente proyecto.

2.2.2. Estructura y desarrollo del análisis forense

El análisis forense digital se trata de la disciplina que tiene la finalidad de identificar, adquirir, examinar, analizar y presentar una serie de evidencias obtenidas ante la corte correspondiente.

Esta disciplina tiene como objetivo final la obtención de ciertas evidencias digitales con el fin de esclarecer unos hechos que han ocurrido en el pasado en un dispositivo digital.

De igual manera, en el proyecto se va analizar esta disciplina en el ámbito del IoT puesto que es un ámbito emergente y en el que queremos enfatizar la relevancia del mismo.

[9] Buró Federal de Investigaciones.

[10] Tecnología que permite el acceso remoto tanto a archivos como a programas con el fin de no depender de equipos locales. [32]

Fases del análisis forense

Según define el NIST *(National Institute of Standards and Technology)*, acepta el análisis forense como una ciencia que trata de aunar en 5 fases, como se muestra en la Figura 2.5, el proceso de obtención de tratamiento de evidencias:

Figura 2.5: Fases del análisis forense digital. Fuente propia.

- **Identificación:** Todo análisis forense digital empieza por la fase de identificación en la que el analista examina el contexto en el que se ha producido el incidente, los elementos involucrados y trata de identificar las evidencias y realizar la correlación de los hechos acontecidos. En esta fase también se realiza un análisis de las posibles herramientas que harán falta para la investigación y adquisición de las evidencias.
- **Adquisición:** En esta fase, el analista trata de adquirir las evidencias. Se trata de una fase crítica ya que se debe utilizar el método adecuado de obtención para no contaminar o destruir las evidencias.
- **Evaluación de evidencias:** En esta fase el analista examina las evidencias obtenidas en el punto anterior. El análisis puede realizarse en un laboratorio o en el propio sitio. En este punto cabe remarcar que es muy importante garantizar la trazabilidad de las evidencias obtenidas y garantizar la cadena de custodia[11], con el fin de que sean válidas ante una corte.
- **Análisis:** En este punto los analistas investigan y correlacionan las evidencias obtenidas para poder esclarecer los hechos ocurridos.
- **Presentación:** En este punto, el analista presenta las conclusiones obtenidas respecto a las evidencias adquiridas. Cabe remarcar que es muy importante mantener la

[11]Concepto utilizado para garantizar la trazabilidad de las evidencias obtenidas y probar de este modo que no han sido manipuladas.

integridad de las evidencias para que, de esta forma, se pueda dar validez a las conclusiones obtenidas.

Adicionalmente, se puede dar el caso de una sexta fase al principio de todo llamada preprocesamiento, en la que el investigador organiza y prepara todos los elementos necesarios para realizar el posterior análisis forense. Se trata pues, de una fase de preparación y planificación de todo el proceso de análisis.

Posibles fuentes de las evidencias

Respecto a la generación de evidencias, debido a la naturaleza de los dispositivos IoT se pueden originar desde diferentes fuentes [25]:

- **Dispositivos finales de usuarios:** Se trata de los dispositivos más habituales de generación de evidencias, como equipos portátiles, equipos de sobremesa, impresoras, etc. Estos equipos pueden proporcionar una gran cantidad de datos valiosos.
- **Dispositivos inteligentes:** En los últimos años estos dispositivos se han incrementado exponencialmente y son de los que más información puede aportar. Se trata de dispositivos como relojes, enchufes, asistentes virtuales, televisores inteligentes, etc. que almacenan información de hábitos cotidianos y que son de gran valor para investigaciones futuras.
- ***Cloud:*** La nube es un lugar en el que se puede encontrar gran cantidad de evidencias debido a la gran cantidad de datos almacenados que contiene y la trazabilidad de los mismos. Sin embargo, presenta serios problemas, así como la obtención de datos en esta, ya que hay que tener en cuenta factores legislativos respecto al país de almacenamiento de los datos, permisos de la organización que los almacena, etc.
- **Dispositivos de red:** Uno de los lugares de los que se pueden obtener evidencias muy valiosas es en los dispositivos de la red y analizando el tráfico de esta. De esta forma se puede analizar la trazabilidad de las comunicaciones y las conexiones realizadas desde un dispositivo.
- **Controladores:** En el ámbito IoT, es uno de los dispositivos más relevantes puesto que son el principal objetivo de los atacantes. Los controladores son los encargados de recoger la información desde los sensores, procesarla y tomar decisiones.
- **Sensores:** Se trata de los dispositivos que reciben la información del entorno. Estos dispositivos representan dificultades importantes en el momento de la obtención de evidencias debido a sus complicadas localizaciones.
- **Actuadores:** Los actuadores son los dispositivos finales encargados de realizar las acciones físicas determinadas por el controlador.

Tipos de análisis forense en IoT

En el análisis forense digital focalizado a dispositivos IoT, se pueden distinguir tres tipos de análisis [9] dependiendo de la casuística del entorno en el que se produzca el incidente:

- **Análisis en el propio dispositivo IoT:** Se obtienen las evidencias en el propio dispositivo IoT. En este análisis se pueden obtener evidencias físicas del propio dispositivo o de la memoria almacenada en el mismo.

- **Análisis en la red de comunicaciones:** El análisis de las comunicaciones es una fuente muy importante de evidencias ya que los dispositivos durante sus comunicaciones, ya sean permanentes o puntuales, dejan *logs* que permiten rastrear las comunicaciones realizadas. De este modo se puede analizar el tráfico de red o de los dispositivos de red intermedios para observar el comportamiento que han tenido o tienen los dispositivos.

- **Análisis en la nube:** Muchos de estos dispositivos, por su naturaleza, almacenan sus datos y los comunican constantemente con un servidor o directamente con la nube haciendo que gran cantidad de la obtención de evidencias se deba realizar en la misma.

2.3 Diferencias entre análisis forense digital tradicional, análisis forense en IoT y análisis forense en el *cloud*

Con el fin de entender este proyecto, es de gran relevancia entender las diferencias existentes entre el análisis forense digital convencional, el análisis forense digital en dispositivos IoT y el análisis forense en el *cloud*.

De esta forma, como se muestra en las Tablas 2.1, 2.2 y 2.3, se observan grandes diferencias entre los tres campos mencionados respecto al nombre de equipos a analizar, los propietarios de los mismos, la cantidad de datos generados a analizar, el tipo de redes y protocolos empleados y, finalmente, la fuente de generación de evidencias.

Por ejemplo, no será el mismo número de equipos a analizar en un entorno *cloud* o tradicional que en el ámbito del IoT, puesto que entran en consideración gran cantidad dispositivos de distinta naturaleza. Además, el volumen de datos generado en el ámbito IoT y *cloud* es considerablemente mayor que en los análisis convencionales de equipos de uso personal.

Características análisis forense tradicional	
Elemento	Análisis Forense Tradicional
Número de dispositivos	Baja cantidad de dispositivos a analizar
Propietarios de los dispositivos	Usuarios finales, empresas o instituciones
Cantidad de datos	Hasta pocos terabytes de datos
Tipo de redes y protocolos	Ethernet, WiFi y Bluetooth con protocolos IPv4 e IPv6
Fuente de evidencias	Portátiles, PC, smartphones, redes sociales, email, etc.

Tabla 2.1: Características análisis forense tradicional.

Características análisis forense en IoT	
Elemento	Análisis Forense en IoT
Número de dispositivos	Gran cantidad de dispositivos a analizar
Propietarios de los dispositivos	Usuarios finales, empresas o instituciones
Cantidad de datos	Hasta pocos exabytes de datos
Tipo de redes y protocolos	Protocolos y redes RFID, LoRaWAN, VSAT, 5G, LTE y Rime
Fuente de evidencias	Sensores

Tabla 2.2: Características análisis forense en IoT.

Características análisis forense en el *cloud*	
Elemento	Análisis Forense en el *cloud*
Número de dispositivos	Baja cantidad de máquinas virtuales y servidores
Propietarios de los dispositivos	Proveedores *cloud*, usuarios, empresas o instituciones
Cantidad de datos	Hasta pocos petabytes de datos
Tipo de redes y protocolos	Privadas/públicas con protocolos SSHP, MTP, XMPP, etc.
Fuente de evidencias	Máquinas virtuales, electrónica de red y servidores

Tabla 2.3: Características análisis forense en el *cloud*.

Asimismo, se puede concluir que un análisis forense en el IoT no se adapta a los estándares normativos existentes de hoy en día para el análisis forense digital convencional puesto que difieren notoriamente en el número de dispositivos a analizar y sobre todo a la cantidad de datos a analizar.

Finalmente, cabe destacar que al tener una gran cantidad de dispositivos a analizar y cada uno perteneciente a un fabricante, también se hace más compleja la fase de adquisición de evidencias puesto que los datos estarán estructurados de forma distinta para cada dispositivo.

2.4 Retos que afrontar en el análisis forense en IoT

En esta sección se muestra una de las mayores problemáticas encontrada en el análisis forense en IoT y en las que se debe fundamentar la metodología que se propondrá con el fin de subsanar estas deficiencias identificadas. [12]

2.4.1. Diferentes tipos de dispositivos y fabricantes

Una de las fortalezas del IoT se basa en la gran variedad de dispositivos que se pueden interconectar. Asimismo, también conlleva grandes problemáticas puesto que cada tipo de dispositivo tiene un modo de funcionamiento totalmente distinto.

De este modo, una nueva metodología tiene que tener en cuenta esta complejidad y poder adaptarse a los distintos dispositivos y modalidades de funcionamiento existentes.

2.4.2. Gran cantidad de dispositivos

Otra gran característica del IoT es la gran cantidad de dispositivos existentes. A diferencia del análisis forense digital tradicional, nos encontramos ante un modelo en el que coexisten gran cantidad de dispositivos y en el que se tienen que tener en cuenta en el momento de realizar un análisis.

De este reto, se tiene que tener en cuenta la gran cantidad de dispositivos existente y la consiguiente correlación entre los dispositivos.

2.4.3. Gran cantidad de datos generados

Uno de los grandes retos detectados es la inmensa cantidad de datos que se genera a través del uso del IoT. En una nueva propuesta de metodología de análisis se debe tener en cuenta la gran cantidad de datos generados para poder procesarlos y obtener únicamente la información necesaria.

2.4.4. Topología de red complicada

Reconocer la topología de red en un entorno IoT no es una tarea sencilla, puesto que únicamente con los *logs* que proporcionan los dispositivos no son suficientes para determinarla. Asimismo, descubrir y analizar el tipo de protocolos utilizados también es un desafío importante.

2.4.5. Integridad de las evidencias

Garantizar la integridad de las evidencias con el fin de mantener la cadena de custodia, es uno de los puntos clave en las investigaciones forenses para garantizar la validez de las evidencias.

Este reto debe tratarse con detenimiento para la definición de una nueva metodología puesto que el análisis forense se fundamenta en poder mantener y garantizar la integridad y la validez de las evidencias.

2.4.6. Creación de imágenes de dispositivos IoT

Un reto importante en el análisis forense digital es poder extraer imágenes de los dispositivos IoT para copiar su contenido a un fichero para poder analizar evidencias.

2.4.7. Localización de los datos

Uno de los puntos conflictivos es el de la localización de los datos de los dispositivos. Los datos pueden encontrarse en diferentes centros de datos que están localizados en diferentes regiones de las que se originan los mismos. De esta forma, se muestra la problemática de la legislación aplicable para la obtención de evidencias en un análisis forense.

Además, la organización propietaria de estos centros también puede decidir si se permite la obtención de los datos que tienen almacenados en sus infraestructuras.

2.4.8. Formato de los datos

En la definición de una nueva metodología, se debe tener en consideración que cada dispositivo IoT genera los datos en un formato diferente y posteriormente, se almacena en otro formato en la nube.

2.4.9. Identificación de los dispositivos

La identificación de los dispositivos IoT y de los usuarios que acceden a ellos, es fundamental para poder realizar una trazabilidad de los hechos ocurridos. De este modo, es de vital importancia poder observar la trazabilidad de las acciones que realiza cada dispositivo y poder vincularlas al mismo o a un usuario identificado.

2.4.10. Herramientas de análisis forense

Finalmente, se encuentra el problema de las herramientas de análisis forense tradicional. Éstas basan sus búsquedas en entornos homogéneos en el que la estructuración de los datos es similar. Sin embargo, en el IoT se encuentra la dificultad de tratarse de un

entorno heterogéneo en el que cada dispositivo estructura los datos de un modo y en el que la utilización de las herramientas convencionales puede no ser suficiente.

2.5 Carencias y problemática en el campo forense-informático de dispositivos IoT

Una vez puesto en contexto y basándonos en lo expuesto en los puntos anteriores, podemos identificar los aspectos más relevantes para este proyecto.

De este modo, en este proyecto se trata de identificar y remediar las siguientes problemáticas observadas en el campo forense-informático de dispositivos IoT:

- **Incremento exponencial de dispositivos IoT:** Cada vez existe un mayor número de dispositivos IoT interconectados y que están generando datos constantemente.

- **Baja seguridad y poca concienciación del fabricante y usuario:** Los usuarios no están concienciados de la información que recopilan los dispositivos ni tampoco de los riesgos que puede acarrear la exfiltración de estos. Además, una mala práctica del uso de esos dispositivos como es la de no cambiar las contraseñas por defecto hace todavía más vulnerables estos dispositivos.

- **Ausencia de actualizaciones de *firmware*:** Esta problemática es doble. Por una parte, muchos fabricantes no prestan soporte de *software* a los dispositivos desarrollados por lo que muchos tampoco ofrecen actualizaciones de *firmware* periódicas con el fin de subsanar vulnerabilidades. Por otra parte, los usuarios tampoco prestan importancia en la aplicación de los parches de seguridad en los dispositivos por lo que la gran mayoría de ellos, aún teniendo actualizaciones de seguridad, no acaban actualizándose a versiones no vulnerables.

- **Incremento de ataques e incidentes de seguridad:** En los últimos años se han incrementado exponencialmente los ciberataques, hecho que afecta también a los dispositivos IoT ya que se encuentran interconectados a Internet.

- **Falta de un marco legislativo y un estándar de análisis para IoT:** No existe un estándar a seguir respecto al análisis forense en dispositivos IoT, si bien es cierto que hay varios modelos en desarrollo, aun no hay uno oficial que seguir [20]. Además, la ausencia de un marco legislativo respecto a la obtención de evidencias acarrea un problema respecto a la interpretación legislativa sobre las leyes de cada estado, ya que los dispositivos IoT pueden estar deslocalizados.

- **Diversidad de dispositivos IoT:** Desde un reloj inteligente hasta una nevera, existe una gran cantidad de dispositivos IoT y cada uno recopila una serie de datos y funciona de forma distinta, haciendo que sea todavía más difícil la obtención de evidencias al no servir las técnicas convencionales.

- **Asegurar la integridad de evidencias y garantizar la cadena de custodia:** Este es sin lugar a duda, uno de los puntos más relevantes en este proyecto. Una de las mayores problemáticas encontradas es que en el caso de obtener evidencias de algún dispositivo IoT, se debe poder demostrar ante la autoridad competente que ésta evidencia es lícita y no ha sido modificada. Es pues, un gran reto en la definición de una metodología ya que es muy importante poder verificar la validez de las evi-

dencias. En este punto se va a valorar cómo la tecnología del Blockchain[12] podría contribuir a mejorar la integridad de las evidencias.

- **Muchos de los *frameworks* existentes son más teóricos que prácticos:** Finalmente, nos encontraremos durante el desarrollo del proyecto que muchos de los *frameworks* encontrados son puramente teóricos y apenas se han puesto en práctica lo cual complica su valoración al no tener casos prácticos donde poder ver el comportamiento de los distintos modelos.

2.6 Análisis de posibles soluciones

Las posibles soluciones para solventar las problemáticas detectadas en el punto anterior, se deben afrontar desde el punto de vista de la definición de una nueva metodología que contemple todas las deficiencias del sistema actual, pudiendo así garantizar de forma intrínseca la integridad de las evidencias y permitan garantizar la cadena de custodia de las mismas y a su vez fortalecer su validez.

De este modo, para definir esta metodología de análisis mejorada, se deben tener en cuenta las singularidades del IoT y encontrar un punto de inflexión en el que se puedan mejorar los modelos existentes en la actualidad.

En el capítulo 3, se va a mostrar con más detalle los modelos existentes en la actualidad y realizar una taxonomía de modelos con el fin de mostrar las deficiencias y virtudes de cada uno para poder llevar a cabo, de una forma más acertada, la definición de un nuevo modelo en el que se minimicen las carencias existentes en los modelos actuales.

[12] Conjunto de tecnologías que a través de una cadena de bloques permiten mantener un registro seguro y distribuido de operaciones digitales.

CAPÍTULO 3
Análisis y comparación diferentes metodologías existentes

Una vez expuestas las diversas problemáticas para el análisis forense digital, en el siguiente capítulo se van a enumerar y analizar los diferentes modelos existentes en la actualidad, así como también se realizará una comparación entre ellos para poder tener una visión global del contexto actual y los puntos en los que se debe mejorar.

3.1 Análisis de metodologías existentes

A pesar de que existe una variedad importante respecto a modelos propuestos, en el siguiente apartado se han seleccionado un conjunto de modelos existentes que han destacado en su campo.

3.1.1. Identificación de evidencias en redes IoT basado en la evaluación de amenazas

Uno de los modelos más interesantes y que todavía sirven de referencia es el propuesto por Nikolay Akatyev y Joshua I. James en 2017 [4], cuyo objetivo se basa en el análisis de los dispositivos IoT en el ámbito de las *Smart Homes.*

En este modelo, los autores consideran el IoT como una tecnología heterogénea y lo fundamentan como un modelo UCIoT *(User Centered IoT)*, es decir, el IoT centrado en el usuario y que se centra únicamente en el ámbito interno, de forma que solamente tienen en cuenta los eventos que ocurren dentro de un sistema o red.

La finalidad de este modelo, es la utilización de casos de uso de evaluación de amenazas para el usuario. Este modelo es muy relevante todavía hoy en día, sin embargo, tiene ciertas carencias ya que es un modelo que se centró en los sistemas existentes en el momento de su publicación y no tomó en consideración el modelo para futuros sistemas.

Asimismo, no se introducen términos como integridad o privacidad dado que este modelo se centra únicamente en las casuísticas que puedan afectar al usuario.

3.1.2. Marco forense para identificar entre artefactos locales y sincronizados

El siguiente modelo fue propuesto por Jacques Boucher y Nhien-An Le-Khac en 2018 [6] y basa su metodología en el análisis de la sincronización de datos.

Se trata de un modelo muy relevante ya que introduce los términos del *cloud* y del almacenamiento de datos distribuido.

Este modelo se basa en la investigación forense centrada en la sincronización de datos, es decir, pretende averiguar si los datos han estado en un equipo de forma local o si han sido sincronizados externamente desde otro dispositivo. También se observa la importancia respecto a la cantidad de dispositivos que pueden estar sincronizados y entre los que se debe realizar la investigación.

Sin embargo, presenta varias carencias ya que es muy difícil poder verificar que las evidencias provienen de una fuente sincronizada, al igual que tampoco tiene en consideración conceptos tan importantes como son la integridad y privacidad de las mismas.

3.1.3. Análisis forense de dispositivos IoT y reducción de datos

Daren Quick y Kim-Kwang Raymond, propusieron en 2018 [29] un modelo basado en la reducción de datos.

Los dispositivos IoT generan una cantidad inmensa de datos que en ocasiones puede alcanzar varios *Terabytes* con tan solo el análisis de unos pocos dispositivos.

En esta metodología, se analiza esta problemática y para solucionarlo hacen uso de DRnSI *(Data Reduction by Selective Imaging)*. DRnSI consiste en un proceso de reducción de datos con el fin de que la cantidad de datos a analizar sea únicamente la necesaria, descartando datos irrelevantes.

De esta forma, consiguen junto a un proceso semiautomático utilizando programas como Bulk Extractor[1] y Pajek64[2], el análisis y correlación de datos de forma más eficaz.

Como consecuencia de centrarse únicamente en la reducción de datos en la fase de adquisición, en este modelo no se abordan conceptos como la privacidad de los usuarios ni la integridad de las evidencias.

3.1.4. Testigo digital protegiendo las evidencias digitales mediante el uso de arquitecturas seguras en dispositivos personales

El modelo propuesto en 2016 [23] por Ana Nieto, Rodrigo Román y Javier López, se trata de uno de los modelos con mayor repercusión en la actualidad.

A pesar de tratarse de un modelo publicado hace varios años, ha sido tomado como referencia para muchos modelos posteriores. En este caso, se introduce el concepto de *Digital Witness* (Testigo digital). La finalidad de este concepto, es introducir este elemento en una arquitectura IoT cuya función principal es la de conservar y monitorizar las evidencias.

Este testigo digital tiene como finalidad que todos los elementos IoT se conecten a través de este para, de este modo, controlar todo el flujo de información.

Asimismo, este modelo presenta grandes ventajas respecto a la integridad y privacidad de las evidencias ya que garantiza la integridad al tener todas las evidencias controladas y, por otra parte, también garantiza la privacidad, puesto que se implementa tras la confirmación del usuario.

[1] Programa utilizado por analistas forenses para escanear imágenes de discos y archivos con el fin de obtener información valiosa.

[2] Paquete de programas para realizar un análisis de redes.

Finalmente, a pesar de que se trate de un modelo de referencia, también presenta una importante carencia, ya que el modelo está centrado para dispositivos IoT personales por lo que, como ya indican en su propuesta, no es válido para otros elementos IoT.

3.1.5. Un marco de investigación forense basado en un libro de registro público para IoT

El siguiente modelo publicado en 2018 [13] por Mahmud Hossain, Ragib Hasan y Shams Zawoad, centra su metodología en la integridad, confidencialidad y anonimidad de las evidencias.

Este modelo se centra en mantener y conservar la cadena de custodia de las evidencias, haciendo uso de transacciones mediante la tecnología Blockchain[3] de una forma descentralizada. De este modo, puede garantizar la cadena de custodia y por lo tanto la validez e integridad de las evidencias.

Sin embargo, este modelo se centra únicamente en este aspecto, por lo que la privacidad de los usuarios y la gran cantidad de datos en origen no se tienen en consideración.

3.1.6. Metodología para el análisis forense de IoT centrado en la privacidad

El siguiente modelo propuesto en 2017 [22] por Ana Nieto, Ruben Rios y Javier López, introduce por primera vez la privacidad como una parte fundamental de la metodología.

De este modo, en este modelo definen el PRoFIT*(Privacy-aware IoT-Forensic Model)*, una metodología que se centra en garantizar la privacidad de los usuarios.

El objetivo final de este modelo es el de involucrar a los usuarios para que de forma voluntaria colaboren con la investigación y den su consentimiento ante la obtención de cualquier dato.

De esta forma, esta metodología únicamente está basada en garantizar la privacidad de los usuarios por lo que no se tiene en cuenta aspectos como la recolección de datos e integridad y gestión de las evidencias.

3.1.7. Análisis forense del Internet de las cosas

A pesar de ser uno de los modelos más longevos ya que fue publicado en 2013 [25], el modelo propuesto por Edewere Oriwoh, David Jazani y Gregory Epiphaniou, introdujo conceptos fundamentales que todavía se usan de referencia en la gran mayoría de metodologías.

En su artículo, describen dos tipos de modelos potenciales: *The 1-2-3 Zones of Digital Forensics* y *The Next Best Thing Model*.

El primero, trata de describir dónde buscar las evidencias. Para ello, define tres zonas fundamentales en las que realizar la búsqueda: Una zona interna en la que se encuentra todo el Software, Hardware y las redes; una segunda zona intermedia en la que se encuentran los dispositivos de red que comunican la zona interna con la externa; y, finalmente, una zona externa, en la que el software se encuentra fuera de la red.

Por otra parte, *The Next Best Thing*, se trata de un modelo que se fundamenta en que en muchas ocasiones el dispositivo en el que se deben buscar evidencias puede no estar disponible. De este modo, se propone buscar en dispositivos que tengan relación con él para localizar en ellos las evidencias.

[3] Conjunto de bloques que garantizan la integridad y trazabilidad de la información.

Como punto negativo de este modelo, es que únicamente se basa en la localización y obtención de evidencias, dejando de lado conceptos como la privacidad de los datos e integridad de las evidencias.

3.1.8. Modelo de investigación forense digital del Internet de las cosas

El siguiente modelo [28] propuesto por Sundresan Perumal, Norita Md Norwawi y Valliappan Raman, se basa en la integración del modelo *1-2-3 Zones of Digital Forensics*.

De este modo, este modelo se centra en la obtención de evidencias fundamentado en las 3 zonas diferenciadas. Además, contempla y adapta todas las fases de obtención de evidencias: desde la planificación de la intervención, hasta finalmente el archivo de las evidencias.

Aunque este método se presente como muy completo, cabe destacar que este modelo es puramente teórico y no se han tenido en cuenta aspectos como la privacidad, integridad y recolección heterogénea de datos.

3.1.9. Un modelo forense en IoT basado en la tecnología Blockchain

El modelo publicado en 2021 [18] por Anca Delia, proporciona una visión totalmente diferente a todos los modelos vistos hasta la fecha: el uso de tecnología Blockchain.

Este modelo se fundamenta en garantizar la cadena de custodia de las evidencias y para ello recurre a la tecnología Blockchain de forma descentralizada con el fin de garantizar la cadena de custodia a través de la cadena de bloques.

De este modo, se propone un modelo que garantiza totalmente la integridad de las evidencias y el proceso de su tratamiento, ya que todas las transacciones que ocurran serán registradas en la cadena de bloques. Sin embargo, no se tiene en consideración la privacidad de los usuarios ni la obtención de los datos.

3.2 Clasificación de los modelos

En la actualidad, existe una gran variedad de modelos. Sin embargo, se especializan únicamente en solucionar una parte del problema.

A partir de los diversos modelos analizados que han sido seleccionados debido a su relevancia para este estudio, y sobre los que se han fundamentado multitud de investigaciones, se ha realizado una taxonomía de los mismos como se muestra en la Figura 3.1. De esta forma se puede comparar las carencias y fortalezas de cada uno y poder obtener un panel donde se indican los aspectos considerados en los modelos seleccionados.

Para ello, se han escogido como referencia los siguientes aspectos a tener en cuenta y en los que un modelo debe basarse:

- Integridad de las evidencias.
- Privacidad de los datos obtenidos de los usuarios teniendo en cuenta la legalidad existente.
- Verificabilidad de las evidencias para poder probar su validez.
- Adaptabilidad del modelo para distintos entornos IoT.
- Se tiene en cuenta la heterogeneidad de datos de distintas fuentes.
- Utilizan algoritmos de inteligencia artificial.
- Hacen uso de la tecnología Blockchain.

Taxonomía de modelos

Modelo	Integridad evidencias	Privacidad	Verificabilidad	Adaptabilidad	Heterogeneidad de datos	IA	BlockChain
Nikolay Akatyev (2017)					✓		
Jacques Boucher (2018)				✓	✓		
Daren Quick (2018)				✓	✓	✓	
Ana Nieto (2016)	✓	✓	✓				
Mahmud Hossain (2018)	✓		✓	✓			✓
Ana Nieto (2017)		✓		✓			
Edewere Oriwoh (2013)				✓	✓		
Sundresan Perumal (2015)				✓			
Anca Delia (2021)		✓	✓	✓			✓

Figura 3.1: Taxonomía de los modelos. Fuente propia.

Como se puede observar, los modelos existentes presentan carencias, dejando sin cubrir aspectos importantes en el tratamiento de las evidencias.

A partir de los valores obtenidos en la taxonomía de modelos, en la Figura 3.2 se ha realizado una gráfica de pesos en la que se puede observar de una forma más visual el campo mayoritario de investigación en el que se centran los principales modelos existentes.

Figura 3.2: Campos de investigación de los modelos actuales. Fuente propia.

Otra clasificación relevante de las metodologías, es en función de su campo mayoritario de estudio. En la Figura 3.3, se muestra una clasificación de las metodologías en función de su línea de investigación.

Clasificación de modelos según su campo de investigación			
Basado en la adquisición	Basado en la privacidad	Basado en la integridad/CoC	Basado en el análisis
Daren Quick (2018) Jacques Boucher (2018)	Ana Nieto (2017)	Ana Nieto (2016) Mahmud Hassain (2018) Anca Delia (2021)	Nikolay Akatyev (2017) Edewere Oriwoh (2013) Sudresan Perumal (2015)

Figura 3.3: Campos de investigación de los modelos actuales. Fuente propia.

De esta forma, tras el análisis de las modelos seleccionados, se puede concluir que la gran mayoría de los modelos existentes basan su campo de estudio en abordar una única problemática. Además, la integridad de los datos y la privacidad de los usuarios rara vez se tienen en consideración, por lo que el futuro modelo propuesto, debe centrarte en intentar reducir todas estas carencias.

CAPÍTULO 4
Análisis y definición de una nueva metodología

Una vez analizados y contrastados los diferentes modelos más relevantes en la actualidad, en el siguiente capítulo se propone definir una nueva metodología que solvente la gran mayoría de las carencias detectadas, haciendo uso de tecnologías de vanguardia como son algoritmos de inteligencia artificial de detección de patrones y cadenas de bloques de Blockchain.

Además, también se tomarán modelos ya existentes como referencia y cuya aportación es muy relevante para este proyecto.

El modelo propuesto está constituido de la siguiente forma:

- Se propone una metodología a dos niveles haciendo uso para ello de algoritmos de clasificación de datos y detección de anomalías.
- Se analiza la legislación actual para garantizar la privacidad y protección de datos de los usuarios afectados.
- Se hace uso de la tecnología Blockchain para garantizar la cadena de custodia, la integridad y la verificabilidad de las evidencias.
- La metodología debe poder adaptarse a la gran variedad de dispositivos IoT.
- Se considera la escalabilidad de los datos y la heterogeneidad de dispositivos.

Con todo ello, el modelo propuesto se va a explicar a través del desarrollo e interacción de las tecnologías escogidas para las distintas etapas clásicas del análisis forense mostradas en la Figura 4.1.

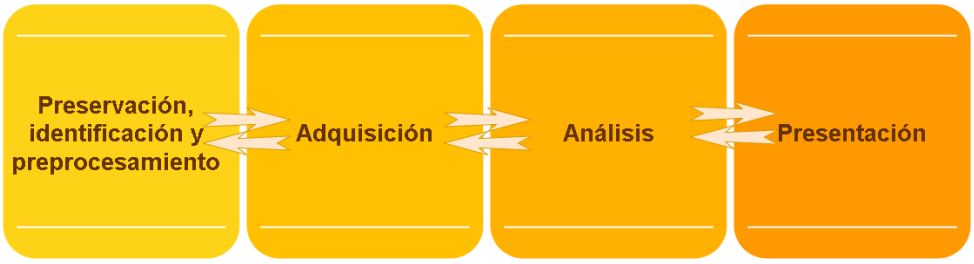

Figura 4.1: Etapas análisis forense digital clásico. Fuente propia.

4.1 Etapas del análisis forense según la nueva metodología

4.1.1. Etapa de preservación, identificación y preprocesamiento

El objetivo fundamental en esta etapa es la identificación de las fuentes de evidencias. En un entorno IoT, esta identificación puede resultar una tarea complicada dada la heterogeneidad de los dispositivos que lo forman, junto al gran volumen de datos que pueden proporcionar. Es por ello que nuestra propuesta aborde esta identificación a dos niveles.

El primer nivel, tiene como objetivo la identificación de patrones y/o comportamientos anómalos, de modo que podamos determinar los agentes involucrados en el incidente. El segundo nivel, se centrará en los agentes identificados, con el objetivo de obtener evidencias relevantes al incidente. De este modo se pretende limitar la adquisición de datos a aquellos que presenten relevancia en el incidente. Este segundo nivel se implementará en la etapa de adquisición de datos, por lo que se describirá en ella.

Nivel I

En este primer nivel se propone el uso de la inteligencia artificial (IA). Su aplicación, proporcionará efectividad, precisión y reducción de costes, si lo comparamos con la realización de esta tarea manualmente.

El uso de algoritmos de inteligencia artificial ha demostrado ser de gran eficacia en el ámbito forense-informático [16] aunque en muy pocas ocasiones se haya incluido en las metodologías actuales. Es esta eficacia la que ha motivado que este tipo de algoritmos sea un pilar fundamental para la metodología propuesta.

Dentro de la inteligencia artificial, podemos diferenciar el *Machine Learning*[1] y el *Deep Learning* [30].

En el primero, se precisa de aportar un conjunto de datos como entrenamiento para que a partir de este, el modelo sea capaz de reconocer ciertos patrones y realizar las predicciones, mientras que el segundo, siendo un subconjunto del *Machine Learning*, el algoritmo es capaz de realizar sus propias conclusiones sin necesidad de aportar datos para entrenarlo.

Nosotros nos centraremos en el primero, en los algoritmos de *Machine Learning*, entre los cuales pueden diferenciarse cuatro grandes grupos: supervisados, semi-supervisados, aprendizaje por refuerzo y los no supervisados [17].

- **Supervisados:** Los modelos supervisados, utilizan probabilidades de datos proporcionados, que están etiquetados, esto es, datos de los que ya se sabe a qué grupo, valor o categoría pertenecen para obtener probabilidades de los nuevos eventos.

  Algunos de los modelos incluidos en esta familia son los árboles de decisión, algoritmos de regresión logística, los algoritmos SVM (Support Vector Machines), Naive Bayes, Redes Neuronales, etc

- **Semi-supervisados:** Los modelos semi-supervisados, hacen uso en el entrenamiento tanto de conjuntos de datos etiquetados como no etiquetados. Estos modelos basan su funcionamiento en entrenar un algoritmo supervisado con una pequeña parte del conjunto de datos, que ha sido etiquetado manualmente, para posteriormente proporcionar el resultado al algoritmo no supervisado.

[1] Algoritmo perteneciente a la familia de la inteligencia artificial que se basa en el aprendizaje automático.

- **Aprendizaje por refuerzo:** Se trata de un método en el que basa su funcionalidad a base de prueba y error. En este caso trata de penalizar los comportamientos que no son esperados y de beneficiar a los que son correctos.

  De esta forma el algoritmo va aprendiendo a base de prueba-error. El inconveniente de este método, es que al inicio del análisis, sus predicciones son muy erróneas por lo que debe considerarse como un método de aplicación a largo plazo.

- **No supervisado:** Los modelos no supervisados, realizan predicciones desde conjuntos de datos no etiquetados y que posteriormente lo aplican a los nuevos datos.

  En esta familia podemos encontrar algoritmos como K-means, DBSCAN, One-Class SVM, Isolation Forests, etc.

La importancia del uso de *Machine Learning* en el ámbito de la ciberseguridad, es que permite realizar reconocimiento de patrones y detección de anomalías, dos campos que son de gran importancia para este proyecto.

De este modo, tras el análisis de distintos modelos de inteligencia artificial tanto supervisados como no supervisados, se ha escogido un modelo llamado *Isolation Forests* debido a que presenta las siguientes características:

- Algoritmo de *Machine Learning* no supervisado. Debido a la heterogeneidad de los datos obtenidos, el modelo debe ser no supervisado ya que las muestras no van a poder etiquetarse en la mayoría de casos.
- Se trata de un algoritmo de detección de anomalías.
- Funciona muy bien para grandes cantidades de datos. Su debilidad es que haya pocos datos de entrada.

Este tipo de algoritmo se ha escogido por una finalidad doble: trabaja muy bien con grandes cantidades de datos y, al no tratarse de un modelo supervisado, podemos trabajar con conjuntos de datos que no estén etiquetados o que lo estén mínimamente.

Además, se ha contrastado [31] la viabilidad de este algoritmo respecto a otros algoritmos como Árboles de decisión, *Random forest*, *Bagging classifier* y otros más, ya que tanto su precisión de la predicción como sus tiempos de ejecución son óptimos.

De este modo, se mejoran dos de las principales carencias de los modelos actuales: el manejo de grandes cantidades de datos y su gran heterogeneidad.

4.1.2. Etapa de adquisición de evidencias

En este punto, se presenta el otro nivel propuesto en la metodología. Una vez identificadas las posibles fuentes de evidencias, proponemos aplicar nuevamente inteligencia artificial, aunque en este caso el algoritmo tiene diferentes objetivos, lo que requerirá distintas características, como es el que sea parametrizable [33] para poder ajustar el modelo a las necesidades del sistema IoT concreto. El objetivo de ello es detectar las anomalías, susceptibles de ser evidencias que justifiquen la hipótesis del incidente o la desmientan, o si por el contrario no presentan relevancia y que sean descartadas.

De este modo, una vez analizados los datos en la etapa anterior, en este caso se va a disponer de una menor cantidad de datos que ya van a poder estar etiquetados. Para este punto, se va a poder hacer uso de un algoritmo de *Machine Learning* supervisado basado en árboles de decisión dado que ha sido comprobado que aporta buenos resultados respecto a la precisión [21] y que presentan las siguientes ventajas:

- Son algoritmos de clasificación cuya decisión es explicable.
- Facilmente entrenables.

Asimismo, se deben tener en cuenta también si la adquisición se va a realizar sobre dispositivos en ejecución, es decir, una adquisición en vivo, o si por el contrario las evidencias se extraerán de dispositivos apagados, es decir, una adquisición *post-mortem*[2]. De igual modo, también se debe especificar si la granularidad de los datos va a ser fina o gruesa.

- En caso de ser un análisis *post-mortem*, el proceso va a ser mucho más sencillo puesto que se va a poder analizar más en profundidad los datos.
- Un análisis en vivo, presentará las ventajas de poder acceder a datos en memoria, en transito o todavía no cifrados. Sin embargo, la extracción de estos es mucho más delicada y complicada.

Otro aspecto fundamental a tener en cuenta en esta fase es la privacidad y protección de datos de los usuarios, así como la legalidad aplicable y el *Cloud*.

Para ello se deberán tener en consideración todos los aspectos referentes a la Ley Orgánica de Protección de datos del país en el que se encuentren estos.

En el caso de España, se va a regir por la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales [5] y cuya aplicación será de obligatoriedad para garantizar la validez y legalidad de las evidencias.

En este caso se tiene que tener en cuenta aspectos como el deber de confidencialidad, la transparencia e información al afectado y el tratamiento de datos entre otros.

De esta forma, en la investigación forense será imprescindible tener en cuenta los siguientes aspectos:

- Tener el consentimiento expreso de los propietarios de los datos.
- Garantizar, controlar e implantar las medidas de seguridad necesarias para garantizar la protección de los datos obtenidos.
- Ser transparente en todos los pasos de la investigación respecto a los datos obtenidos.

Adicionalmente, en el ámbito europeo, se establece el Reglamento UE (Unión Europea) 2016/679 del parlamento europeo y del consejo denominado RGPD [27] (Reglamento general de protección de datos) el cual dicta unas normas básicas que se regirán en todos los países pertenecientes en la UE y que deberá tenerse en cuenta para investigaciones en esta zona geográfica.

Otro aspecto relevante a tener en cuenta son los *Data Leaks*[3] que se hayan podido producir, al igual que su mitigación.

Este aspecto no es sencillo de abordar puesto que no es fácil poder localizar las fugas de información en estos dispositivos debido a la dificultad de implantación de herramientas preventivas. [8]

[2] Análisis realizado sin arrancar el sistema.
[3] Exfiltración de datos. Se trata de la fuga de datos de un sistema a través de una ataque informático.

Como posible solución, se propone la consulta en bases de datos de IoC *(Index of Compromise*[4]*)* para observar si las comunicaciones de estos dispositivos se realizan contra alguna fuente que ya está considerada maliciosa. De este modo, es posible localizar las comunicaciones contra servidores maliciosos y en los que comúnmente se realizan este tipo de acciones de exfiltración.

Finalmente, también se deberá garantizar la ISO/IEC 27037 [24], directrices para la identificación, recopilación, adquisición y preservación de evidencia digital, la cual proporciona una serie de pautas para el tratamiento de las evidencias digitales. Algunas de las pautas sobre las que se fundamenta la norma son las siguientes:

- Las evidencias deben ser adquiridas de la forma menos intrusiva posible con el fin de preservar su integridad.
- Todo el proceso de tratamiento de las evidencias debe ser auditable, es decir, se debe poder seguir todo el proceso por el que ha transitado la evidencia.
- Las técnicas aplicadas deben ser reproducibles, argumentables y verificables.
- Todo el proceso de obtención debe poder ser defendible.

4.1.3. Etapa de análisis y evaluación de evidencias

En esta etapa se va a tener que garantizar la integridad y verificabilidad de las evidencias. Para ello y también para poder garantizar la cadena de custodia, se propone el uso de la tecnología Blockchain.

La tecnología Blockchain puede definirse como una estructura matemática para almacenar datos de forma que crea un registro inmutable de cada una de las transacciones. De esta forma, esta tecnología puede definirse como un libro digital público, compartido por varios usuarios, que garantiza que cada una de las transacciones (o bloques) quede registrado de forma inequívoca.

El uso de Blockchain se va a realizar de forma descentralizado por lo que cualquier usuario podrá verificar la validez e integridad de la evidencia en cualquier etapa y desde cualquier lugar. Un esquema de ejemplo del uso de esta tecnología se muestra en la Figura 4.2.

[4]Índice que indica que un recurso es malicioso.

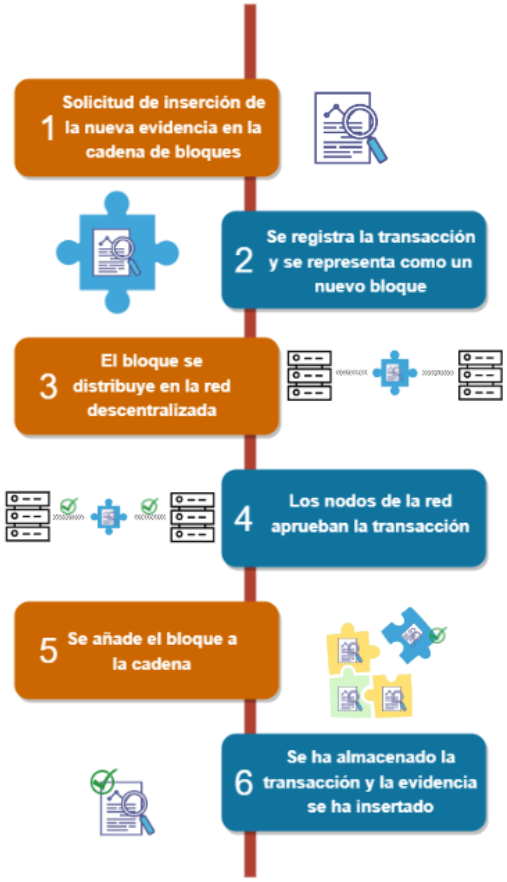

Figura 4.2: Etapas análisis Blockchain. Fuente propia.

Para este caso, se va a tomar como referencia el modelo basado en Blockchain de Anca Delia [18] debido a que proporciona una visión muy clara respecto al uso de esta tecnología para garantizar la CoC.

Para ello, en esta parte intervienen cuatro activos de gran relevancia:

- La figura del investigador y las autoridades.
- La figura del núcleo de Blockchain.
- La figura de los elementos procesadores de transacciones y *logs*.
- Generación de una BBDD externa de seguridad.

La figura del investigador y las autoridades

Son los elementos que precisan de verificar la integridad y validez de las evidencias. Estos usuarios necesitan acceder a las evidencias para investigarlas y por lo tanto deben cerciorarse de que son válidas.

La figura del núcleo de Blockchain

Se trata del centro de la cadena de bloques y cuya función es almacenar todas las transacciones producidas en la cadena, de forma distribuida, con el fin de que en el caso de que caiga un nodo, toda la información de las evidencias se encuentra disponible.

La figura de los elementos procesadores de transacciones y logs

Son las encargadas de generar los logs y las claves SHA256[5] de las evidencias para transferirlas a la red de Blockchain.

Base de datos Blockchain

Todas las transacciones producidas en la cadena Blockchain se van a almacenar en una BBDD (Base de datos) externa con el fin de poder verificar en el caso de corrupción de datos que no han sido modificados.

Normativa aplicable

El análisis también se deberá ceñir a las siguientes normas existentes en la actualidad:

- Norma UNE71506:2013 [2], Metodología para el análisis forense de las evidencias electrónicas por AENOR. (Asociación Española de Normalización y Certificación)

  Esta norma define el proceso del análisis forense en todas sus etapas y la gestión de las evidencias en las distintas fases.

- Norma UNE71505 [3], Sistema de Gestión de Evidencias Electrónicas por AENOR.

  Esta norma pretende garantizar la disponibilidad de las evidencias y su correcta gestión. Para ello esta nueva metodología debe ser estrictamente respetuosa con esta norma.

4.1.4. Etapa de presentación de evidencias

En la etapa de presentación, se debe poder demostrar toda la cadena de custodia que han seguido las evidencias al igual que su presentación ante una corte.

Para ello los investigadores y agentes deberán poder consultar cualquier tipo de evidencias en la cadena de bloques para poder garantizar su validez.

Además, para poder representar y redactar el informe del análisis se deberán seguir la siguientes norma:

- Norma Española UNE197001:2019 [1], Criterios generales para la elaboración de informes periciales.

  En la etapa de presentación de evidencias, esta norma es de gran relevancia y debe tenerse en consideración. Esta norma incluye los aspectos generales para la elaboración de informes periciales y por lo tanto que sean aceptados por la administración correspondiente.

4.2 Arquitectura de la metodología propuesta

Tras definir la nueva metodología propuesta, se puede observar cómo el uso de tecnologías emergentes como son el Blockchain y la inteligencia artificial pueden resultar de gran ayuda para el análisis forense en IoT.

[5] Algoritmo criptográfico basado en un conjunto de funciones *hash*.

Para ello, se ha representado en la Figura 4.3, el esquema del modelo definido en el punto anterior donde se pueden observar de una forma resumida todos los aspectos a tener en cuenta.

En él, se definen los 3 posibles tipos de fuentes de evidencias y el posterior tratamiento de las evidencias a través del consentimiento y conocimiento de los usuarios afectados.

Además, posteriormente se incluye el doble análisis de las evidencias encontradas:

- Un primer análisis para identificar las posibles fuentes de evidencias a través de un algoritmo de *Machine Learning* no supervisado.
- Un segundo análisis de clasificación de patrones a través de un algoritmo de *Machine Learning* supervisado.

Todos estos datos obtenidos se procesarán a través de un nuevo elemento denominado *Procesador de logs* y que irá trasmitiendo todos los nuevos *logs* y *hashes* generados a la cadena de nodos descentralizada de Blockchain.

Asimismo, se considera el uso de una BBDD de seguridad que contenga una copia aislada de todas las transacciones realizadas con el único fin de poder verificar de nuevo la validez de las evidencias en el caso de corrupción de la cadena de custodia.

Como parte final el esquema, se muestra el acceso tanto de los peritos informáticos como de los agentes a la cadena de custodia con el fin de poder verificar la información de las evidencias.

Además, en la Figura 4.4, se observa el diagrama de flujo de la metodología definida, en el cual se puede observar el flujo de todo el proceso de obtención de evidencias, desde la fase de preprocesamiento hasta la última fase de presentación.

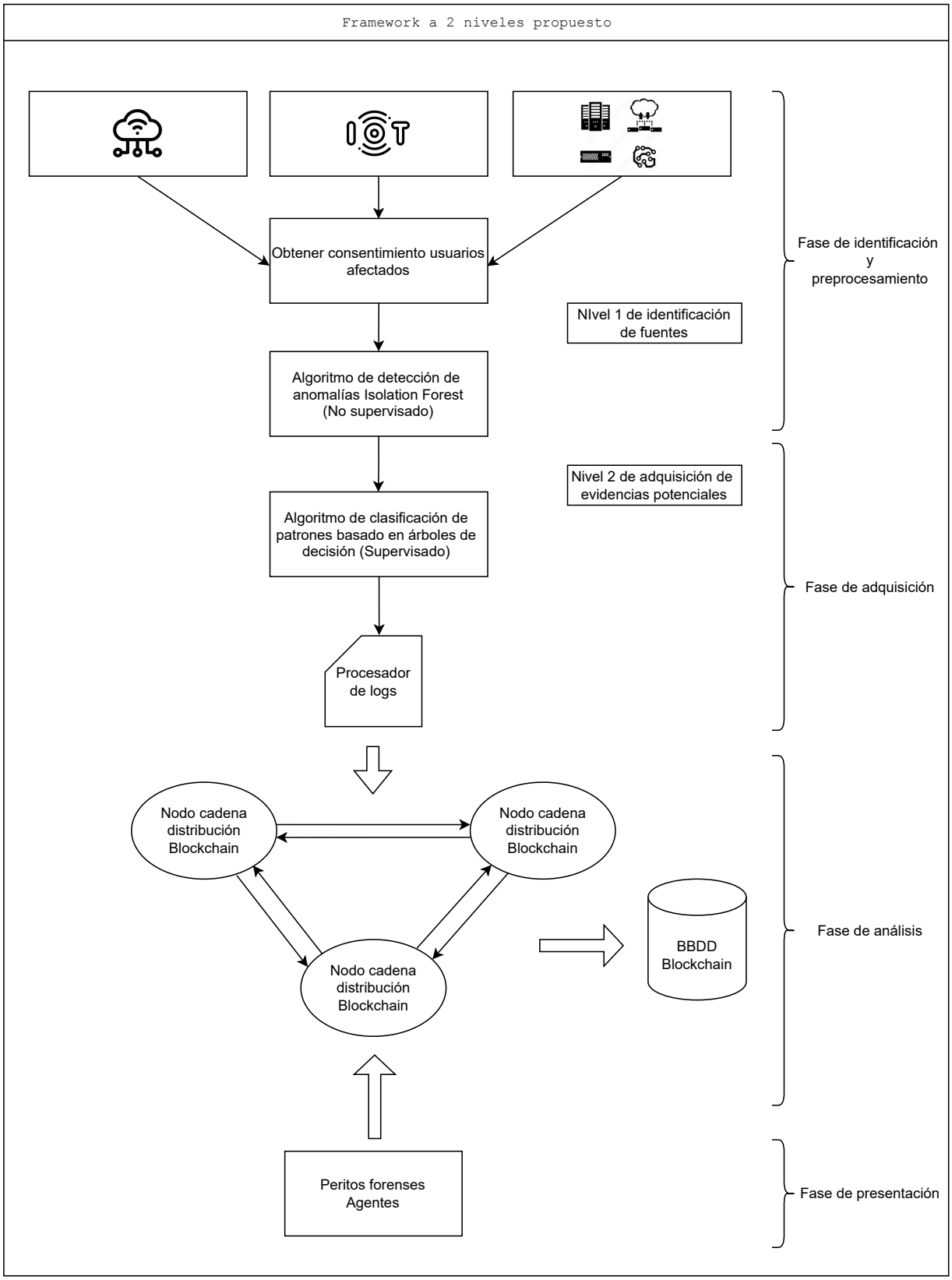

Figura 4.3: Diagrama del modelo propuesto. Fuente propia.

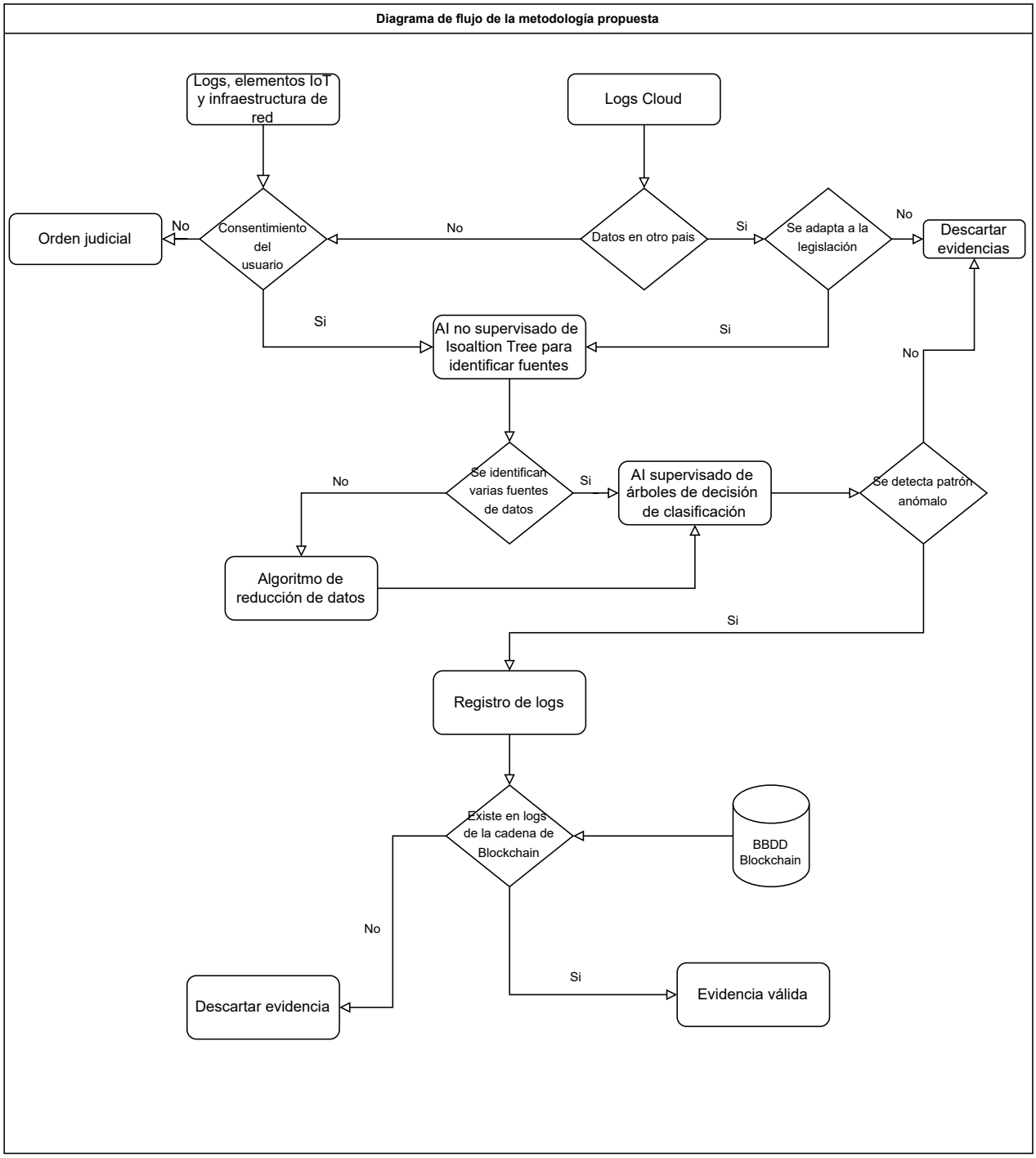

Figura 4.4: Diagrama de flujo del modelo propuesto. Fuente propia.

CAPÍTULO 5
Conclusiones

Vivimos en una sociedad cada vez más digitalizada en el que el uso de dispositivos inteligentes ya forma parte esencial de nuestros quehaceres diarios. Parte de estos dispositivos, forman parte de la familia del IoT y su uso ha aumentado exponencialmente en los últimos años.

Por otra parte, los ataques cibernéticos también han aumentado de forma considerable, afectando en gran medida a esta tecnología emergente. Asimismo, surge la necesidad de describir una metodología para la realización de un análisis forense para poder obtener evidencias con total seguridad y certeza de que puedan verificarse ante cualquier agente y por lo tanto ser demostrables ante la autoridad competente.

De este modo, en un contexto de una falta de metodología estándar que ampare el análisis forense-informático en el ámbito del IoT, se presenta una metodología a dos niveles que trata de solventar las siguientes deficiencias existentes:

- Gran heterogeneidad de datos.
- Gran cantidad de datos generados.
- Adaptabilidad a diferentes dispositivos IoT.
- Preservar la privacidad de los usuarios.
- Garantizar la integridad, verificabilidad y cadena de custodia de las evidencias.

En el marco propuesto, se hace uso de la tecnología de la inteligencia artificial en su modalidad de detección de patrones y anomalías, con el fin de subsanar dos de los grandes problemas mencionados: la heterogeneidad y la gran cantidad de datos generados.

Además, a través del uso de otra tecnología emergente como es el Blockchain, se pretende garantizar en todo momento que las evidencias obtenidas tengan validez ante cualquier autoridad competente.

De este modo, este proyecto ha tratado de solventar todas estas problemáticas y enmarcarlas en una nueva metodología para la realización del análisis forense focalizada en el análisis IoT.

Sin embargo, quedan todavía aspectos relevantes a tener en cuanta que, debido a su magnitud, no son fáciles de solucionar. Por ejemplo, la localización de datos en la nube, dificulta ampliamente las investigaciones ya que dependerá su obtención tanto de la legislación del país en el que se encuentre, si se encuentran de forma distribuida entre diferentes CPDs (Centro de Procesamiento de Datos) y también si el propietario del CPD autoriza el acceso a los datos.

Estos aspectos, junto a una legislación cambiante respecto a la privacidad de los datos dependiendo de la zona geográfica en la que se encuentren, son dos de los grandes retos que deberán ser afrontados en futuros marcos normativos.

CAPÍTULO 6
Trabajos futuros

Debido a que este proyecto se realiza la definición de una nueva metodología para el análisis forense, como trabajos futuros se propone la implementación de la misma para la realización de una investigación forense en un entorno IoT comprometido-controlado para poder verificar y validar su efectividad.

Esta implementación no es una tarea sencilla puesto que se requiere tener un laboratorio de pruebas con distintos dispositivos IoT, entre los que se encuentre alguno infectado con el fin de poder tener un conjunto de datos que poder proporcionar a los algoritmos de inteligencia artificial para poder realizar sus predicciones y poder evaluar su efectividad. Además, la implantación de la cadena de Blockchain, también deberá realizarse de forma minuciosa con el fin de poder demostrar que todas las transacciones se almacenan de forma correcta y que las evidencias sean verificables y accesibles.

Bibliografía

[1] AENOR. Criterios generales para la elaboración de informes periciales. https://www.une.org/encuentra-tu-norma/busca-tu-norma/norma?c=N0062378. [Consulta: 29 de Mayo de 2022].

[2] AENOR. Tecnologías de la información (ti). metodología para el análisis forense de las evidencias electrónicas. https://www.une.org/encuentra-tu-norma/busca-tu-norma/norma/?c=N0051414. [Consulta: 29 de Mayo de 2022].

[3] AENOR. Tecnologías de la información (ti). sistema de gestión de evidencias electrónicas (sgee). https://www.une.org/encuentra-tu-norma/busca-tu-norma/norma/?c=N0051411. [Consulta: 29 de Mayo de 2022].

[4] Nikolay Akatyev and Joshua I. James. Evidence identification in iot networks based on threat assessment. *Future Generation Computer Systems*, 93:814–821, 2019.

[5] BOE. Ley orgánica 3/2018, de 5 de diciembre, de protección de datos personales y garantía de los derechos digitales. https://www.boe.es/buscar/pdf/2018/BOE-A-2018-16673-consolidado.pdf. [Consulta: 29 de Mayo de 2022].

[6] Jacques Boucher and Nhien-An Le-Khac. Forensic framework to identify local vs synced artefacts. *Digital Investigation*, 24:S68–S75, 2018.

[7] Brainsins. La gamificación entra en el hype cycle de gartner. https://www.brainsins.com/es/blog/gamificacion-hype-gartner/5897/. [Consulta: 23 de Abril de 2022].

[8] Dmitry Raidman. Detecting and mitigating iot breaches require an "inside-out" approach to security. https://www.infosecurity-magazine.com/opinions/detecting-mitigating-iot-breaches/. [Consulta: 11 de Junio de 2022].

[9] Gregory Epiphaniou Paul Sant Edewede Oriwoh, David Jazani. Internet of things forensics: Challenges and approaches. pages 609 –615, 2013.

[10] EuropaPress. Los ataques a dispositivos del iot se duplicaron en el primer semestre de 2021. https://cutt.ly/XGAg79B. [Consulta: 05 de Febrero de 2022].

[11] FBI. Forensic science communications. https://archives.fbi.gov/archives/about-us/lab/forensic-science-communications/fsc/april2000/swgde.htm#:~:text=The%20International%20Organization%20on%20Computer%20Evidence%20(IOCE)%20was%20established%20in,other%20computer%2Drelated%20forensic%20issues. [Consulta: 23 de Abril de 2022].

[12] AhmedAlenezi Madini O.Alassafi Gary B.Wills Hany F.Atlam, EzzEl-Din Hemdan. Internet of things forensics: A review. pages 13 – 15, 2020.

[13] Mahmud Hossain, Ragib Hasan, and Shams Zawoad. Probe-iot: A public digital ledger based forensic investigation framework for iot. In *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pages 1–2, 2018.

[14] Incibe. La importancia de la seguridad en iot. principales amenazas. https://www.incibe-cert.es/blog/importancia-seguridad-iot-principales-amenazas/. [Consulta: 05 de Febrero de 2022].

[15] Informática Forense. Historia informática forense. http://informaticaforense1.blogspot.com/2013/11/historia-informatica-forense.html. [Consulta: 23 de Abril de 2022].

[16] Aaron Jarrett and Kim-Kwang Raymond Choo. The impact of automation and artificial intelligence on digital forensics. *WIREs Forensic Science*, 3(6):e1418, 2021.

[17] Jose Antonio Sanchez. ¿cómo aprenden las máquinas? machine learning y sus diferentes tipos. https://datos.gob.es/es/blog/como-aprenden-las-maquinas-machine-learning-y-sus-diferentes-tipos. [Consulta: 10 de Junio de 2022].

[18] Anca Delia Jurcut. *BLOFF: A Blockchain-Based Forensic Model in IoT*. IGI Global, 2021.

[19] Livinginternet. The internet toaster. https://www.livinginternet.com/i/ia_myths_toast.htm. [Consulta: 23 de Abril de 2022].

[20] Spyridon Panagiotakis Evangelos Markakis Maria Stoyanova, Yannis Nikoloudakis. *A Survey on the Internet of Things (IoT) Forensics: Challenges, Approaches, and Open Issues*, pages 9–12. 04 2022.

[21] Muhammad Shoaib Mazhar, Yasir Saleem, Ahmad Almogren, Jehangir Arshad, Mujtaba Hussain Jaffery, Ateeq Ur Rehman, Muhammad Shafiq, and Habib Hamam. Forensic analysis on internet of things (iot) device using machine-to-machine (m2m) framework. *Electronics*, 11(7), 2022.

[22] Ana Nieto, Ruben Rios, and Javier Lopez. A methodology for privacy-aware iot-forensics. In *2017 IEEE Trustcom/BigDataSE/ICESS*, pages 626–633, 2017.

[23] Ana Nieto, Rodrigo Roman, and Javier Lopez. Digital witness: Safeguarding digital evidence by using secure architectures in personal devices. *IEEE Network*, 30(6):34–41, 2016.

[24] Organización Internacional de Normalización. Iso/iec 27037:2012. https://ciberseguridad.com/normativa/espana/iso-iec-27037-evidencia-digital/. [Consulta: 29 de Mayo de 2022].

[25] Edewede Oriwoh, David Jazani, Gregory Epiphaniou, and Paul Sant. Internet of things forensics: Challenges and approaches. In *9th IEEE International Conference on Collaborative Computing: Networking, Applications and Worksharing*, pages 608–615, 2013.

[26] Paloma Recuero de los Santos. Breve historia de internet de las cosas (iot). https://empresas.blogthinkbig.com/breve-historia-de-internet-de-las-cosas-iot/. [Consulta: 23 de Abril de 2022].

[27] Parlamento Europeo. Reglamento (ue) 2016/679 del parlamento europeo y del consejo de 27 de abril de 2016. https://eur-lex.europa.eu/legal-content/ES/TXT/?qid=1532348683434&uri=CELEX%3A02016R0679-20160504. [Consulta: 29 de Mayo de 2022].

[28] Sundresan Perumal, Norita Md Norwawi, and Valliappan Raman. Internet of things(iot) digital forensic investigation model: Top-down forensic approach methodology. In *2015 Fifth International Conference on Digital Information Processing and Communications (ICDIPC)*, pages 19–23, 2015.

[29] Darren Quick and Kim-Kwang Raymond Choo. Iot device forensics and data reduction. *IEEE Access*, 6:47566–47574, 2018.

[30] Rodrigo Alonso. Ia, machine learning y deep learning, ¿cuál es la diferencia? https://hardzone.es/tutoriales/rendimiento/diferencias-ia-deep-machine-learning/. [Consulta: 10 de Junio de 2022].

[31] Charulatha S., Neela Madheswari, Shanthi K., and Chamundeswari Arumugam. *Crucial Role of Data Analytics in the Prevention and Detection of Cyber Security Attacks*, pages 67–80. 01 2021.

[32] Salesforce. Cloud computing: Aplicaciones en un solo lugar. https://www.salesforce.com/mx/cloud-computing/. [Consulta: 23 de Abril de 2022].

[33] Biasiotti Maria Angela Solanke Abiodun. Digital forensics ai: Evaluating, standardizing and optimizing digital evidence mining techniques. *Kanstliche Intelligenz*, 2022.

[34] SonicWall. 2022 sonicwall cyber threat report. https://www.sonicwall.com/2022-cyber-threat-report/. [Consulta: 20 de Mayo de 2022].

[35] Statista. Dispositivos conectados (internet de las cosas) a nivel mundial de 2019 a 2030. https://es.statista.com/estadisticas/517654/. [Consulta: 23 de Abril de 2022].

[36] U.S. Department of Justice. Computer analysis and response team (cart): The microcomputer as evidence. https://www.ojp.gov/ncjrs/virtual-library/abstracts/computer-analysis-and-response-team-cart-microcomputer-evidence. [Consulta: 23 de Abril de 2022].

APÉNDICE A
Acrónimos y términos de interés

- **Cloud:** Tecnología de alojamiento de datos en servicios externos y que permite el acceso remoto.
- **Evidencia:** Conjunto de información que indican si una acción es cierta o no.
- **Firmware:** Programa básico de dispositivos electrónicos.
- **Framework:** Marco de trabajo.
- **IoT:** Internet of Things (Internet de las cosas).
- **NIST:** Instituto Nacional de Estándares y Tecnología.
- **Software:** Conjunto de programas cuya finalidad es la automatización y realización de tareas.
- **WIFI:** Wireless Fidelity.
- **Malware:** Programa malicioso.
- **Log:** Registro de una acción en un sistema.
- **UCIoT:** User Centered IoT (IoT Centrado en el usuario).
- **PRoFIT:** Privacy-aware IoT-Forensic. Análisis forense en IoT centrado en la privacidad.
- **Digital Witness:** Testigo Digital.
- **DRnSI:** Data Reduction by Selective Imaging (Reducción de la información a través de selección de imágenes).
- **UE:** Unión Europea.
- **RGPD:** Reglamento general de protección de datos.
- **AENOR:** Asociación Española de Normalización y Certificación.
- **BBDD:** Base de Datos.
- **CPD:** Centro de procesamiento de datos.
- **IoC:** Índice de Compromiso.

APÉNDICE B
Planificación del trabajo

Este apéndice muestra cómo ha sido la planificación llevada a cabo durante el proyecto, con el fin de distribuir de forma eficiente las horas asignadas al trabajo.

B.1 Planificación inicial

La planificación inicial consta de siete puntos principales que agrupan las tareas más importantes y en el que se planifica un inicio del proyecto desde Diciembre de 2021 hasta Abril de 2022.

El primer punto de la planificación aúna 50 horas de dedicación exclusiva a la definición del proyecto, planificación de los objetivos, revisión de artículos iniciales para definir el proyecto y revisión de vídeos de conferencias impartidas por profesionales del sector.

El segundo punto, cuya finalidad es exclusivamente la búsqueda de artículos científicos, se ha planificado un total de 60h con el fin de recabar la mayor información posible de diferentes fuentes para poder enriquecer la investigación final.

El tercer punto se trata de uno de los más importantes por lo que es uno de los que se ha empleado un mayor esfuerzo de dedicación. En total, se han destinado 75 horas de investigación sobre los diferentes modelos existentes, las carencias del análisis forense en dispositivos IoT y en general, recabar y contrastar la información de los diferentes artículos escogidos.

Por otra parte, el apartado cuarto se han empleado 30 horas dedicadas al estudio y comparación de los diferentes modelos de análisis forense de dispositivos IoT para observar y comparar las carencias de unos y los beneficios de otros.

Por lo que respecta al punto quinto, se trata de uno de los más relevantes ya que tras estudiar todo lo anterior se intenta construir una nueva metodología de análisis IoT. En este punto se han invertido un total de 60 horas.

Respecto al punto sexto, se ha incluido un total de 137 horas para redactar y elaborar la memoria del trabajo final de máster.

Finalmente, el punto número 7 con 38 horas se ha dedicado a preparar la presentación y exposición oral del proyecto.

A continuación se puede observar tanto la Tabla B.1 como la Figura B.1 con la planificación inicial detallada.

Planificación horaria por tareas			
Nº	**Tarea**	**Tiempo (h.)**	**Dependencia**
1	**Planificación**	**50 horas**	
1.1	Definición del proyecto con la tutora	2 horas	
1.2	Revisión artículos iniciales	10 horas	
1.3	Revisión vídeos de expertos	10 horas	
1.4	Análisis del proyecto y búsqueda de información	24 horas	
1.5	Redactar propuesta del proyecto	4 horas	
2	**Búsqueda de artículos**	**60 horas**	
2.1	Búsqueda de artículos divulgativos	20 horas	
2.2	Estudio y lectura de los artículos	40 horas	2.1
3	**Investigación**	**75 horas**	**2**
3.1	Extraer aspectos importantes de los artículos	35 horas	
3.2	Crear brainstorming de los aspectos relevantes	20 horas	
3.3	Crear un boceto inicial de las ideas obtenidas	20 horas	
4	**Comparación modelos**	**30 horas**	**3**
4.1	Comparación de los modelos existentes	16 horas	
4.2	Identificar pros y contras de los modelos	14 horas	4.1
5	**Desarrollo metodología**	**60 horas**	**4**
5.1	Elaborar nueva metodología	40 horas	
5.2	Comparar la metodología con las existentes	17 horas	5.1
5.3	Elaboración de gráficos y figuras	3 horas	5.1
6	**Redacción memoria**	**137 horas**	**5**
6.1	Redactar la memoria	100 horas	
6.2	Creación de gráficos	37 horas	6.1
7	**Preparar presentación**	**38 horas**	**6**
7.1	Preparar diapositivas presentación ante tribunal	18 horas	
7.2	Preparar exposición oral	20 horas	7.1

Tabla B.1: Desglose de tareas de la planificación horaria inicial.

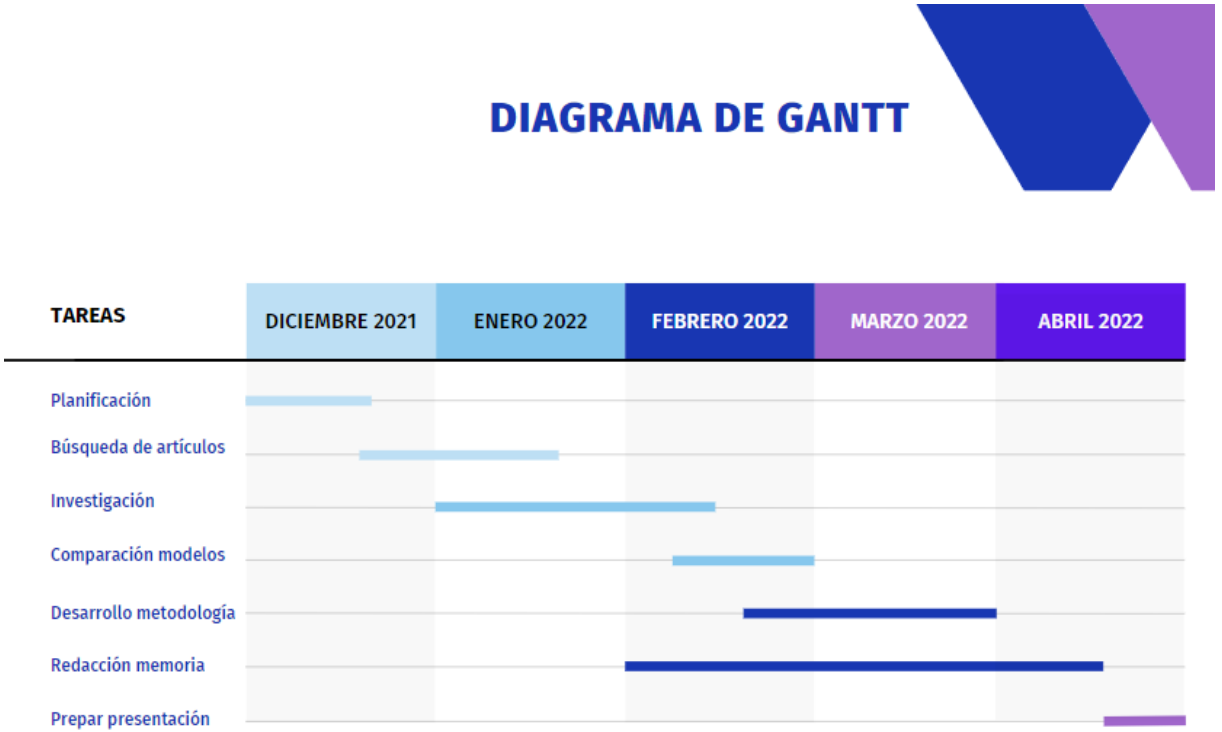

Figura B.1: Diagrama de Gantt inicial.

B.2 Planificación final

La planificación final del proyecto ha presentado diversas desviaciones debido a factores externos al mismo.

Por lo tanto, respecto a la planificación inicial, el proyecto ha extendido su duración hasta el mes de Junio, siendo el mes de Enero en el que se empezaron las labores de búsqueda de artículos de divulgación científica.

De este modo, en la nueva planificación, se ha aumentado sobre todo el tiempo de dedicación a la búsqueda de artículos e investigación, puesto que son dos puntos fundamentales en los que a partir de toda la información elaborada en este punto, se decidirá la hoja de ruta de la propuesta de la nueva metodología.

A continuación se puede observar tanto la Tabla B.2 como la Figura B.2 con la planificación final detallada.

Planificación horaria por tareas			
Nº	**Tarea**	**Tiempo (h.)**	**Dependencia**
1	**Planificación**	**52 horas**	
1.1	Definición del proyecto con la tutora	2 horas	
1.2	Revisión artículos iniciales	10 horas	
1.3	Revisión vídeos de expertos	8 horas	
1.4	Análisis del proyecto y búsqueda de información	28 horas	
1.5	Redactar propuesta del proyecto	4 horas	
2	**Búsqueda de artículos**	**83 horas**	
2.1	Búsqueda de artículos divulgativos	40 horas	
2.2	Estudio y lectura de los artículos	43 horas	2.1
3	**Investigación**	**97 horas**	**2**
3.1	Extraer aspectos importantes de los artículos	62 horas	
3.2	Crear brainstorming de los aspectos relevantes	20 horas	
3.3	Crear un boceto inicial de las ideas obtenidas	15 horas	
4	**Comparación modelos**	**30 horas**	**3**
4.1	Comparación de los modelos existentes	26 horas	
4.2	Identificar pros y contras de los modelos	4 horas	4.1
5	**Desarrollo metodología**	**53 horas**	**4**
5.1	Elaborar nueva metodología	43 horas	
5.2	Comparar la metodología con las existentes	7 horas	5.1
5.3	Elaboración de gráficos y figuras	3 horas	5.1
6	**Redacción memoria**	**113 horas**	**5**
6.1	Redactar la memoria	100 horas	
6.2	Creación de gráficos	13 horas	6.1
7	**Preparar presentación**	**22 horas**	**6**
7.1	Preparar diapositivas presentación ante tribunal	15 horas	
7.2	Preparar exposición oral	7 horas	7.1

Tabla B.2: Desglose de tareas de la planificación horaria final.

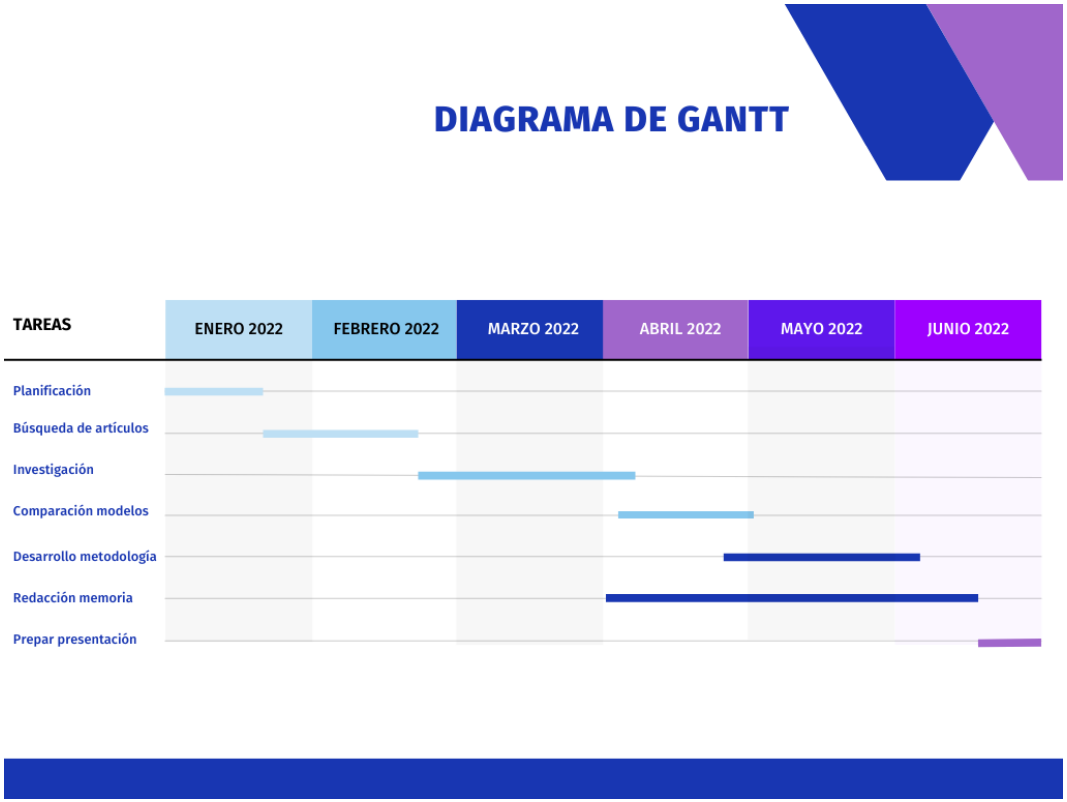

Figura B.2: Diagrama de Gantt final.

B.3 Seguimiento del proyecto

El seguimiento del proyecto se ha desarrollado siguiendo las directrices de la planificación inicial detallada, al igual que siguiendo los tiempos marcados en el diagrama de Gantt.

Asimismo, se han ido planificando reuniones periódicas con la tutora del proyecto para actualizar los avances sobre el mismo y detallar la hoja de ruta de este. Éstas reuniones también han servido de inspiración y para la resolución de dudas que han ido surgiendo.

APÉNDICE C
Gestión de riesgos

Con el fin de gestionar eficientemente el tiempo dedicado al proyecto, se han analizado los posibles riesgos en los que podría verse envuelto por posibles variaciones respecto a la planificación inicial del proyecto. Es pues, en las tablas C.1, C.2 y C.3 se muestra de forma detallada la evaluación y análisis de los posibles riesgos y sus correspondientes tareas de prevención y mitigación.

GR01 - Cambio de la planificación inicial
Identificador: GR01 Nombre: Modificación en la planificación inicial
Afectación: Variable, dependiendo del estado de desarrollo del proyecto.
Descripción: Una modificación de la planificación puede conllevar problemas para la finalización. Prevención: Dedicación de más tiempo en el comienzo del proyecto para determinar la estructuración del proyecto para realizar una planificación correcta. Mitigación: Reestructuración del proyecto optimizando la información ya analizada.

Tabla C.1: GR01 - Modificación en la planificación inicial.

GR02 - Falta de planificación horaria para concluir el proyecto
Identificador: GR02 Nombre: Falta de planificación horaria para ejecutar el proyecto
Afectación: Baja
Descripción: Falta de tiempo por una mala planificación inicial. Prevención: Utilizar herramientas de seguimiento y planificación. Mitigación: Rehacer la planificación y poder obtener tiempo remanente.

Tabla C.2: GR02 - Falta de planificación horaria para concluir el proyecto.

GR03 - Definición de una metodología ya existente
Identificador: GR03 Nombre: Definición de una metodología ya existente
Afectación: Alta.
Descripción: En el momento de la definición de una nueva metodología se puede detectar grandes similitudes con alguna ya existente. Prevención: Dedicar tiempo al análisis exhaustivo de las metodologías existentes. Mitigación: Analizar las deficiencias existentes y localizar el punto de inflexión.

Tabla C.3: GR03 - Definición de una metodología ya existente.

APÉNDICE D
Objetivos de desarrollo sostenible

Grado de relación del trabajo con los Objetivos de Desarrollo Sostenible (ODS).

Objetivos de Desarrollo Sostenible	**Alto**	**Medio**	**Bajo**	**No procede**
ODS 1. **Fin de la pobreza.**				X
ODS 2. **Hambre cero.**				X
ODS 3. **Salud y bienestar.**				X
ODS 4. **Educación de calidad.**				X
ODS 5. **Igualdad de género.**				X
ODS 6. **Agua limpia y saneamiento.**				X
ODS 7. **Energía asequible y no contaminante.**				X
ODS 8. **Trabajo decente y crecimiento económico.**				X
ODS 9. **Industria, innovación e infraestructuras.**	X			
ODS 10. **Reducción de las desigualdades.**				X
ODS 11. **Ciudades y comunidades sostenibles.**				X
ODS 12. **Producción y consumo responsables.**				X
ODS 13. **Acción por el clima.**				X
ODS 14. **Vida submarina.**				X
ODS 15. **Vida de ecosistemas terrestres.**				X
ODS 16. **Paz, justicia e instituciones sólidas.**	X			
ODS 17. **Alianzas para lograr objetivos.**	X			

Reflexión sobre la relación del TFG/TFM con los ODS y con el/los ODS más relacionados.

El trabajo final de máster propuesto en este proyecto, trata de abordar una de las problemáticas más importantes en el ámbito de la obtención de evidencias en el campo del IoT, que se trata de la falta de un *Framework* para realizar el análisis para la obtención de evidencias.

El proyecto, se desarrolla en un contexto de una tecnología IoT cada vez más demandada y cuya previsión de crecimiento se espera que sea exponencial. Asimismo, el número de ciberataques sigue aumentando de forma muy considerable año tras año. Es por este motivo por el que surge la necesidad de poder determinar si un sistema ha sido comprometido ya que un dispositivo IoT puede ser desde un reloj inteligente hasta una planta productiva que tiene interconectados miles de dispositivos IoT.

De este modo, en el proyecto se trata de resolver la problemática de la falta de una metodología para el análisis de evidencias en el ámbito forense y por lo tanto poder remitir pruebas que se consideren válidas ante un tribunal.

El proyecto en mayor o menor medida, tiene un firme compromiso con varios puntos de la Agenda 2030 sobre el Desarrollo Sostenible. Entre ellos, algunos de los objetivos de desarrollo sostenible que pueden relacionarse con el proyecto pueden ser los siguientes:

Industria, innovación e infraestructuras

Se trata del objetivo más relacionado con el proyecto ya que el análisis forense en dispositivos IoT se centra mayoritariamente en dispositivos IoT desplegados en la industria. Además, se trata de poder analizar los hechos ocurridos ante un ciberataque ya sea en la industria como en infraestructuras críticas.

En este punto, se busca una industrialización inclusiva, innovadora y sostenible en la que se propone desarrollar infraestructuras sostenibles, de calidad y fiables.

Este proyecto apoya firmemente el desarrollo tecnológico y sobre todo, la innovación para generar un clima de seguridad y fiabilidad ante las infraestructuras tecnológicas.

Paz, justicia e instituciones sólidas

El análisis de evidencias en un entorno IoT trata de buscar la justicia sobre los hechos ocurridos. Por lo tanto, uno de los puntos más relacionados es este, dado su fin de buscar el origen de lo ocurrido para poder esclarecer los hechos.

Ante este contexto, este proyecto aporta transparencia y seguridad jurídica respecto a la privacidad de los datos involucrados en el análisis. Asimismo, pretende combatir el terrorismo y la delincuencia, si bien no directamente, de una forma indirecta y secundaria, a través del análisis de las evidencias en redes IoT comprometidas.

Alianzas para lograr objetivos

En este campo, la generación de alianzas es fundamental. El uso de varias metodologías existentes permite que no se tenga que reinventar la rueda y poder avanzar sobre estudios ya realizados con el fin de alcanzar ciertos objetivos todavía por cubrir.

Del mismo modo, y tomando este punto de desarrollo sostenible desde una visión tecnológica, las investigaciones de este proyecto, también pretenden servir como una base de ayuda con un fin divulgativo para otras investigaciones en este ámbito tecnológico.

Como puede observarse, el proyecto puede relacionarse en mayor o menor medida con varios objetivos de desarrollo sostenible. El campo del IoT y el análisis forense en estos sistemas, permite generar un clima de seguridad que en su mayor medida beneficia a la justicia, el fortalecimiento de la innovación e infraestructuras, y a través de la divulgación científica, pretende generar alianzas con el fin de lograr objetivos.

ETS Enginyeria Informàtica
Camí de Vera, s/n, 46022, València
T +34 963 877 210
F +34 963 877 219
etsinf@upvnet.upv.es - www.inf.upv.es