Aplicación de la Ingeniería Ontológica para representar la trazabilidad de un Correo Electrónico

Beatriz P. de Gallo[1], Horacio Leone[2]

[1]I.Es.I.Ing. /Facultad de Ingeniería, Universidad Católica de Salta
Campo Castañares S/N, Salta, Argentina
bgallo@ucasal.edu.ar
http://www.ucasal.edu.ar
[2]INGAR/ Facultad Regional Santa Fe UTN
Avellaneda 3657, Santa Fe, Argentina
hleone@santafe-conicet.gov.ar
http://www.ingar.santafe-conicet.gov.ar/

Resumen. Si bien la Forensia Digital avanzó en concordancia con la tecnología, aún se debe trabajar para que los resultados periciales se presenten no como información técnica sino sistemáticamente y semánticamente en el marco de la causa judicial. Resulta conveniente contar con un marco de referencia común para todos los actores judiciales. Este trabajo propone una ontología para el análisis forense de correos electrónicos, focalizándose en la representación de la *trazabilidad de un correo electrónico*. Se toma como caso de estudio el análisis de la cabecera de un correo electrónico y el camino de recorrido inverso de dos destinatarios del correo de ejemplo, estableciendo el modelo ontológico que permite seguir esa trazabilidad.

1 Introducción

El correo electrónico (o e-mail) se ha transformado en el medio de comunicación más utilizado en el tráfico de red facilitando grandemente la comunicación entre las personas. Además de acortar tiempos y distancia, permite el intercambio de múltiples tipos de datos (video, imagen, audio) y se encuentra accesible en prácticamente todos los medios de comunicación tecnológicos, habiendo avanzado rápidamente en la telefonía celular. Esta última característica de "portabilidad" abre instancias de comunicación que antes no estaban presentes, reforzando la inmediatez de la comunicación interpersonal, con el agregado de que ahora existe un registro de esta comunicación. Si bien en el correo electrónico se adopta lenguaje coloquial, y es muy utilizado para la comunicación informal, es importante reconocer que es posible recuperar la conversación y utilizarla como prueba de que tal comunicación existió.

Desde el punto de vista legal, el correo electrónico tiene interés como documento probatorio en un juicio, por lo que resulta importante introducirlo con la fuerza y el rigor técnico suficiente para que actúe en el proceso judicial de igual manera que lo hiciera cualquier otra prueba material. Es importante señalar que la característica de

volatilidad de los datos digitales, impacta negativamente en el reconocimiento de un correo electrónico (o cualquier otro componente digital) como prueba documental. Los profesionales del foro judicial quieren "ver" la prueba, darle forma, buscar el origen, su historia, imaginar todo lo que hay alrededor de este componente, cual si fuera un "arma homicida". Recién en esta instancia pueden reconocer la validez de la prueba digital, i.e., una vez que logran ver la "sustancia" y "consistencia" de elemento probatorio.

Si bien la Forensia Digital avanzó en concordancia con la tecnología, es necesario aún trabajar un aspecto que no es propiamente del ámbito tecnológico y que genera un conjunto de interrogantes que impactan grandemente en los resultados que se obtienen, i.e., la ***interpretación de los resultados***. Harichandran et al. [1] señalan la importancia de mejorar las instancias de comunicación entre los técnicos y los profesionales del derecho, mejorando la accesibilidad y usabilidad de las herramientas de análisis forense para facilitar su interpretación por parte de los no técnicos. El volumen de datos que se obtiene al realizar el análisis forense debe ser interpretado a la luz de la pesquisa. La enorme cantidad de información técnica resultante del análisis de un correo electrónico debe insertarse en el conjunto de pruebas documentales de la causa judicial, colocándolo en un estadio de lectura que facilite la interpretación de esos datos técnicos por parte de los profesionales de la criminalística y el derecho. Se requiere mucho más que la identificación de una dirección IP (Internet Protocol) del correo electrónico. Hoy en día se exige que estos datos se presenten ***sistemáticamente*** y ***semánticamente*** en el marco de la causa judicial, no como información técnica, sino como elemento documental. En el contexto de este requerimiento "no técnico", se encuentra la motivación de este trabajo. Resulta conveniente contar con un marco de referencia basado en la conceptualización formal del universo de discusión. Y en particular, las ontologías resultan una herramienta universal o pluridisciplinar para facilitar el análisis de la prueba documental, por parte de todo los actores (abogados, jueces, investigadores y peritos).

Este artículo se organiza de la siguiente manera. La sección 2 introduce la problemática del análisis forense del correo electrónico. La sección 3 describe brevemente la aplicación de ontologías al análisis forense. Por su parte, la sección 4 presenta la ontología propuesta poniendo énfasis en la representación de los conceptos relacionados con la trazabilidad de un correo. La propuesta es ejemplificada mediante un caso de estudio que se introduce en la sección 5. Finalmente, la sección 6 presenta las conclusiones y los trabajos a futuro.

2 Análisis Forense de un correo electrónico

Tomando como base la tipificación propuesta por Banday [2] para el análisis forense de un correo electrónico se puede identificar tres componentes principales: los actores participantes en la transmisión, la arquitectura lógica y la estructura interna de un correo electrónico. De todos ellos, solo es de interés para una pericia la identificación de los ***actores principales,*** o sea, de aquellas personas que actuaron como probables

emisores y/o receptores del correo electrónico bajo análisis. En referencia a esto se puede decir que, si bien la comunicación de un correo electrónico requiere de personas que actúan como emisor y receptor del mensaje, no son éstos los únicos partícipes de la transmisión. Banday [2] establece un mapa de relaciones y caminos posibles que puede recorrer un correo durante el proceso de transmisión e identifica los procesos responsables de sostener el servicio –denominados *actores*- que actúan internamente durante la transmisión.

Básicamente, un correo electrónico es manejado por un mínimo de cuatro equipos distintos: el equipo emisor, el servidor de correo del remitente, el servidor de correo del receptor y el equipo receptor. En todos ellos, el proceso de transmisión *deja una huella* del correo emitido, que se encuentra en la *cabecera del correo*. Los MTA[1] añaden una etiqueta de identificación en la cabecera del correo cada vez que el mail ingresa a un servidor. En base a estos elementos mencionados, es posible realizar la ***trazabilidad*** de un correo electrónico. La norma ISO 9000:2015 [3] define trazabilidad como la "capacidad para seguir el histórico, la aplicación o la localización de un objeto; al tratarse de un producto o servicio, la trazabilidad puede estar relacionada con el origen de los materiales y las partes, el histórico del proceso y la distribución y localización del producto o servicio después de la entrega". La principal ventaja que reporta la trazabilidad (o logística inversa) es poder conocer a ciencia cierta la procedencia y la historia que atañe a un producto. Existen investigaciones que relacionan la trazabilidad y las ontologías. En [4] se propone una ontología para identificar la trazabilidad a lo largo del ciclo de vida del proceso unificado, mientras que en [5] formulan una ontología para identificar la trazabilidad del código fuente y la documentación para apoyar procesos de ingeniería inversa en el mantenimiento de software. Aplicado a un correo electrónico, la trazabilidad permitiría establecer el proceso desarrollado durante la comunicación que lleva del receptor al emisor, siendo éste –en última instancia- el objeto esencial del análisis forense de un correo electrónico.

Guo et al. [6] describe un procedimiento para el análisis forense de un correo electrónico a partir de los datos de la cabecera y resalta la necesidad de validar la autenticidad e integridad del correo a partir de estos metadatos. Por su parte, Daniel et al. [7] señala que durante el proceso de transmisión el correo electrónico puede quedar almacenado en distintos espacios: servidores de correo para cuentas empresariales/laborales, servidores de correo para cuentas de uso gratuito, servidores de los ISP[2], memoria cache del equipo emisor, almacenamiento en el cliente local del equipo emisor, entre otros, señalando además que existen características propias de software para la gestión del correo residente en cada tipo de espacio.

[1] MTA: Message Transfer Agents, agente de transferencia del mensaje. Proceso responsable de registrar en la cabecera del mensaje los datos referidos al servidor en donde se almacenó el correo durante el camino de transmisión.

[2] ISP = Internet Service Provider, proveedor del servicio de internet.

2.1 Los puntos de pericia de un correo electrónico

Una pericia es un conjunto de operaciones técnicas científicas puestas en práctica para el esclarecimiento de un posible hecho ilícito y ordenadas por el Tribunal interviniente [8].

En cuanto a los ***puntos de pericia***, su ofrecimiento permitirá al Juez determinar la procedencia de la prueba, es decir, la congruencia entre los aspectos a conocer y la necesidad de un técnico para que lo asesore. Los puntos periciales se proponen en un pliego que señala las cuestiones técnicas, de manera clara y precisa, siempre referidas al tema que se dilucida en la litis y que técnicamente puedan ser respondidas por el Perito. Usualmente los puntos de pericia referidos a correos electrónicos abordan cuestiones relacionadas con la verificación de la *autenticidad y existencia* de un correo electrónico.

Los elementos que permiten verificar la ***autenticidad*** de un correo electrónico son los siguientes:

- la identificación de los datos del remitente (nombre de usuario, cuenta de correo y dirección IP),
- la trazabilidad del mismo (diferentes servicios o agentes que intervienen en la transmisión), y
- los datos del destinatario (nombre de usuario, cuenta de correo y dirección IP.

En cuanto a la ***existencia*** de un correo electrónico, ésta se puede probar fehacientemente cuando se comprueba la presencia del archivo digital del mismo tanto en el dispositivo emisor (o en el servidor del ISP del emisor) como en el dispositivo receptor del correo (o en el servidor ISP del receptor); y ambos archivos digitales son idénticos.

Desde el punto de vista de la forensia digital, existen muchas técnicas y herramientas que ayudan al Perito Informático en el análisis de un correo electrónico, algunas ya fueron analizadas en [9], a los citados allí se agrega el trabajo de Devendran, Shahriar y Clincy [10] acerca de un estudio comparativo de varios software open source para el análisis de correos electrónicos. Sobre este particular, se está trabajando en el análisis de un conjunto de herramientas propias de la forensia de correos electrónicos, a fin de identificar los datos que se pueden obtener con cada una (encabezado, IP, id del mensaje, casillas de correo intervinientes, fechas, hora, etc.) y que puedan ser utilizados para poblar la ontología que se propone en este trabajo.

3 Aplicación de Ontologías al Análisis Forense

Se recurre a las ontologías para representar la multiplicidad de dominios expertos que se requieren en el análisis forense (desde conocimiento de redes hasta conocimientos de sistemas contables) [11], o bien para el diseño de un sistema inteligente en red aplicado a la forensia [12]. Zhu [13] propone una ontología para compartir información sobre patrones de ciberataques y realizar un análisis sistemático y más eficien-

te de la información. Para el caso particular de los correos electrónicos, es de interés el trabajo de Balakumar et al. [14] sobre la definición de una ontología para la clasificación y categorización de e-mail con el objetivo de conformar un filtro para la detección de spam mediante la conformación de una *whitelists* de remitentes conocidos, así como el trabajo de clasificación de e-mails propuesto por Taghva [15] en referencia a la exigencia legal de resguardar ciertos registros de datos residentes o anexados a correos electrónicos.

Si bien existen algunas propuestas que aplican ontologías al análisis forense, éstas no tienen en cuenta la necesidad de desarrollar un espacio *integrado y orientado al sujeto en litis* que les permita a los profesionales del derecho y peritos de otras especialidades una mejor interpretación de los datos técnicos obtenidos como resultado de la pericia de un correo electrónico. A fin de dar respuesta a esta necesidad se propone una ontología que actúa como marco de interpretación pluridisciplinar, la cual se introduce en la siguiente sección.

4 Ontología propuesta para el Análisis Forense de Correos Electrónicos

La ontología que se aborda en este trabajo se formuló inicialmente en [9], sentando las bases de los principales componentes: preguntas de competencias, conceptos, relaciones, etc. Ahora se propone avanzar en la construcción de la ontología mediante el refinamiento de estos primeros conceptos incorporando todo lo relacionado con ocurrencias de correos y su trazabilidad y ejemplificar con un caso de estudio.

Los interrogantes base de esta ontología se pueden buscar en los puntos de pericia que usualmente se proponen al solicitar un análisis forense de un correo electrónico y que se han enunciado en el apartado anterior, de allí se extractan las ***preguntas de competencia*** en lenguaje natural:

1. ¿Cuáles son las partes de un correo electrónico que resultan de interés para un análisis forense?
2. ¿Cuáles son los componentes informáticos a través de los cuales se escribe y se lee un correo electrónico?
3. ¿Cuáles son los datos o componentes que permiten validar la existencia de un correo electrónico? Esta pregunta puede descomponerse en:
   3.1. ¿Cuál es la fecha, hora y dirección IP de emisión del correo electrónico?
   3.2. ¿Cuál es la fecha, hora y dirección IP de recepción del correo electrónico?
4. Dado un correo electrónico ¿Cuáles son los datos que permiten identificar la autoría y recepción del mismo? Esta pregunta puede descomponerse en:
   4.1. ¿Cuál es el nombre de usuario y dirección de e-mail del Autor del mismo?
   4.2. ¿Cuál es el nombre de usuario y dirección de e-mail del Receptor del mismo?
   4.3. ¿Es posible establecer el seguimiento del mensaje desde que se envía hasta que se recibe?

4.4. ¿Cuáles son los diferentes actores/servicios que participaron de la transmisión?

El desarrollo de las diferentes etapas de la ontología se muestra en los trabajos [9], [16] y [17] en los que se describen los principales conceptos de la ontología, así como el diccionario de conceptos y la tabla de definición de objetos y relaciones del modelo. A continuación se introducen los conceptos y relaciones de la ontología propuestos para dar respuesta a estas preguntas de competencia a través del concepto de trazabilidad aplicado a un correo electrónico.

Considerando el "camino" que realiza un correo electrónico desde su emisión hasta la recepción por parte del destinatario, Banday [2] menciona diferentes actores y procesos que se van desarrollando durante la transmisión. De esta secuencia de acciones, interesan en particular aquellas que pueden impactar en la *modificación del paquete de datos* que circula. Entonces, considerando que durante la transmisión el correo electrónico va residiendo en diferentes dispositivos de almacenamiento (equipo emisor, servidor de correo, servidores de paso, equipo receptor) es posible tomar cada archivo almacenado en estos dispositivos y verificar si el paquete de datos "original" fue modificado en algún punto durante la transmisión hasta llegar a destino. La verificación de una posible alteración del correo se realiza chequeando que el correo es el *mismo* en cada dispositivo en donde se va almacenando durante la transmisión. Esto es en cuanto al *cuerpo del mensaje,* que debería mantenerse inalterable mientras que la *cabecera* del correo se va extendiendo a medida que se va almacenando en los servidores durante todo el camino de la transmisión.

Con independencia de las diversas definiciones de *correo electrónico* que ya se han formulado, en el marco de este trabajo se define al mismo en función de los elementos necesarios para la realización del análisis forense, i.e., *un correo electrónico es un documento digital que consta de dos partes: a) una cabecera que contiene información sobre el proceso de transmisión que se desarrolla con identificación de las cuentas intervinientes y los distintos servidores en que el correo se fue almacenando durante la transmisión; y b) un cuerpo que contiene el mensaje que se transmite.*

A partir de esta definición, en este trabajo se propone utilizar un enfoque ontológico para representar la trazabilidad de un correo electrónico basado en los siguientes aspectos de interés: a) el correo *original* se descompone en *ocurrencias*, tantas como veces se almacena en los distintos dispositivos durante el proceso de transmisión; b) existen tres tipos de ocurrencias: de emisión, de transmisión y de recepción; c) mientras que las ocurrencias de emisión y recepción son únicas, las de transmisión serán tantas, como veces en que el correo en transmisión quedó almacenado en un dispositivo durante su camino; d) las ocurrencias se asocian mediante una *secuencia* que establece el orden de aparición de cada ocurrencia, y con tantos *hilos* como receptores tenga el correo; y e) por otra parte, desde el punto de vista pericial, se requiere un conjunto de datos que permitan identificar plenamente al hardware (equipo emisor, servidor de correo, servidor de paso, equipo receptor) en donde el correo estuvo almacenado. Estos datos se incluyen en la clase localización que además recoge información derivada de los datos originales (geolocalización de la IP, ISP, entre otros).

El modelo conceptual especificado debe representarse por medio de un lenguaje formal, dando lugar a los componentes de la ontología: clases, atributos, conceptos, relaciones, funciones, axiomas e instancias. Las Fig. 1 y 2 muestran sendas vista par-

ciales de la ontología propuesta[3]. En tanto, en el anexo A se presenta una vista parcial de la formalización de la ontología propuesta, que incluye los principales axiomas relacionados con los conceptos presentados a continuación.

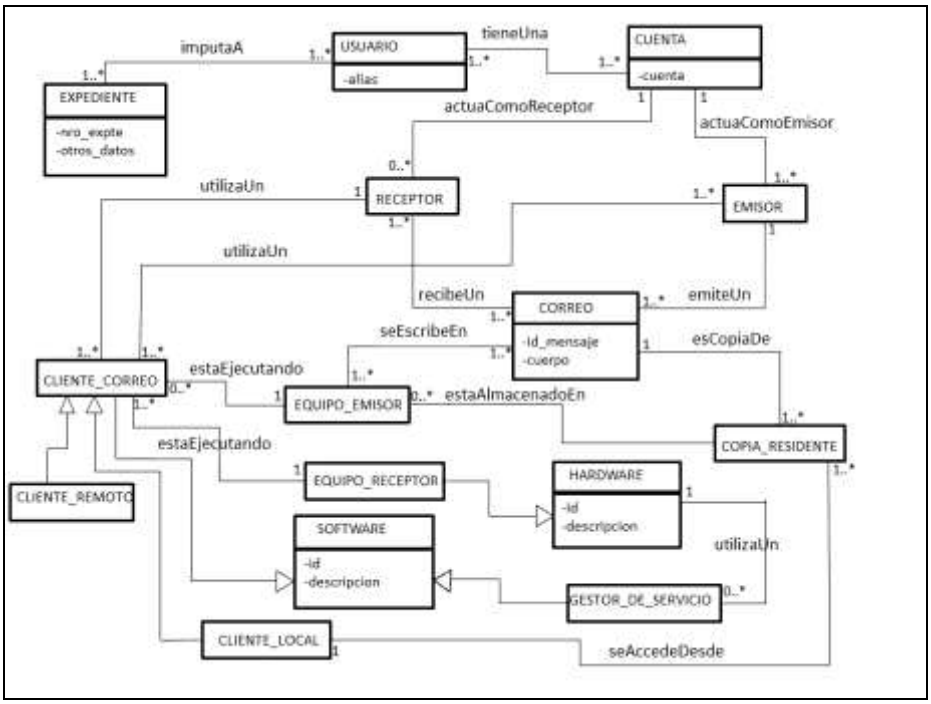

Figura 1: Vista Parcial de la ontología propuesta

Las principales clases identificadas son *CORREO, EMISOR y RECEPTOR.* El concepto *CORREO* se definió en los párrafos precedentes, *EMISOR* es la cuenta de correo desde la cual se emite el correo electrónico *y RECEPTOR* es la cuenta de correo destinataria del mismo. *CORREO* es una clase que se asocia con las clases *EMISOR* y *RECEPTOR* mediante las relaciones de *emitUN* y *recibeUn* respectivamente. A su vez las clases *EMISOR* y *RECEPTOR* se vinculan con la clase *CUENTA* mediante la relación *actuaComoEmisor* y *actuaComoReceptor* pues una misma cuenta puede actuar como emisor en una oportunidad y como receptor en otra. Un correo tiene un único emisor y podría tener uno o más receptores (ver axiomas 1 y 2 en el anexo A).

En el momento de la emisión del correo interactúan varios elementos, representados en el modelo conceptual por las clases *EQUIPO_EMISOR* y *CLIENTE_CORREO* mediante la asociación señalada como *estaEjecutando*. Cuando se escribe el correo se puede utilizar un *CLIENTE_CORREO* que sea *CLIENTE_REMOTO* en caso de utilizar un sistema web mail o que sea un *CLIENTE_LOCAL* en caso de un software de

[3] Por razones de espacio no se incorporan los resultados de las primeras etapas de la metodología (Glosario de Términos, Taxonomía de Conceptos y definición de las relaciones binarias ad-hoc) que sustentan el modelo que aquí se describe.

correo instalado en el equipo y con una cuenta de mail permanente, de allí que *CLIENTE_REMOTO* y *CLIENTE_LOCAL* se describen como subclases de *CLIENTE_CORREO*. Si se utiliza un *CLIENTE_LOCAL*, entonces existirá una *COPIA_RESIDENTE* del correo en el *EQUIPO_EMISOR,* asociado al *CORREO* mediante la relación *esCopiaDe* y al *EQUIPO_EMISOR* con la relación *estaAlmacenadaEn.*

Como se mencionara anteriormente, la ontología propuesta incorpora conceptos que permiten realizar la trazabilidad de un correo entre el emisor y un receptor. Los mismos se describe a continuación y se ilustran en la Fig. 2.

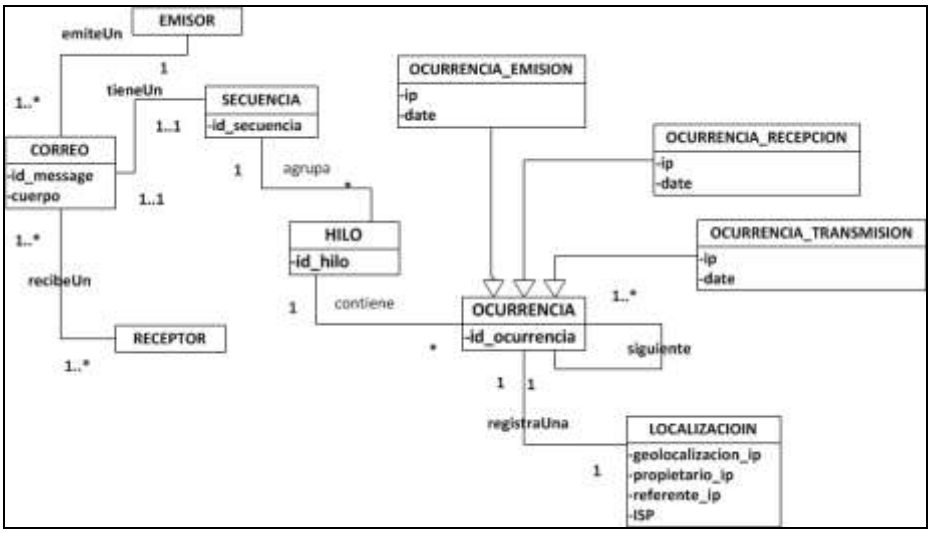

Figura 2: Representación de las ocurrencias de un correo electrónico

Si bien durante el proceso de transmisión el contenido del correo no se altera, cada vez que éste pasa por un servidor se va actualizando los datos de cabecera. A fin de lograr la trazabilidad será necesario identificar estas sucesivas versiones del mismo correo describiéndolas en la ontología propuesta mediante el concepto de OCURRENCIA. Así, un correo es una secuencia de ocurrencias representada mediante la relación *tieneUna* entre *CORREO y SECUENCIA*. Por otra parte, al representar el "camino recorrido" desde el emisor a un receptor, las ocurrencias conforman un *HILO.* Para representar esto, la ontología propone dos relaciones: i) *contiene* entre *HILO* y *OCURRENCIA* que especifica las ocurrencias que conforman un hilo y ii) la relación recursiva *siguiente* que establece el orden en que se fueron generando cada una de las ocurrencias en cada hilo. Para un correo electrónico, existirán tantos hilos como receptores tenga dicho correo ya que para ubicar a cada receptor puede ser necesario recorrer distintas vías de comunicación. Es así que, una secuencia agrupa uno o más hilos, esto se describe en la ontología mediante la relación de composición *agrupa* entre *SECUENCIA e HILO.* Todo correo tiene una y sólo una secuencia, la cual tiene al menos un hilo de ocurrencias. (ver axiomas 4 - 6 en el anexo A).

Por otra parte, en un HILO es posible identificar diferentes tipos de ocurrencias: *OCURRENCIA_EMISION, OCURRENCIA_TRANSMISION y OCURREN-*

CIA_RECEPCION, los cuales se definen a continuación y se formalizan en los axiomas (7), (8) y (9) del anexo A. La *OCURRENCIA_EMISION* es la primera ocurrencia que se genera para el correo al momento de emitirlo, y es siempre única. Es la primera ocurrencia de la *SECUENCIA* y es compartida por todos los hilos que se generan en una misma SECUENCIA. La *OCURRENCIA_RECEPCIÓN* se genera al momento de la recepción del correo y es la última de cada *HILO*. Si bien es única para cada *HILO*, existen tantas *OCURRENCIA_RECEPCION* como hilos integren la secuencia del correo. La *OCURRENCIA_TRANSMISION* es el conjunto de ocurrencias que se generan durante los pasos internos del correo de servidor a servidor. Se distinguen porque siempre tienen una ocurrencia anterior y una ocurrencia siguiente y además porque no pueden estar en dos hilos diferentes. Se ha incorporado una clase *LOCALIZACION* asociada por medio de la relación *RegistraUna* a la clase *OCURRENCIA*. Cada instancia de la clase *LOCALIZACION* permite representar los datos que ayudan a la identificación del servidor en donde se almacena la ocurrencia asociada a dicha instancia. En el caso particular de la localización asociadas a ocurrencias de emisión y recepción, se podrá registrar también otra información de interés para la causa judicial, como ser la localización geográfica del servidor, el proveedor del servicio de internet, el propietario registrado para la dirección IP, entre otros.

5 Caso de Estudio

A fin de ejemplificar esta propuesta, se formula un caso de estudio de un correo electrónico público dirigido desde el emisor *beagallo@gmail.com* a dos receptores: *erivetti83@gmail.com* y *bgallo@ucasal.edu.ar*.

Return-Path: <beagallo@gmail.com>
Delivered-To: bgallo@ucasal.edu.ar
Received: …[4]
Received: from mail.ucasal.edu.ar (mail.ucasal.edu.ar [127.0.0.1])
by mail.ucasal.edu.ar (Postfix) with ESMTP id 9F9CFEB2A8
for <bgallo@ucasal.edu.ar>; Sat, 7 May 2016 15:49:36 -0300 (ART)
X-Virus-Scanned: ….[4]
Received: from mail-oi0-f48.google.com (mail-oi0-f48.google.com [209.85.218.48])
by mail.ucasal.edu.ar (Postfix) with ESMTP id CF8D9EB263
for <bgallo@ucasal.edu.ar>; Sat, 7 May 2016 15:49:35 -0300 (ART)
X-Google-DKIM-Signature:…[4]
Received: by 10.182.24.100 with HTTP; Sat, 7 May 2016 11:49:34 -0700 (PDT)
Reply-To: beagallo@gmail.com
Date: Sat, 7 May 2016 15:49:34 -0300
Message-ID: <CAH18OQUxurYQNd0b-+N7YO2zRTLWhPcQd3P6bA+D85F5VXzptg@mail.gmail.com>
Subject: Mail de prueba para el trabajo de JAIIO
From: "Ing. H. Beatriz P. de Gallo" <beagallo@gmail.com>
To: Esteban Rivetti <erivetti83@gmail.com>
Cc: bgallo <bgallo@ucasal.edu.ar>

Figura 3: Cabecera correo recibido por *bgallo@ucasal.edu.ar*

En la Figura 3 se representa la cabecera del correo recibido por bgallo@ucasal.edu.ar conformado lo que denominamos hilo_1 y en la Figura 4 se representa la cabecera del correo recibido por erivetti83@gmail.com conformado el hilo_2.

[4] Se recortaron partes del archivo cabecera que no son de interés para el ejemplo

La cabecera de un correo debe leerse “de abajo hacia arriba” para identificar los diferentes servidores que se utilizaron durante la transmisión, ya que el MTA actualiza los datos agregando una línea siempre en el extremo superior de la cabecera. Con el fin de dar claridad al ejemplo, se ha señalado con barras sombreadas los datos agregados por el MTA en cada caso.

```
Delivered-To: erivetti83@gmail.com
Received: by 10.194.68.33 with SMTP id s1csp907415wjt;
        Sat, 7 May 2016 11:49:34 -0700 (PDT)
X-Received: by 10.202.64.132 with SMTP id n126mr11108346oia.80.1462646974689;
        Sat, 07 May 2016 11:49:34 -0700 (PDT)
Return-Path: <beagallo@gmail.com>
Received: from mail-oi0-x22f.google.com (mail-oi0-x22f.google.com. [2607:f8b0:4003:c06::22f])
        by mx.google.com with ESMTPS id a12si9412619oez.93.2016.05.07.11.49.34
        for <erivetti83@gmail.com>
        (version=TLS1_2 cipher=ECDHE-RSA-AES128-GCM-SHA256 bits=128/128);
        Sat, 07 May 2016 11:49:34 -0700 (PDT)
Received-SPF…[4]
X-Received: by 10.157.16.27 with SMTP id h27mr12724802ote.118.1462646974360;
 Sat, 07 May 2016 11:49:34 -0700 (PDT)
Received: by 10.182.24.100 with HTTP; Sat, 7 May 2016 11:49:34 -0700 (PDT)
Reply-To: beagallo@gmail.com
Date: Sat, 7 May 2016 15:49:34 -0300
Message-ID: <CAH18OQUxurYQNd0b-+N7YO2zRTLWhPcQd3P6bA+D85F5VXzptg@mail.gmail.com>
Subject: Mail de prueba para el trabajo de JAIIO
From: "Ing. H. Beatriz P. de Gallo" <beagallo@gmail.com>
To: Esteban Rivetti <erivetti83@gmail.com>
Cc: bgallo <bgallo@ucasal.edu.ar>
```

Figura 4: Cabecera correo recibido por *erivetti@gmil.com*

La Fig. 5 esquematiza el proceso de transmisión, describiendo la secuencia y los hilos que se generan. Nótese que los datos referentes a la *OCURRENCIA_EMISION* no están incluidos en las cabeceras, los mismos se obtienen de la revisión de los equipos que actuaron como receptores, y que esta ocurrencia es la misma para ambos hilos. Se puede observar que la cabecera del receptor *bgallo@ucasal.edu.ar* está representada por el HILO_1 del que participan 4 ocurrencias (la de emisión, recepción y 2 de transmisión), mientras que la cabecera del receptor *erivetti83@gmail.com* representada por el HILO_2 participan 6 ocurrencias (la de emisión, recepción y 4 de transmisión).

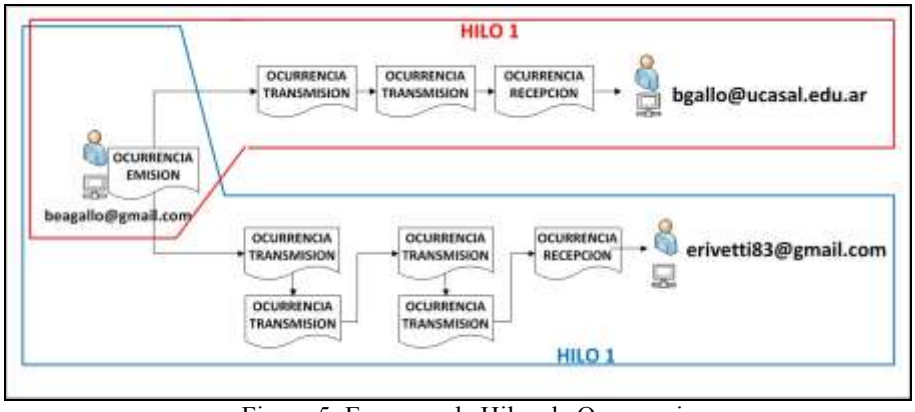

Figura 5: Esquema de Hilos de Ocurrencias

En la Figura 6 se describe las instancias de objetos generadas para el caso de estudio. Obsérvese que ajustándose al caso de estudio, se ha representado una instancia de *CORREO*, de *EMISOR*, dos instancias de *RECEPTOR*, y una instancia de *SECUENCIA*, que contiene dos instancias de *HILO*, cada una de las cuales tiene 4 y 6 *OCURRENCIAS* respectivamente.

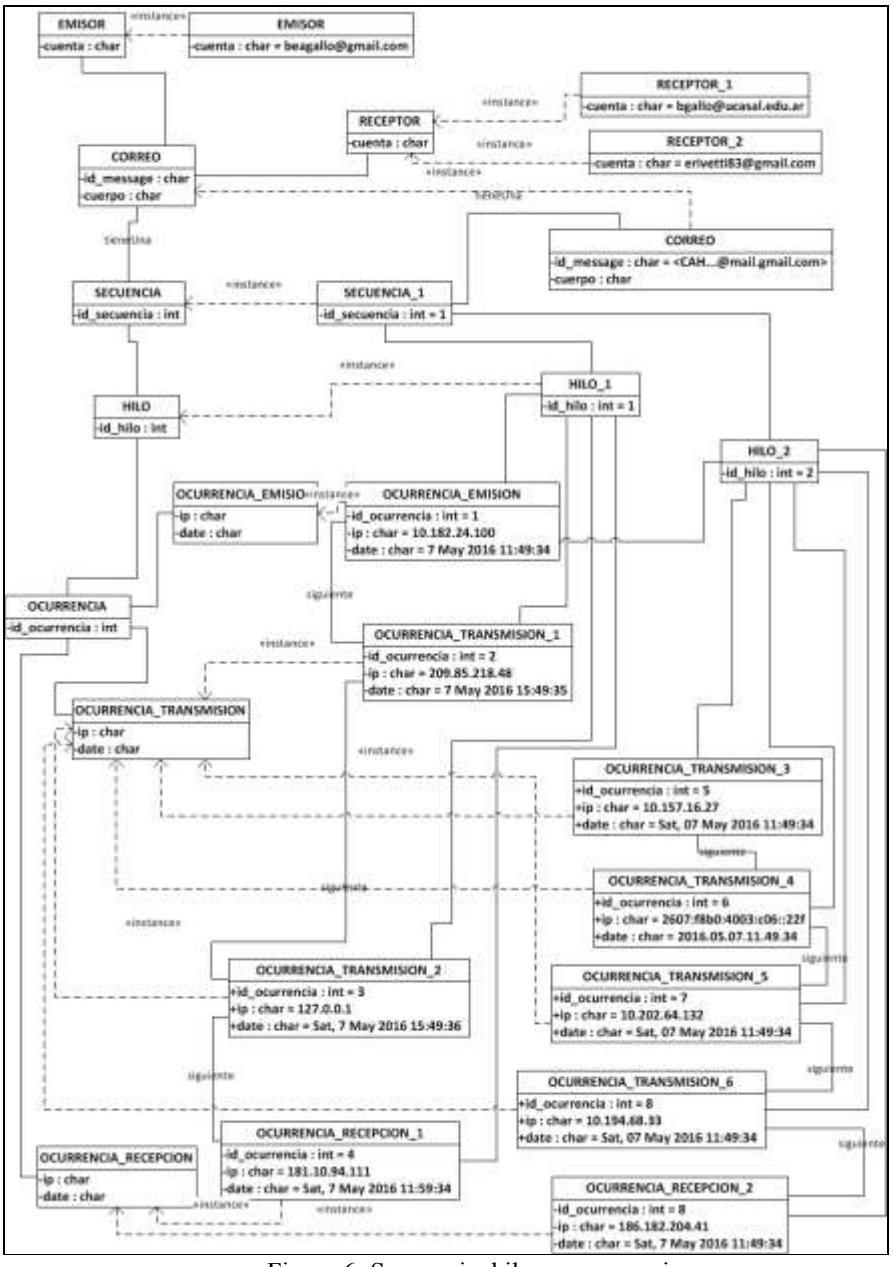

Figura 6: Secuencia, hilos y ocurrencias

6 CONCLUSIONES

En este trabajo se presentó el avance logrado hasta el momento para representar el proceso de transmisión de un correo electrónico, desde el punto de vista de la trazabilidad. La necesidad de reconstruir el camino de inverso de un correo electrónico recibido se sustenta en que de este modo se puede probar la *existencia y autenticidad* del correo, avalando el carácter probatorio de este documento digital y reforzando la capacidad de *no repudio* del documento digital.

Por otra parte, representar esta característica de trazabilidad del correo electrónico mediante una ontología, permite establecer un marco referencial científico y metodológico validado, que sirve de entorno de comunicación entre los partícipes informáticos y no informáticos del proceso judicial, al incorporar al proceso de análisis forense los beneficios de la ingeniería ontológica.

Entre las líneas previstas para la continuación de este trabajo se señala:

- Continuación del ciclo de construcción de la ontología avanzando en la formalización de las restantes reglas que permitan restringuir el modelo e inferir nuevo conocimiento, instanciación y validación de la ontología.
- Verificación de los axiomas de autenticidad y existencia mediante una contrastación de puntos de pericia relevados entre los profesionales de la informática forense.
- Desarrollo del marco jurídico a partir de la incorporación de ontologías existentes sobre la temática, como por ejemplo SC-ONT propuesta por Kalemi et al. [18] que representa el dominio criminal a partir de las redes sociales en línea, en las que el correo electrónico es profusamente utilizado.

Referencias

1. Harichandran, V. S., Breitinger, F., Baggili, I., & Marrington, A. (2016). A cyber forensics needs analysis survey: Revisiting the domain's needs a decade later.Computers & Security, 57, 1-13.
2. Banday, M. Tarik, "TECHNIQUES AND TOOLS FOR FORENSIC INVESTIGATION OF E-MAIL", International Journal of Network Security & Its Applications (IJNSA), Vol.3, No.6, November 2011.
3. ISO 9000-2015 (Traducción Oficial), Instituto Argentino de Normalización, Argentina, 2015
4. Noll, R. P., & Ribeiro, M. B. (2007, March). Enhancing traceability using ontologies. In Proceedings of the 2007 ACM symposium on Applied computing(pp. 1496-1497). ACM.
5 Kim, H. M., Fox, M. S., & Grüninger, M. (1999). An ontology for quality management—enabling quality problem identification and tracing. BT Technology Journal, 17(4), 131-140.
6. Guo, H., Jin, B., & Qian, W. (2013, April). Analysis of Email Header for Forensics Purpose. In Communication Systems and Network Technologies (CSNT), 2013 International Conference on (pp. 340-344). IEEE.

7. Daniel, L. Daniel, L. Digital Forensics for Legal Professionals, Understanding Digital Evidence From The Warrant To The Courtroom, 2012, ISBN: 978-1-59749-643-8
8. Fernández, Eduardo Enrique: "Aspectos legales del peritaje". Revista INDICIOS, Año 2. Vol. 2. La Rioja (Argentina) 2011. pp. 24-33.
9. Gallo Beatriz P. de, Vegetti Marcela, Leone Horacio, "Ontología para el Análisis Forense de Correo Electrónico", CoNaIISI 2014 Actas del 2° Congreso Nacional de Ingeniería Informática/Sistemas de Información, San Luis, Argentina, ISSN: 2346-9927, 2014
10. Devendran, Vamshee Krishna, Hossain Shahriar, and Victor Clincy. "A Comparative Study of Email Forensic Tools." Journal of Information Security6.2 (2015): 111
11. Schatz, B., Mohay, G. M., & Clark, A. (2004). Generalising event forensics across multiple domains. School of Computer Networks Information and Forensics Conference, Edith Cowan University.
12. Saad, S., & Traore, I. (2010, August). Method ontology for intelligent network forensics analysis. In Privacy Security and Trust (PST), 2010 Eighth Annual International Conference on (pp. 7-14). IEEE.
13. Zhu, Y. (2015). Attack pattern ontology: A common language for attack information sharing between organizations (Doctoral dissertation, TU Delft, Delft University of Technology).
14. Balakumar, M., & Vaidehi, V. (2008, January). Ontology based classification and categorization of email. In Signal Processing, Communications and Networking, 2008. ICSCN'08. International Conference on (pp. 199-202). IEEE
15. Taghva, K., Borsack, J., Coombs, J., Condit, A., Lumos, S., & Nartker, T. (2003, April). Ontology-based classification of email. In Information Technology: Coding and Computing, International Conference on (pp. 194-194). IEEE Computer Society.
16 Gallo Beatriz P. de, Vegetti Marcela, Leone Horacio, "Población de ontologías con datos no estructurados utilizando herramientas de minería de datos", CoNaIISI 2015 Actas del 3° Congreso Nacional de Ingeniería Informática/Sistemas de Información, Buenos Aires, Argentina, ISBN: 978-987-1896-47-9, 2015
17. Gallo Beatriz P. de, Vegetti Marcela, Leone Horacio, "Avances en la Construcción de una Ontología para el Análisis Forense de Correo Electrónico", VI CIIDDI, Actas del 6° Congreso Ibeoramericano de Investigadores y Docentes de Derecho e Informática, Santa Fe, Argentina, en proceso de publicación, 2016.
18.Kalemi, E., & Yildirim-Yayilgan, S. (2016). Ontologies for Social Media Digital Evidence. World Academy of Science, Engineering and Technology, International Journal of Computer, Electrical, Automation, Control and Information Engineering, 10(2), 335-340.

Anexo A: Vista parcial de la Formalización propuesta

Para la formalización se representan los conceptos como predicados unarios, por ejemplo: *Correo(x)* significa que *x* es una instancia del concepto *CORREO*. Las relaciones se representan como predicados binarios donde el símbolo predicativo es el nombre de la asociación y los términos del predicado representan las instancias de los conceptos vinculados, por ejemplo: *emiteUn(e,c)* representa la relación entre una instancia de *EMISOR* y una instancia de *CORREO*. A continuación se presentan los axiomas más relevantes relacionados con los conceptos presentados en este artículo.

Todo correo tiene un emisor y es único"

$$\forall x Correo(x) \exists e / Emisor(e) \wedge emiteUn(x,e) \quad (1)$$

$$\forall x / Correo(x) \exists e_1, e_2 / Emisor(e_1) \wedge Emisor(e_2) \wedge emiteUn(x, e_1) \wedge emiteUn(x, e_2) \Rightarrow e_1 = e_2 \quad (2)$$

"Todo correo debe tener al menos un receptor"

$$\forall x Correo(x) \rightarrow \exists r / \text{Re}ceptor(r) \wedge tieneUn(x, r) \quad (3)$$

Todo correo tiene una secuencia de ocurrencias

$$\forall x Correo(x) \exists s / Secuencia(s) \wedge tieneUna(x, s) \quad (4)$$

$$\forall x / Correo(x) \exists s_1, s_2 / Secuencia(s_1) \wedge Secuencia(s_2) \wedge tieneUna(x, s_1) \wedge tieneUna(x, s_2) \Rightarrow s_1 = s_2 \quad (5)$$

"Toda secuencia está conformada por al menos un hilo de ocurrencias"

$$\forall x Secuencia(x) \exists h Hilo(h) \wedge contiene(x, h) \quad (6)$$

"Toda ocurrencia que no tenga una anterior será una ocurrencia de emisión"

$$\forall x Ocurrencia Emision(x) \Leftrightarrow \exists x, x_1 Ocurrencia(x) \wedge not \exists Ocurrencia(x_1) \wedge siguiente(x_1, x) \quad (7)$$

"Toda ocurrencia que tenga una anterior y una siguiente será una ocurrencia de transmisión"

$$\forall x Ocurrencia Transmision(x) \Leftrightarrow \exists x_1, x_2 Ocurrencia(x_1) \wedge Ocurrencia(x_2) \wedge siguiente(x_1, x) \wedge siguiente(x, x_2) \quad (8)$$

"Toda ocurrencia que no tenga una siguiente será una ocurrencia de recepción"

$$\forall x Ocurrencia Emision(x) \Leftrightarrow \exists x, x_1 Ocurrencia(x) \wedge not \exists Ocurrencia(x_1) \wedge siguiente(x, x_1) \quad (9)$$