

Processing of elicited information: good practices in requirements engineering

Procesamiento de la información elicitada: buenas prácticas en la ingeniería de requisitos

Graciela D. S. Hadad¹[0000-0003-4909-9702], Jorge H. Doorn^{1,2}, y M. Celia Elizalde¹

¹ Instituto de Ingenierías y Nuevas Tecnologías,
Universidad Nacional del Oeste, Merlo, Argentina

² Departamento de Ciencia y Tecnología,
Universidad de Tres de Febrero
{ghadad, jdoorn, melizalde}@uno.edu.ar

Abstract. The elicitation of information in the requirements process is usually done through interviews. However, the practical indications about how to carry interviews out and how to use the elicited information when developing models do not prevent the appearance of some relevant systematic errors. Both in the literature and in the experiments carried out, it has been detected that the interviewer and/or the modeler usually introduce subjective information not obtained from the source of information, but from their previous knowledge. This distortion in the modeled information is usually transferred to other models and software artifacts and is detected too late. Therefore, an experiment was carried out that studied in detail how the information was obtained through an interview and how it was processed to build a model written in natural language. The use of incorrectly named concepts, non-existent concepts, poorly understood information, and vocabulary induced by the interviewer were detected in the model. Based on the distortions found in the modeled information, preliminary recommendations were made on how to conduct interviews and how to process the information for modeling, to mitigate cognitive biases of requirements engineers.

Keywords: Requirements Engineering, Interviews, Natural Language Models, Cognitive Factors

Resumen. La obtención de información en el proceso de requisitos se realiza usualmente mediante entrevistas. Sin embargo, pese a su uso frecuente, las indicaciones prácticas disponibles acerca de cómo llevarlas a cabo y cómo usar posteriormente la información obtenida para elaborar modelos que reflejen el conocimiento adquirido, no impiden la aparición de algunos errores sistemáticos relevantes. Tanto en la literatura como en los propios experimentos realizados se ha detectado que el entrevistador y/o el modelador suelen introducir información subjetiva no obtenida de la fuente de información, sino proveniente

Received January 2024; Accepted March 2024; Published May 2024

de su conocimiento previo. Esa distorsión en la información modelada suele trasladarse a otros modelos y artefactos de software y detectarse tardíamente. Por lo tanto, se realizó un experimento que estudió en detalle cómo la información fue elicitada a través de una entrevista y cómo fue procesada para construir un modelo escrito en lenguaje natural. Se detectó en el modelo el uso de conceptos nombrados incorrectamente, conceptos inexistentes, información mal comprendida, y vocabulario inducido por parte del entrevistador. A partir de las distorsiones encontradas en la información modelada, se elaboraron recomendaciones preliminares acerca de cómo realizar entrevistas y cómo procesar la información para su modelado, con el fin de mitigar los sesgos cognitivos de los ingenieros de requisitos.

Keywords: Ingeniería de Requisitos, Entrevistas, Modelos en Lenguaje Natural, Factores Cognitivos

1 Introducción

Es bien sabido que tener un conjunto de requisitos de software de alta confiabilidad propende a un desarrollo de software exitoso [1]. Sin embargo, lograr ese punto de partida de calidad no es una tarea simple; requiere tener un proceso de construcción de requisitos bien establecido que contemple las actividades necesarias y que ellas se desarrollen siguiendo guías o heurísticas apropiadas. Como actividades centrales de este proceso están la elicitación y el modelado, que son el foco del presente trabajo.

A pesar de la existencia de abundante literatura sobre cómo construir determinados modelos en la Ingeniería de Requisitos, se ha comprobado a través de varios estudios previos [2], [3], [4], [5], que el modelo resultante está muy influenciado por quién lo elabora, encontrándose diversas debilidades tales como: conceptos nombrados incorrectamente, conceptos inexistentes, información mal comprendida, y la inducción de vocabulario por parte del elicitor. Para entender cómo se generan esas debilidades, resulta de suma relevancia revisar cómo se lleva a cabo la elicitación y cómo se procesa la información elicitada para ser registrada en modelos. Pues gran parte de la literatura en Ingeniería de Requisitos está dedicada a proveer heurísticas o guías que describen cómo modelar orientadas casi exclusivamente a la sintaxis y la semántica de modelos [19], [20], [21] [25], [32], [33] y, en cuanto a la elicitación, se presentan las cuestiones procedurales de las diversas técnicas existentes [12], [13], [14], [15], [34], dejando de lado los aspectos cognitivos y sociales inherentes a cualquier actividad desarrollada por personas y donde hay importantes interacciones.

A partir de un experimento controlado dedicado a estudiar la realización de una entrevista y su posterior modelado por varios sujetos independientes, se detectaron una variedad de defectos en los modelos, cuyos resultados cuantitativos se presentaron en [6]. Ello dio origen a un estudio más profundo que ha llevado a identificar algunas de las causas de esos defectos. A su vez, esto permitió elaborar un conjunto de recomendaciones para mitigar los problemas identificados en la entrevista y en el posterior procesamiento de la información elicitada.

En la sección siguiente se presenta la fundamentación del uso de entrevistas y su

tratamiento en la Ingeniería de Requisitos y en otras disciplinas, junto con los beneficios del modelado en lenguaje natural. En la sección 3 se describe el experimento realizado usando un caso real desde una entrevista en un ambiente de trabajo hasta la construcción de un modelo basado en la información obtenida en la entrevista. En la sección 4 se detallan las buenas prácticas que se recomiendan, presentando el problema, un ejemplo y su prevención. Finalmente, se presentan conclusiones y futuros trabajos que involucrarán la replicación del experimento y nuevos experimentos o estudios de caso, para extender y refinar las buenas prácticas que se proponen.

2 Marco Conceptual

Varios estudios de distintos autores [2], [6], [7], [8], [9] realizados en la Ingeniería de Requisitos, muestran la presencia de diversos defectos en los modelos construidos a partir de la información recolectada, lo que señala que no se ha logrado una comprensión acabada del contexto donde se pondrá en servicio el futuro sistema de software. Parte de ello se debe a sesgos cognitivos [10] de los propios ingenieros de requisitos, no contemplados en las técnicas utilizadas [4], y de los usuarios que tienen sus propias falencias (conocimiento parcial, conocimiento tácito, inconsistencias entre usuarios, dificultades de transmisión de su conocimiento) [11]. Esto indica falencias no solo al modelar, sino al elicitar y al elaborar la información obtenida para trasladarla al modelo.

2.1 Entrevistas y su Procesamiento

Las entrevistas son la técnica de elicitación más difundida en la literatura en Ingeniería de Requisitos [12], [13], [14] y en la práctica en la industria [9], [11], [15]. Además, son utilizadas en diversas disciplinas como técnica de adquisición de conocimiento o de recolección de información, tales como en la investigación cualitativa [29], [30], [36], en el diagnóstico y orientación [31], en gestión del conocimiento [35], [37], entre otras.

Las recomendaciones sobre el manejo de entrevistas en la literatura de la Ingeniería de Requisitos suelen ser livianas, ya que se centran en qué preguntar y cómo preguntar, sin dar pautas para que la pregunta sea comprensible por el interlocutor y no se introduzcan sesgos en base a presupuestos del entrevistador, entre otras cuestiones. Se privilegia el análisis del tipo de entrevista a realizar, la manera de preparar la misma, el cumplimiento en el desarrollo de la entrevista en cuanto a tomar notas, hacer resúmenes y actas [11], [16], [17], [34], y eventualmente grabar o filmar [14]. Casi no se considera importante realizar una transcripción de la entrevista [18]. Tampoco se presentan en la literatura pautas claras sobre cómo procesar la información elicitada y volcarla en modelos de manera tal que éstos reflejen la información provista por las fuentes, pues como ya se mencionó, habitualmente las recomendaciones se centran, casi exclusivamente, en describir la sintaxis y semántica del modelo.

Por otro lado, las Ciencias Sociales han perfeccionado el uso de las entrevistas, dando indicaciones sobre la necesidad de establecer un vínculo de empatía con el entrevistado, de dejar registro fehaciente de la entrevista a través de la grabación, de transcribir el audio a texto incluyendo anotaciones sobre aspectos no verbales (gestos, tonos de voz, emociones), y de cómo elaborar conclusiones basadas estrictamente en la información obtenida dejando rastros a las partes del texto involucrado [22], [23], [24].

Esta última recomendación está vinculada a lo que en Ingeniería de Requisitos se denomina *rastreabilidad de requisitos* [40] y [41], más específicamente referida a la *pre-rastreabilidad*, es decir, vincular las partes del modelo con las fuentes de información que proveyeron tales partes. La pre-rastreabilidad es una cuestión de suma relevancia, que frecuentemente no es puesta en práctica o es realizada parcialmente en la Ingeniería de Requisitos; esto ocurre por el esfuerzo y costo que conlleva, cuando no se evalúan los beneficios que provee, aunque también ocurre por desconocimiento sobre su uso y sus beneficios [42].

2.2 Modelos en Lenguaje Natural

La comunicación es una parte esencial del proceso de requisitos, tanto para que los desarrolladores se informen sobre el contexto en el que se desenvolverá el futuro sistema de software y los objetivos de ese sistema, como para que informen a los clientes y usuarios acerca de cómo planifican atender esos objetivos. Que todos los involucrados en este proceso comprendan acabadamente las consecuencias de las decisiones que se vayan tomando es una de las condiciones necesarias para el éxito tanto del proceso de requisitos como del propio sistema. Esto es especialmente cierto durante las validaciones de los modelos que se construyan y durante las negociaciones que naturalmente tienen lugar para poder definir adecuadamente los servicios que prestará el sistema.

Los involucrados en el proceso de requisitos -desarrolladores, clientes y usuarios, consultores, reguladores y otros- suelen utilizar técnicas de comunicación muy diferentes entre sí. En particular, los profesionales de la informática se suelen expresar mediante esquemas gráficos y otros lenguajes de muy difícil comprensión por los restantes involucrados. Está muy bien que así lo hagan en toda aquella documentación interna al desarrollo propiamente dicho, especialmente por la precisión de sus indicaciones. Sin embargo, en la relación con los restantes involucrados es mandatorio el uso del lenguaje natural.

El uso de modelos escritos en lenguaje natural en el proceso de requisitos, tales como glosarios, historias de usuario, escenarios y casos de uso, son el medio clásico de comunicación y participación de los involucrados [19], [25], [26], [27], pues pueden ser usados como medio de intercambio en la elicitación, en la negociación y en la validación.

Una debilidad natural de estos modelos es la ambigüedad [3], [39]. Más allá de los defectos ortográficos y sintácticos, existen otras fuentes de defectos, usualmente más difíciles de detectar como son las ambigüedades léxicas, ambigüedades sintácticas,

ambigüedades semánticas y ambigüedades pragmáticas [38]. De todas ellas, se ha observado [5] que la ambigüedad léxica, especialmente la homonimia, y la pérdida de contexto o ambigüedad pragmática suelen ser las más perniciosas. Ellas son parte de los desafíos que se abordan en este trabajo.

Asimismo, cabe notar que la ambigüedad no solo está presente en estos modelos, sino que también ocurre durante la comunicación oral, por ejemplo, en las entrevistas. Ello se pone en evidencia al estudiar en profundidad tanto las preguntas como las respuestas de los participantes, y además dichas ambigüedades acaecen independientemente del tipo de modelo que luego se construya.

3 Experimento Controlado

Con el fin de estudiar el tratamiento de la información que se obtiene en las entrevistas y su elaboración para construir modelos en lenguaje natural, se diseñó y llevó a cabo un experimento basado en la realización de una entrevista a un usuario real en un puesto de gestión en una organización privada y en construir un modelo en lenguaje natural por distintos ingenieros de requisitos usando la información elicitada.

Se adoptó de las Ciencias Sociales las prácticas de grabar la entrevista, transcribirla incluyendo anotaciones de aspectos no verbales y marcar las partes de la transcripción utilizadas en el modelo. Esto último tiende a una aproximación hacia la pre-rastreabilidad. Esta es una técnica controversial en el contexto de la heurística utilizada, ya que tiene ventajas e inconvenientes que no han sido estudiadas objetivamente aún. En las Ciencias Sociales, las marcas se utilizan para visualizar categorías de interés en forma rápida, sin necesidad de recurrir a relecturas permanentes mientras que en este experimento se utilizaron para identificar las porciones de la transcripción ya utilizadas. Marcar las porciones de la transcripción de la entrevista ofrece la gran ventaja que dirige la atención a los fragmentos no considerados de la misma. Sin embargo, ocurre que no hay una relación uno a uno entre fragmentos del texto y conceptos de interés, muy por el contrario, es notoriamente frecuente que fragmentos del texto contengan información relevante relacionada con varios conceptos. En particular, en la transcripción de la entrevista del experimento controlado, el 73% de los fragmentos útiles contiene información relevante para modelar más de un concepto. Eso hace que marcar como usada una parte del texto pueda inducir omisiones. Esto no es un imperativo categórico, podría atemperarse o reducirse a un mínimo, pero no se sabe si las ventajas superan a los inconvenientes, ni si el esfuerzo involucrado es compatible con los presuntos beneficios.

3.1 Pasos del Experimento

En la Figura 1 se presenta un esquema que describe los pasos realizados para llevar a cabo el experimento y se visualizan los productos intermedios obtenidos.

En el experimento, participaron cinco ingenieros de requisitos como modeladores,

los que contaban con experiencia similar en la construcción del modelo. De ellos, dos fueron los entrevistadores, con experiencia en realizar entrevistas. A cada modelador se le presentó el proceso de construcción del modelo y se le entregó con anticipación un documento con su definición detallada, que incluía nuevas heurísticas y otras con mayor nivel de detalle, referidas a la selección de información y elaboración, un mecanismo iterativo de modelado, un mecanismo de cortar y pegar desde texto transcrito al modelo y un criterio de parada [5], dándoles un período de lectura y consultas. Los cinco sujetos recibieron además el documento de objetivo general y alcance del sistema. De ellos, los dos entrevistadores accedieron a este documento previamente a la entrevista.

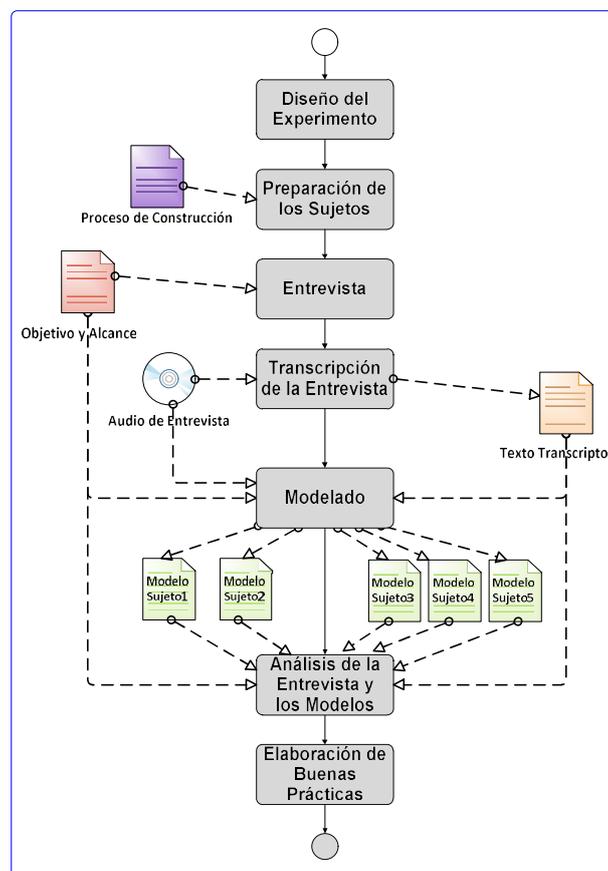


Fig. 1. Proceso del Experimento Controlado.

El análisis de la entrevista a través del texto transcripto y de los cinco modelos estuvo a cargo de dos evaluadores, que no fueron sujetos del experimento y que evaluaron de manera independiente. Estos dos evaluadores fueron quienes habían generado el documento de objetivo y alcance a partir de un contacto inicial con el cliente.

La primera parte del análisis se concentró en la calidad de los modelos y el esfuerzo involucrado en su construcción, en función de los roles cubiertos en el proceso,

obteniéndose datos cuantitativos presentados en [6]. Estos resultados permitieron observar que usar la transcripción de la entrevista mejoró la calidad de los modelos, que quienes entrevistaron y luego modelaron produjeron modelos más compactos con poca información irrelevante y, por el contrario, quienes solo modelaron produjeron modelos con exceso de información irrelevante (más del 47% de los conceptos descartados). Esto pudo llevar a que los esfuerzos de construcción de los modelos de estos sujetos fueran muy superiores a los otros.

La segunda parte del análisis se abocó a estudiar con una visión cualitativa el desarrollo de la entrevista usando la transcripción y la elaboración de la información desde la entrevista hacia el modelo según cada sujeto. Esto permitió obtener conclusiones basadas en problemas detectados que derivaron en un conjunto de buenas prácticas para la entrevista y para la elaboración de información al modelar. En este análisis cualitativo también se consideraron los resultados de experimentos anteriores en la construcción del mismo modelo para otros casos de estudio.

La comparación de los resultados entre quienes participaron en la entrevista y quienes no lo hicieron permitió evaluar la importancia de esa participación en la calidad del resultado final y la comparación de los resultados de quien realizó la corrección de la transcripción de la entrevista con los resultados de quien no lo hizo también proveyó información útil. Claramente todos los resultados obtenidos son preliminares por la escasa cantidad de participantes y la influencia de sus habilidades personales.

3.2 Validez del Experimento

El diseño del experimento permitió seleccionar adecuadamente los sujetos para cubrir los distintos roles necesarios. Posterior a la primera etapa de análisis, a quien se le había encargado la transcripción de la entrevista, se le adicionó el rol de modelador. Es decir, bajo las mismas condiciones de los otros modeladores, el responsable de la transcripción construyó una nueva versión del modelo a partir de la transcripción que él mismo había revisado y corregido para los restantes sujetos, contando también con el documento de objetivo y alcance, y habiendo recibido las mismas instrucciones. Este quinto modelo fue objeto de estudio en la segunda etapa de análisis.

Todos los sujetos fueron seleccionados con similar nivel de experiencia en la construcción del modelo, ninguno lo había construido previamente con el nuevo proceso. El entrevistado fue un usuario trabajando en una organización privada, sin ninguna relación con los evaluadores ni con los sujetos del experimento. Los evaluadores que analizaron la entrevista y los modelos fueron quienes diseñaron el nuevo proceso de construcción.

Debe tenerse presente que las conclusiones obtenidas no pueden generalizarse porque se trata de sujetos expertos, basado en un único caso, aunque cabe notar que en el análisis cualitativo se estudió adicionalmente una amplia gama de otros casos reales de diversa complejidad elaborados en anteriores experimentos.

4 Buenas Prácticas

A continuación, se describen soluciones a los problemas más importantes detectados en el experimento, o detectados revisando retrospectivamente casos previos.

Cada buena práctica está expresada en términos de un problema, su origen y las prevenciones a tomar para evitar o mitigar el problema, acompañada de ejemplos. Estos problemas se han agrupado en aquellos que surgen de la realización de la entrevista (subsección 4.1) y aquellos que se producen al modelar (subsección 4.2). Para el primer caso se presentan prevenciones a considerar tanto durante la entrevista como prevenciones al procesar la información elicitada para modelar.

Para facilitar la comprensión de cada uno de los problemas, se ha recurrido a ejemplos extraídos tanto del experimento como de casos de estudio realizados con anterioridad, de manera de hacer más visible la dificultad.

4.1 Problemas en la Entrevista

Problema Nro. 1. El entrevistador induce en el entrevistado el uso de una determinada palabra o frase.

Ejemplo: El entrevistador pregunta: “¿Los *retrasos* en la recepción de la materia prima requieren alguna acción del *supervisor*?”, pero en ese contexto se usan las frases *demoras del proveedor* y *encargado de stock*. En esa situación el entrevistado contesta utilizando las palabras *retrasos* y *supervisor* en forma inadvertida o para mantener el hilo de la conversación. Posteriormente, el entrevistado utiliza la frase *encargado de stock*. Esto puede originar la suposición que existen dos actores con diferentes responsabilidades o que se cree un sinónimo inexistente. Cuando en otro momento de la entrevista el entrevistado utiliza la frase *demoras del proveedor*, puede no asociarse a lo descripto con anterioridad sobre los *retrasos*, introducidos por el entrevistador.

Origen del Problema: Esto ocurre frecuentemente y tiene sus raíces en las experiencias previas del entrevistador referidas a contextos similares.

Prevenciones:

a) Durante la entrevista esforzarse en minimizar el uso de palabras o frases no escuchadas con anterioridad. Escuchar qué términos específicos maneja el entrevistado y esforzarse en utilizarlos, evitando incorporar sinónimos. También se debe evitar, tanto como sea posible la realización de preguntas específicas, principalmente preguntas sobre tópicos no mencionados aún por el entrevistado. Prefiriendo en todos los casos los estímulos generales, tales como: “Acláreme más este punto” o “¿Y luego que ocurre?” o “¿Eso es todo?”.

b) Al procesar la entrevista para modelar, prestar especial atención a aquellas palabras o frases cuya primera aparición fue en un dicho del entrevistador. Verificar si el entrevistado o el entrevistador siguen usando esa palabra o frase posteriormente. Verificar si el entrevistado utiliza posteriormente otra palabra o frase, en tal caso descartar

el término utilizado por el entrevistador.

Problema Nro. 2. El entrevistador usa una determinada palabra o frase, que el entrevistado interpreta con otro significado.

Ejemplo: El entrevistador pregunta: “Si el dispositivo de la plataforma está en ON y el ómnibus se retira ¿Qué se debe hacer con él?”. El entrevistado interpreta que se está preguntando por el *sensor de ómnibus* y no por el *indicador de plataforma ocupada* por lo que contesta: “Nada, la *indicación* desaparece en forma automática”. El entrevistador interpreta que ese *indicador* no es atendido por el sistema de software.

Origen del Problema: Esto ocurre por falta de atención del entrevistador sobre los dichos del entrevistado y la falta de claridad en sus preguntas.

Prevencciones:

a) Prácticamente la misma que la prevención a) del problema anterior. Asegurarse que la respuesta del entrevistado refleje el tipo de información esperada por el entrevistador, ante la menor duda, reformular la pregunta o preguntar para esclarecer sobre los dichos del entrevistado.

b) Al procesar la entrevista para modelar, prestar especial atención a aquellas palabras o frases cuya primera aparición fue en un dicho del entrevistador. Verificar si la respuesta del entrevistado está en relación con la pregunta, si usa otras palabras o frases en su respuesta, o si le da otro significado a la palabra o frase del entrevistador. No introducir en el modelo palabras o frases solo mencionadas por el entrevistador.

Problema Nro. 3. El entrevistado se controla en el uso de la jerga específica de su actividad y procura utilizar sinónimos de uso más general para facilitar la comprensión por parte del entrevistador.

Ejemplo: El entrevistado dice: “Una vez *separado del agua inyectada para la recuperación secundaria*, el *petróleo crudo* se lava y se envía a otros *tanques para ser separado del agua agregada*”. En el uso corriente el entrevistado hubiera dicho: “Una vez *decantado*, el crudo se *desmineraliza* y se envía a los tanques de *pre-despacho*”.

Origen del Problema: Esto ocurre en situaciones donde el contexto es muy especializado y existe un profuso uso de palabras o frases con un significado preciso.

Prevencciones:

a) Durante la entrevista prestar especial atención a titubeos en el uso de algunos términos o la existencia de una incoherencia entre la complejidad de las actividades descriptas y los términos utilizados para referirse a ellas. Consultar cualquier duda al respecto. Podrían hacerse preguntas tales como “¿Cómo llaman a esto?” o “¿Tienen nombres especiales?”.

b) Al procesar la entrevista para modelar, observar si el entrevistado usó consistentemente las palabras o frases, o si recurrió a muchos sinónimos de uso general para referirse a lo mismo, es decir, términos de uso común en cualquier contexto sin incorporar términos específicos. También evaluar la posibilidad de existencia de incoherencias entre las palabras utilizadas y la complejidad de las actividades. En este caso, es necesario realizar otra elicitación que permita adquirir ese vocabulario específico.

Si se detecta que el entrevistado utilizó términos específicos junto con sinónimos de uso más general, no considerar a éstos últimos como sinónimos reales, aunque es conveniente consultarlo.

Problema Nro. 4. El entrevistador se esfuerza en asociar lo que está recibiendo del entrevistado a su propio conocimiento.

Ejemplo: El entrevistado dice: “Una vez aprobada la orden de compra se emite la factura en forma electrónica, pero una copia en papel se envía a expedición”. Entonces el entrevistador pregunta: “¿En expedición arman los Pallets para su *entrega a los clientes*?”. Y el entrevistado contesta: “Efectivamente, así se hace”. El entrevistador supone que los *clientes pasarán a retirar los productos adquiridos*, cuando en la realidad, *se emiten remitos que se entregan a repartidores propios y de terceros que llevan los productos a los clientes*.

Origen del Problema: Relacionado con las experiencias previas del entrevistador.

Prevencciones:

a) Durante la entrevista evitar hacer asociaciones entre experiencias anteriores y lo que se está escuchando u observando en la entrevista. Evitar preguntas que intenten confirmar esas asociaciones; recurrir a preguntas, tales como: “¿Qué ocurre en ese lugar?” o “¿Qué ocurre posteriormente?” o “¿Algo más ocurre?”.

b) Al procesar la entrevista para modelar, estudiar cuidadosamente las preguntas específicas del entrevistador y evaluar con cuidado las respuestas del entrevistado, evitando introducir en el modelo información proveniente de esas preguntas, cuyas respuestas no amplían lo preguntado. Prestar especial atención a cambios de atención por parte del entrevistador que pudieran inducir a cierto tipo de respuestas.

Problema Nro. 5. El entrevistador inadvertidamente pregunta directa o indirectamente sobre tópicos fuera del alcance establecido para el sistema.

Ejemplo: Después de relatar el entrevistado sobre cómo se relaciona la clínica odontológica con los laboratorios que les proveen los materiales auxiliares y cómo se les liquidan los pagos, el entrevistador dice: “Es complicado. Y con los *odontólogos contratados* ¿cómo se trata este tema?”. El entrevistado empieza a relatar sobre el *pago de servicios a los odontólogos*, que era un tema excluido del alcance del sistema.

Origen del Problema: Esto está relacionado con la habilidad y el entrenamiento del entrevistador, y suele ser motivado por su propia curiosidad y por la falta de concentración en la actividad.

Prevencciones:

a) Durante la entrevista evitar preguntas que traten temas que no estaban previstos en el objetivo de la entrevista ni en el objetivo y alcance del sistema. Ser cuidadoso con los temas de las preguntas, que sean orientadas al tipo de información que se busca obtener. Tener presente en todo momento el objetivo y alcance del sistema, además del objetivo establecido para la entrevista.

b) Al procesar la entrevista para modelar, asegurarse que toda palabra o frase que

se use en el modelo o información que se le incorpore esté dentro del alcance establecido. Tener presente en todo momento el objetivo y alcance del sistema.

Problema Nro. 6. El entrevistado relata espontáneamente temas que están fuera del alcance establecido para el sistema, o que no son parte del objetivo de la entrevista.

Ejemplo: Luego de relatar con qué proveedores trabaja, el entrevistado dice: “Con los proveedores tenemos diferentes *acuerdos de pago* ..., el sistema debe tomar el precio de cada insumo recibido según la factura enviada por el proveedor y emitir la *orden de pago*”. Inicialmente, se había establecido un objetivo general con un alcance donde el *pago a proveedores* quedaba fuera del sistema, solo se incluiría las entradas a stock por compras.

Origen del Problema: Esto suele ocurrir cuando el entrevistado no está directamente relacionado con el área de la organización que requiere el nuevo sistema de software o cuando tiene necesidades adicionales y desea manifestarlas.

Previsiones:

a) Durante la entrevista procurar informar al entrevistado que ese tema no se tratará en esa entrevista. Recurrir a expresiones tales como: “Lamentablemente ese tema no lo trataremos hoy” o “Primero necesitaríamos conocer sobre ...” o “Por ahora, ese tema quedó afuera”.

b) Prácticamente la misma que la prevención b) del problema anterior.

Problema Nro. 7. El entrevistado utiliza una palabra o frase con dos significados diferentes pero cercanos.

Ejemplo: El entrevistado dice: “Para reservar una habitación los *pasajeros* deben pagar un anticipo, que depende de la fecha. En temporada alta, fines de semana o feriados es del treinta por ciento del total, mientras que en el resto de los días se debe anticipar sólo el diez por ciento. Todas las reservas se deben hacer con dos días de anticipación. A los *pasajeros frecuentes* no les exigimos anticipo y pueden hacer sus reservas en cualquier momento. Los *pasajeros* que no se presentan el día acordado, pierden su reserva y el anticipo. Ya alojados, los *pasajeros* podrán acortar su estadía libremente o extender la misma siempre que tengamos disponibilidad”. La primera ocurrencia de la palabra *pasajeros* se refiere sólo a los pasajeros no frecuentes, la segunda forma parte de la frase *pasajero frecuente*. La tercera mención fabrica una ambigüedad porque no se sabe si los pasajeros frecuentes también pierden su reserva. Finalmente, la cuarta se refiere a todos los pasajeros. Acá el entrevistado utiliza la misma palabra *pasajero* para referirse a personas distintas, por un lado, a todos los pasajeros y por el otro a los no frecuentes.

Origen del Problema: Esto ocurre cuando el entrevistado utiliza una palabra o frase acortada con el mismo significado que una frase más larga que la contiene.

Previsiones:

a) Durante la entrevista prestar atención a las frases que se refieren a objetos y actores que estén conformadas por un sustantivo y un adjetivo calificativo. Si además el sustantivo es usado en forma aislada, procurar detectar ambigüedades en el discurso y

pedir aclaraciones en caso de encontrarlas.

b) Al procesar la entrevista prestar atención a la misma situación tratando de deducir del contexto cual es la interpretación correcta. Si eso no es posible, registrar la ambigüedad para tratar de resolverla en otra oportunidad.

4.2 Problemas en el Procesamiento de la Información Elicitada

Problema Nro. 1. El modelador se esfuerza en asociar lo que surge de la entrevista a conceptos de su propio conocimiento, incorporándolos al modelo. Este problema es otra manifestación del Problema Nro. 4 de la Entrevista, pero en otro momento y posiblemente por parte de otra persona.

Ejemplo: En la transcripción de la entrevista se tiene: “Entrevistado: Si un producto elaborado no es despachado a tiempo y se encuentra próximo a su fecha de vencimiento, no se entrega a los distribuidores, sino que se envía directamente a instituciones de caridad”. Al procesar la entrevista el modelador asocia esta situación a una muy similar en la que tuvo alguna intervención en la que utilizaban la frase *descarte preventivo*, interpretando que esa es la forma correcta de referirse a esta actividad, pero para referirse a un tratamiento diferente del mencionado por el entrevistado.

Origen del Problema: Relacionado con las experiencias previas del modelador.

Prevención: Asegurarse que toda palabra o frase que se incorpore al modelo haya sido utilizada por el entrevistado y que lo haya hecho con ese significado.

Problema Nro. 2. El modelador agrega información basado en su conocimiento previo o haciendo suposiciones respecto a información omitida. Este problema es una tercera manifestación del Problema Nro. 4 de la Entrevista, pero en otro momento y posiblemente por parte de otra persona.

Ejemplo: En la transcripción de la entrevista se tiene: “Entrevistado: Las planillas de pedidos de materia prima no se envían al proveedor, vos le mandas la *orden de compra*”. El modelador asume que el *jefe de compras* se encarga del envío de las órdenes, aunque no fue dicho expresamente por el entrevistado, es decir, incorpora quién es el responsable de la actividad. Además, en la transcripción se tiene: “Entrevistado: Para encontrar una *orden de compra* buscamos en las planillas por *su número* y, a veces, por *fecha de pedido*”. El modelador describe que la búsqueda de una *orden de compra* se puede realizar también por *fecha esperada de entrega*, según lo había experimentado en un contexto similar.

Origen del Problema: Ocurre frecuentemente y tiene sus raíces en las experiencias previas del modelador en contextos similares, principalmente cuando hay omisiones.

Prevención: Asegurarse que toda información que se incorpore al modelo haya sido mencionada por el entrevistado, y que no sea de su propia creación.

Problema Nro. 3. El modelador ignora una palabra o frase con un significado específico, por lo que omite modelar un aspecto necesario.

Ejemplo: En la transcripción de la entrevista se tiene: “Entrevistado: Una vez finalizado el sorteo, el *escribano labra un acta* con el resultado del mismo, indicando expresamente los beneficiados y quien es el *responsable de realizar las comunicaciones fehacientes*”. Al procesar la entrevista el modelador deduce que “... todos los *escribanos labran actas*”, por lo que esta acción no tiene ninguna peculiaridad en este contexto. Como consecuencia no sólo se omite el actor *escribano*, sino que se introduce el riesgo de no percibir el actor *responsable de las comunicaciones fehacientes*, el objeto *comunicación fehaciente* y tal vez la actividad *realizar comunicación fehaciente*.

Origen del Problema: Ocurre cuando en la jerga utilizada en el contexto han adaptado palabras o frases de uso corriente dándoles un significado particular y relevante.

Prevención: Prestar especial atención a si una palabra o frase tiene un significado específico en el contexto, examinando si ese concepto tiene alguna consecuencia, impacto o relación con otros conceptos ya identificados como pertenecientes a ese contexto. En tal caso, incorporar esa palabra o frase en el modelo junto con las consecuencias o relaciones identificadas. Si ello no puede establecerse, registrarlo para tratar de resolverlo en otra oportunidad.

Problema Nro. 4. El modelador agrega una palabra o frase proveniente del entrevistado, pero que no posee un significado específico, por lo que no es necesario modelar.

Ejemplo: En la transcripción de la entrevista se tiene: “Entrevistado: Los *resultados del análisis* se podrán entregar en forma *impresa o electrónica* según lo que desee el interesado”. Al procesar la entrevista el modelador deduce que existen tres frases que deben incluirse en el modelo: *resultados del análisis*, *resultados del análisis electrónico* y *resultados del análisis impreso*. También deduce que existe una relación jerárquica entre estas tres frases. Además, existe el riesgo que se agregue información no provista en la entrevista para completar el modelo debido a esas relaciones creadas o a las frases incorporadas.

Origen del Problema: Suele ocurrir cuando el modelador interpreta que una palabra o frase tiene relevancia. Ello puede surgir cuando el entrevistado utiliza palabras o frases con baja frecuencia de uso en el habla corriente para describir cuestiones secundarias.

Prevención: Asegurarse que toda palabra o frase que se incorpore al modelo tenga un significado específico en el contexto y que no sea de su propia creación o basada en experiencias anteriores. Estudiar en detalle el uso de esa palabra o frase y si tiene algún impacto en ese contexto o se relaciona con otros conceptos identificados. Si ello no puede establecerse, no incluir dicha palabra o frase en el modelo y registrar esto para tratar de resolverlo en otra oportunidad.

Problema Nro. 5. El modelador introduce definiciones en el modelo utilizando sus propios términos y no los utilizados por el entrevistado, lo que puede derivar en modificar significados degradando la calidad del modelo.

Ejemplo: En la transcripción de la entrevista se tiene: “Entrevistado: Las cajas con *accesorios* se empaquetan en forma especial respetando la naturaleza de cada componente. Es decir, se empaqueta por un lado la caja, por el otro los *separadores* y final-

mente los *divisores*”. Al procesar la entrevista el modelador indica que “... los *separadores* y *divisores* no se empaquetan junto con sus respectivas cajas ...”, agregando que “... los *separadores* y los *divisores* se empaquetan en forma *separada* ...”. Así no sólo se omitió el uso de la palabra *accesorios* que es relevante, eliminando además una relación entre esos tres términos, sino que además se agregó una ambigüedad entre el objeto *separadores* y la palabra *separada* que no es relevante en este contexto.

Origen del Problema: Ocurre cuando el modelador no es lo suficientemente cuidadoso en respetar el vocabulario utilizado por el entrevistado, y posiblemente considere que reescribir con un vocabulario de uso corriente facilita la comprensión.

Prevención: Durante la descripción relacionada con una palabra o frase utilizar al máximo la técnica de cortar y pegar trasladando literalmente la porción de texto de la entrevista al modelo, realizando luego las modificaciones mínimas para adaptar los tiempos verbales y estructura del texto evitando tanto como sea posible el agregado de nuevas palabras.

Problema Nro. 6. El modelador agrega en el modelo una palabra o frase que está mencionada por el entrevistador, pero no usada por el entrevistado.

Ejemplo: En la transcripción de la entrevista se tiene: “Entrevistador: El tema de turnos, *prestaciones de obra social*, la historia clínica, el registro de lo que se le hace a cada paciente ...”. En la transcripción, el entrevistado durante sus relatos no menciona el término *prestaciones de obra social* sino *prestaciones*, es decir, ese concepto no está directamente asociado a la *obra social*. Al no advertir esa intrusión, el modelador traslada la frase del entrevistador y posiblemente introduzca alguna relación inexistente con otros conceptos modelados.

Origen del Problema: Esto está relacionado con la habilidad y el entrenamiento del modelador cuando no advierte que el entrevistador hace preguntas específicas con su propio vocabulario, pero no tiene correlato con el vocabulario usado por el entrevistado en ese contexto.

Prevención: Asegurarse que toda palabra o frase que se incorpore al modelo sea utilizada por el entrevistado.

5 Conclusiones

Los problemas detectados en el experimento, cuya existencia había sido observada en otros estudios realizados con anterioridad [2], [3], [4], [5], [7], giran alrededor de las personas involucradas en la actividad de elicitar información mediante entrevistas: el entrevistado, el entrevistador y el modelador. Obviamente no es posible pretender modificar el desempeño del entrevistado más allá de realizar algunas sugerencias acerca de ceñirse, dentro de lo posible, al objetivo de la entrevista y a no restringirse en el uso de la jerga propia del contexto que está describiendo.

Todas las demás recomendaciones y heurísticas a seguir aplican a las personas per-

tenecientes al equipo de Ingeniería de Requisitos. Es claro que también es difícil que los ingenieros de requisitos puedan dejar de lado las asociaciones que naturalmente realizan debido a sus experiencias previas. El mecanismo introducido en el proceso de construcción del modelo respecto a usar la transcripción de la entrevista y “copiar y pegar” desde dicho texto al modelo, colabora en tal sentido. Sin embargo, con pautas recordatorias y con una preparación adecuada tanto para entrevistar como para modelar, es posible reducir este sesgo en el tratamiento de la información elicitada y elaborada.

En ese sentido es muy importante entender que todas las conclusiones que se refieren a los problemas relacionados con el procesamiento de la entrevista aplican también, casi sin cambios, al procesamiento de los resultados de varias otras técnicas de elicitación como es el caso de la lectura de documentos, la observación y la revisión de sistemas preexistentes, entre otras.

De los problemas reportados, nueve son adjudicables a los ingenieros de requisitos, ya sea por dificultades asociadas a sus habilidades, cuidado de atención o influencia de sus experiencias previas, mientras que los cuatro restantes dependen más del contexto en sí propio. En alguna manera esto es una buena noticia porque es de esperar que el respeto de las buenas prácticas permita obtener modelos de mejor calidad.

Es notoria la escasez de buenas prácticas que consideren los factores humanos intervinientes en la Ingeniería de Requisitos. En [28] se describen buenas prácticas para elaborar un modelo en lenguaje natural centradas exclusivamente en cuestiones relativas al modelo. Lo mismo ocurre con las heurísticas presentadas en [25] y en [19].

En resumen, se pueden esbozar algunas conclusiones iniciales:

- Las deficiencias detectadas giran en torno a las personas involucradas.
- El modelo resultante depende de si el modelador participó o no de la entrevista, como también si utilizó el audio o la transcripción.
- Es difícil que los ingenieros de requisitos dejen de lado asociaciones debido a experiencias previas.

Una cuestión que ha quedado pendiente de estudio es los beneficios de marcar el texto de la transcripción a medida que se usa en el modelo, los cuidados que debieran tenerse al llevar adelante esta práctica y el esfuerzo involucrado.

En los próximos trabajos se hará un nuevo experimento incluyendo las buenas prácticas presentadas. Asimismo, se intentará confirmar y ampliar los resultados en el experimento controlado anterior en cuanto a la incidencia en la calidad del modelo cuando el mismo sujeto participa en más de una actividad: en la elicitación, la transcripción y/o el modelado. Otro factor que se estudiará es el nivel de experiencia de los entrevistadores y de los modeladores.

Referencias

1. Nuseibeh, B., Easterbrook, S.: Requirements Engineering: A Roadmap. ICSE - Future of SE Track 2000, Limerick (2000) 35-46
2. Hadad, G.D.S., Litvak, C.S., Doorn, J.H.: Problemas y Soluciones en la Completitud de Modelos en Lenguaje Natural. II Congreso Argentino de Ingeniería, Tucumán, T366 (2014)
3. Sebastián, A., Hadad, G.D., Robledo, E.: Inspección centrada en Omisiones y Ambigüedades de un Modelo Léxico. 20th Workshop on Requirements Engineering, Buenos Aires (2017)
4. Doorn, J.H., Hadad, G.D.S., Elizalde, M., García, A., Carnero, L.: Críticas Cognitivas a Heurísticas Orientadas a Modelos. 22nd Workshop on Requirements Engineering, Recife (2019)
5. Elizalde, M.C., Hadad, G.D.S., Doorn, J.H.: Incorporación de heurísticas lingüístico-cognitivas en el Proceso de Requisitos. Memoria de Trabajos de 9º Congreso Nacional de Ingeniería Informática / Sistemas de Información, Mendoza (2021) 312-320
6. Doorn, J.H., Hadad, G.D.S., Elizalde, M.C., Ridaio, M.N., Casafuz, D., Sebastián, A., Riera, G.A.: Impacto del Proceso de las Entrevistas en la Calidad de los Modelos. 10º Congreso Nacional de Ingeniería Informática / Sistemas de Información, Entre Ríos (2022)
7. Ridaio, M.N., Doorn, J.H.: Estimación de Completitud en Modelos de Requisitos Basados en Lenguaje Natural. 11th Workshop on Requirements Engineering, Río de Janeiro (2006) 146-157
8. Martínez, S.N., Oliveros, A., Zuñiga, J.A., Corbo, S., Forradelas, P.: Aprendizaje de la elicitación y especificación de requerimientos. XX Congreso Argentino de Ciencias de la Computación, Buenos Aires (2014)
9. Bano, M., Zowghi, D., Ferrari, A., Spoletini, P., Donati, B.: Teaching requirements elicitation interviews: an empirical study of learning from mistakes. Requirements Engineering Journal, Vol. 24. Springer-Verlag (2019) 259-289
10. Sternberg, R.J., Sternberg, K.: Cognitive Psychology. 6ta. edición, CENGAGE Learning, Boston (2009) 185-227
11. Ferrari, A., Spoletini, P., Gnesi, S.: Ambiguity and tacit knowledge in requirements elicitation interviews. Requirements Engineering Journal, Vol. 21, Nro. 3, Springer (2016) 333-353
12. Coulin, C., Zowghi, D.: Requirements Elicitation for Complex Systems: Theory and Practice. En: Maté, J.L., Silva, A. (eds.): Requirements Engineering for Sociotechnical Systems. Information Science Publishing, Hershey (2005) 37-52
13. Carrizo, D., Dieste, O., Juristo, N.: Systematizing requirements elicitation technique selection. Information and Software Technology, Vol. 56, Nro. 6, (2014)
14. Wieggers, K., Beatty, J.: Software Requirements. 3º edición, Microsoft Press (2013)
15. Oliveros, A., Antonelli, L.: Técnicas de elicitación de requerimientos. XXI Congreso Argentino de Ciencias de la Computación, Junín (2015)
16. Alexander, I.F., Beus-Dukic, L.: Discovering Requirements: How to Specify Products and Services. John Wiley & Sons, Chichester (2009) 261-274
17. Ferrari, A., Spoletini, P., Bano, M., Zowghi, D.: SaPeer and ReverseSaPeer: teaching requirements elicitation interviews with role-playing and role reversal. Requirements Engineering Journal, Vol. 25, Nro. 4, Springer-Verlag (2020) 417-438
18. Thew, S., Sutcliffe, A.: Value-based requirements engineering: method and experience. Requirements Engineering Journal, Vol. 22, Nro. 4, Springer-Verlag (2018) 443-464
19. Seyff, N., Maiden, N., Karlsen, K., Lockerbie, J., Grünbacher, P., Graf, F., Ncube, C.: Exploring how to use scenarios to discover requirements. Requirements Engineering Journal, Vol. 14, Nro. 2. Springer-Verlag (2009) 91-111
20. Yu, E., Giorgini, P., Maiden, N. et al. (eds.): Social Modeling for Requirements Engineering. MIT Press (2011)
21. Jacobson, I., Spence, I., Bittner, K.: Use Case 2.0 The Guide to Succeeding with Use Cases. Ivar Jacobson International (2011)
22. Roulston, K.: Reflective interviewing: A guide to theory and practice. Sage Publications,

- Londres (2010)
23. Mann, S.: *The Research Interview Reflective Practice and Reflexivity in Research Processes*. Palgrave Macmillan, Basingstoke (2016) 199-211
 24. Buriro, A.G., Awan, J., Lanjwani, A.: *Interview: A Research Instrument for Social Science Researchers*. Intl Journal of Social Sciences, Humanities and Education, Vol. 1 (2017) 1-14
 25. Leite, J.C.S.P., Doorn, J.H., Kaplan, G.N., Hadad, G.D.S., Ridaio, M.N.: *Defining System Context using Scenarios*. En: Leite, J.C.S.P., Doorn, J.H. (eds.): *Perspectives on Software Requirements*. EE.UU.: Kluwer Academic Publishers (2004) 169-199
 26. Antonelli, L., Rossi, G., Leite, J.C.S.P., Oliveros, A.: *Deriving requirements specifications from the application domain language captured by Language Extended Lexicon*. 15th Workshop on Requirements Engineering, Buenos Aires (2012)
 27. Mighetti, J.P., Hadad, G.D.S.: *A Requirements Engineering Process Adapted to Global Software Development*. CLEI Electronic Journal, Vol. 19, Nro. 3, paper 7 (2016)
 28. Antonelli, L., Rossi, G., Leite, J.C.S.P., Oliveros, A.: *Buenas Prácticas en la Especificación del Dominio de una Aplicación*. 16th Workshop on Requirements Engineering, Montevideo (2013) 319-332
 29. Díaz-Bravo, L., Torruco-García, U., Martínez-Hernández, M., Varela-Ruiz, M.: *La entrevista, recurso flexible y dinámico*. Investigación en educación médica, Vol. 2, Nro. 7, (2013) 162-167.
 30. Vargas-Jiménez, I.: *La entrevista en la investigación cualitativa: nuevas tendencias y retos*. Revista Electrónica Calidad en la Educación Superior, Vol. 3, Nro. 1, (2012) 119-139.
 31. Araújo, M.F.: *Estratégias de diagnóstico e avaliação psicológica*. Revista Psicologia: Teoria e Prática, Vol. 9, Nro. 2, (2007) 126-141.
 32. Barcelos, L., Penteadó, R.: *Elaboration of Software Requirements Documents by Means of Reuse Patterns*. Journal of Software Engineering Research and Development, Vol. 5, Nro. 3, (2017)
 33. Antonelli, L., Leite, J.C.S.P., Oliveros, A., Rossi, G.: *Specification Cases: a Lightweight Approach based on Natural Language*. 24th Workshop on Requirements Engineering (WER21), Brasilia, Brasil (2021)
 34. Dick, J., Hull, E., Jackson, K.: *Requirements Engineering*, Springer-Verlag London Limited, 4^o edición (cap. 5) (2017)
 35. Norese, M.F., Salassa, F.: *A System to Analyse Interviews and Structure the Acquired Knowledge Elements*. En: *Bridging the Socio-technical Gap in Decision Support Systems*, Vol. 212, IOS Press (2010) 461-472.
 36. Patel, T., Thakkar, S.: *Knowledge models, current knowledge acquisition techniques and developments*. Oriental Journal of Computer Science & Technology, Vol. 6, Nro. 4, (2013) 467-472.
 37. Peet, M. R., Walsh, K., Sober, R., Rawak, C. S.: *Generative Knowledge Interviewing: A method for knowledge transfer and talent management at the University of Michigan*. International Journal of Educational Advancement, Vol. 10, Nro. 2, (2010) 71-85.
 38. Berry, D.M., Kamsties, E.: *Ambiguity in Requirements Specification*. En: Leite, J.C.S.P., Doorn, J.H. (eds.): *Perspectives on Software Requirements*. EE.UU.: Kluwer Academic Publishers. (2004) 7-44
 39. Kaplan, G.N., Doorn, J.H., Hadad, G.D.S.: *Ambigüedad léxica en los modelos de requisitos en lenguaje natural*. XII Workshop de Investigadores en Ciencias de la Computación, El Calafate, (2010) 440-444
 40. Pinheiro, F.A.C.: *Requirements Traceability*. En: Leite, J., Doorn, J.H. (eds.): *Perspectives on Software Requirements*. EE.UU.: Kluwer Academic Publishers. (2004) 91-113
 41. Cleland-Huang, J., Gotel, O., Zisman, A., (eds.): *Software and Systems Traceability*. Springer, Londres (2012)
 42. Kannenberg, A., Saiedian, H.: *Why Software Requirements Traceability Remains a Challenge*. The Journal of Defense Software Engineering, Nro. 22, (2009) 14-19