

# Un abordaje al problema de completitud en requisitos de software

Claudia S. Litvak<sup>1,2</sup>, Graciela D. S. Hadad<sup>1,2</sup>, Jorge H. Doorn<sup>2,3</sup>

<sup>1</sup>Facultad de Tecnología Informática, Universidad de Belgrano, Argentina

<sup>2</sup>DIIT, Universidad Nacional de La Matanza, Argentina

<sup>3</sup>Fac. Ciencias Exactas, Universidad Nacional del Centro de la Provincia de Buenos Aires  
{claudia.litvak,graciela.hadad}@comunidad.ub.edu.ar, jdoorn@exa.unicen.edu.ar

**Abstract.** La completitud es uno de los temas imposibles o casi imposibles de ponderar en la Ingeniería de Software en general y en la Ingeniería de Requisitos en particular. Se han hecho algunos intentos de estimar la completitud en los casos de prueba en inspecciones de sistemas mediante el uso de técnicas predictivas. Estas técnicas permiten estimar el grado de completitud alcanzado. Entre ellas se encuentra Detection Profile Method, que fue con anterioridad aplicada a un modelo de requisitos escrito en lenguaje natural. Los resultados fueron muy promisorios, pese a basarse sólo en aspectos formales y cuantitativos del modelo. Una revisión ulterior de estos resultados ha permitido detectar que existen factores cualitativos, como distintas personas observan distintos problemas o partes del mismo problema, que podrían haber distorsionado los resultados alcanzados. En el presente artículo, se reportan los resultados de reanalizar el problema incorporando algunos elementos semánticos a las estrategias involucradas.

**Keywords:** Ingeniería de Requisitos, Modelado de Requisitos, Completitud de Modelos, Método de Captura y Recaptura.

## 1 Introducción

Diversos estudios, realizados a lo largo de las últimas décadas, referidos a fracasos en proyectos de software concluyen que los requisitos son una de las fuentes principales de problemas en la mayoría de dichos proyectos: requisitos inadecuados, cambios en los requisitos durante el ciclo de vida, requisitos no bien comprendidos y requisitos incompletos. Algunos de estos estudios son realizados por The Standish Group (<http://blog.standishgroup.com/>) con sus reportes CHAOS, por U.S. Government Accountability Office (GAO) (<http://www.gao.gov/>) y por muchos investigadores en el área de la ingeniería de software [1] [2] [3]. En la mayoría de estos casos, la incompletitud es uno de los principales problemas que se afronta en la fase de producción de requisitos [4] [5] [6] y que inevitablemente se arrastra a las siguientes fases del proceso de desarrollo de software.

## **1.1 Completitud en la Ingeniería de Requisitos**

El problema de la incompletitud reside básicamente en la dificultad en establecer si se ha elicitado y modelado toda la información requerida para desarrollar un sistema de software que cubra las expectativas y necesidades de los clientes y usuarios. Es decir, este problema se vincula directamente con un tipo específico de defecto: las omisiones. En un estudio realizado por Basili y Weiss [7] se presentó que el 31% de los errores en los requisitos provenían de omisiones.

Kotonya y Sommerville [4] entienden que la completitud significa que ningún servicio o restricción necesarios han sido omitidos. Loucopoulos y Karakostas [5] amplían este concepto cuando mencionan que un modelo de requisitos está completo cuando no se omite información esencial acerca del dominio de la aplicación. Firesmith [6] distingue entre la incompletitud del modelo de requisitos y la incompletitud individual de los requisitos. El primero se refiere a la omisión total de uno o varios requisitos, mientras que el segundo se refiere a la falta de información necesaria para que el requisito no sea ambiguo y se pueda implementar sin requerir información adicional. Se considera que ambos tipos de incompletitud deberían ser tratados con el mismo nivel de importancia.

Dado que no es esperable alcanzar la completitud de un modelo de requisitos debido a la complejidad del mundo real [8], una alternativa es poder establecer cómo decidir que se ha logrado una elicitación y un modelado suficiente para construir un software de calidad, concepto conocido como “reglas de parada” [9]. Pero, establecer estas reglas no siempre es viable o fácil de llevar a la práctica. Otra alternativa puede hallarse estableciendo una definición bien precisa de las guías y/o heurísticas para elicitar y modelar requisitos. Las verificaciones y validaciones ayudan a mitigar el problema de la incompletitud [4] [5]. El uso de técnicas de elicitación apropiadas [10] permite lograr una adquisición de conocimiento más acabada sobre el dominio de la aplicación. Pero estas técnicas no permiten eliminar el problema ni tampoco estimar el grado de completitud alcanzado.

## **1.2 Motivación del Caso de Estudio**

Los autores de este artículo han observado que, al construir los modelos de requisitos de innumerables casos, en calidad de ingenieros de requisitos y supervisores de alumnos de grado y postgrado, existe una importante duda por parte de quien elicitó, acerca de los alcances del problema a estudiar, especialmente durante las primeras etapas. Al inicio de un proyecto sólo se dispone de un enunciado usualmente muy breve acerca del objetivo del sistema y se debe proceder a identificar las fuentes de información, a definir el método de elicitación y a construir los primeros artefactos del proceso de requisitos. Todas estas actividades están fuertemente relacionadas con el alcance del problema, el cual aún no se conoce. Uno de los primeros artefactos que se suele construir es un glosario con el vocabulario del dominio del problema. La consecuencia práctica de esta incertidumbre es que al final de la creación del glosario dos personas diferentes han construido el mismo artefacto con una cantidad diferente de términos, pero esto no es todo, ninguno de los glosarios es un sub-conjunto propio del otro, ya que, lo que habitualmente ocurre es que hay una intersección importante

pero una diferencia simétrica también importante. Se considera que en el área de ingeniería de requisitos este fenómeno debe haber sido observado, pero no ha sido lo suficientemente estudiado.

Retomando el tema de la completitud, ¿cuál es el verdadero glosario del dominio del problema? Se podría hipotetizar que la respuesta es la unión de ambos o algún otro conjunto, pero definitivamente no se sabe cuál ni cómo construirlo. Las consideraciones anteriores dieron lugar a varios proyectos y el presente artículo reporta una etapa avanzada en el análisis de este problema.

Para estudiar más en detalle estas situaciones se seleccionó un caso que había sido muy estudiado: Sistema de Planes de Ahorro Previo para la Adquisición de Vehículos OKm [11] [12]. La importancia de este caso reside en que: a) Se habían realizado 9 replicaciones aisladas del mismo trabajo por diferentes grupos de personas (varios grupos de profesores universitarios y alumnos de posgrado); b) estas replicaciones fueron realizadas desconociéndose que posteriormente serían utilizadas para un estudio de completitud por parte de otras personas, por lo que las mismas están absolutamente no contaminadas con el propósito del presente estudio; c) el caso ha sido muy estudiado por los autores; y d) la fuente de información del problema bajo estudio es muy confiable.

### **1.3 Trabajos Previos**

Doorn y Ridao [12] [13] aplicaron una variante del método de captura y recaptura, denominado DPM (Detection Profile Method) [14], sobre modelos de requisitos escritos en lenguaje natural para predecir el tamaño de dichos modelos, es decir, estimar la cantidad de información aún faltante de elicitar. En base a los estudios que realizaron concluyeron que al aumentar el número de elicidores (ingenieros de requisitos) la cantidad de elementos estimados se acercaría significativamente a la cantidad realmente existente en el dominio del problema. Dado los resultados que obtuvieron al aplicar el método DPM sobre el modelo de requisitos, se realizó un análisis semántico sobre los elementos del modelo antes de aplicar el método predictivo con el fin de evitar distorsiones en la predicción del tamaño del mismo. Este análisis semántico se presenta en este artículo. Asimismo, se estudió las visiones de los elicidores respecto al mismo dominio del problema, factor que no es contemplado por las técnicas de elicitación, y se comprobó que no siempre los elicidores están observando el mismo problema, lo cual provoca diferencias entre los modelos construidos por cada elicitor.

### **1.4 Estructura del Artículo**

En la siguiente sección se plantea una hipótesis para poder dar una explicación a la diferencia de elementos observados por los elicidores. En la sección 3 se presenta el análisis semántico realizado y el posterior estudio estadístico, seguido de la sección 4 donde se realizan observaciones al estudio realizado. Finalmente, se mencionan conclusiones del trabajo realizado y se delimitan los próximos pasos a seguir.

## 2 Hipótesis de Trabajo

En [12] se estudió experimentalmente el uso de DPM para estimar el número total de términos de un glosario, denominado Léxico Extendido del Lenguaje (LEL). El LEL es un glosario con la denotación y connotación de los términos (símbolos) utilizados en el dominio del problema [15]. El modelo léxico que utilizaron había sido elaborado con mucha anterioridad por nueve grupos independientes, aplicando todos ellos la misma técnica de elicitación sobre el mismo universo de discurso “Sistema de Planes de Ahorro Previo para la Adquisición de Vehículos 0Km” [11].

Según este estudio [12], habría faltado encontrar aproximadamente 9 términos del LEL considerando el total de términos distintos encontrados en conjunto entre todos los grupos (118 símbolos). Ahora bien, si se considera individualmente a cada grupo de elicidores, el grupo que más símbolos elicitor alcanzó un nivel de completitud del 53%. Este valor parecería indicar que este grupo construyó un modelo con exceso de elementos omitidos.

Si dos o más grupos de ingenieros construyen el LEL del mismo dominio de aplicación, sucede que todos los grupos detectan una cantidad de símbolos diferente y que los conjuntos con menor cantidad de símbolos no son subconjuntos propios de los que tienen mayor cantidad de símbolos. Surgen entonces algunas hipótesis al estudiar qué símbolos de un LEL encuentran diferentes grupos y cómo se relacionan los símbolos hallados por los diferentes grupos. En otras palabras, dos grupos cualesquiera detectan una cantidad diferente de símbolos, pero con la propiedad que el primero de los grupos detecta algunos símbolos que no detecta el segundo y viceversa. Bajo esta circunstancia, se podrían considerar tres posibilidades: 1) Los grupos tienen características que los diferencian; 2) Se trata de dos problemas diferentes; 3) Las diferencias se producen por el mero azar.

Una de las variantes del método de captura y recaptura se basa en la hipótesis 1) y utiliza un coeficiente de corrección que distingue la probabilidad de detección de los grupos. Lo cual ha inducido a centrarse en dicha hipótesis [12]. Frecuentemente, la hipótesis 2) es rechazada por parecer absurda, y la 3) por considerarse de baja incidencia. Analizando estas cuestiones en detalle, se planteó que la diferencia está en la relación entre el grupo y el problema; es decir, que cada grupo tiene cierta tendencia a observar mejor algunos aspectos del problema que otros.

Lo primero que se debe hacer al construir un LEL de un problema es definir los límites del universo de discurso (UdeD). ¿Cuándo se hace esto? ¡En el momento en que se sabe menos del problema! Habitualmente, se están definiendo los límites del UdeD en el momento en que no se conoce casi nada del mismo. Esto es un caso práctico del teorema de Jackson [16], en el sentido de no usar el enfoque “top down” en la Ingeniería de Requisitos, dado que las decisiones más importantes no se deben tomar cuando menos se sabe del problema. Como consecuencia de este análisis, se elaboró una hipótesis de trabajo: “*Los grupos de ingenieros tratan problemas diferentes*”. Esto indicaría que no hay un exceso de términos omitidos sino que los grupos independientes tuvieron visiones disímiles de ese universo.

Bajo esta visión de problemas diferentes, se realizó un estudio estadístico de esos nueve modelos léxicos haciendo una corrección semántica para comparar las distintas visiones. Para ello, se identificaron cinco categorías o sub-problemas que abarcaban el problema en estudio y se analizaron estadísticamente cómo cada grupo profundizó o

no la elicitación de información correspondiente a cada categoría. Las conclusiones preliminares confirmarían promisoriamente la hipótesis planteada.

### 3 Análisis Semántico

Se realizó un estudio de la semántica de todos los términos del LEL definidos por los nueve grupos de elicitadores. En principio, se analizó la denotación y connotación de los términos que se consideraron dudosos. Los términos dudosos fueron examinados para determinar su relevancia en el UdeD (aportan o no conocimiento) y su pertenencia al UdeD (existen realmente en él o fueron creados artificialmente por el grupo de elicitadores).

Se encontraron términos cuyo nombre no existía en el UdeD y, en algunos casos, sus denotaciones y connotaciones contenían información relevante. Asimismo, se analizó la denotación y connotación de todos aquellos términos cuyos nombres no eran exactamente iguales o no eran sinónimos evidentes de símbolos definidos por otros grupos, con el fin de determinar aquellos símbolos que habían sido identificados por un único grupo, es decir, ningún otro grupo los había detectado. En pocos casos, se encontraron sinónimos dentro del mismo grupo: dos términos definidos separadamente pero que tenían el mismo significado (denotación más connotación).

En resumen, los símbolos calificados como no relevantes involucraron distintos aspectos: a) información que estaba fuera del límite del UdeD; b) información con un exceso de nivel de detalle, innecesario para un glosario, que podía ser mencionada dentro de otro símbolo; y c) información que ya estaba contenida en otros símbolos.

Para comprobar la hipótesis planteada, se estudió estadísticamente la visión que tenía cada grupo de ingenieros aplicando la corrección semántica. Las 5 categorías utilizadas fueron: Adhesión, Adjudicación, Gestión de Cuotas, Administración de Bien Tipo y Administración de Plan. Esto se pudo realizar debido a que los autores tienen hoy más conocimiento sobre el UdeD que los nueve grupos de elicitadores.

La Tabla 1 muestra el estudio estadístico PRE y POST análisis semántico de los símbolos del LEL observados por cada grupo en cada categoría, donde se pueden observar las diferencias al descartar los símbolos no relevantes. Se pasó de un total de 129 símbolos sin repetición observados en conjunto por los 9 grupos a un total de 97.

Se observó que las curvas obtenidas para todas las categorías eran aproximadamente gaussianas, con una leve asimetría a la derecha. Esto permitió utilizar la distribución t de Student dado que se tenían muestras de tamaño pequeño [17]. Las categorías Adhesión y Adjudicación fueron percibidas cada una como un único universo por los grupos; esto se comprobó aplicando la prueba chi-cuadrado [17]. En la categoría Gestión de Cuotas se observó la existencia de un UdeD y un valor atípico (“outlier”) correspondiente al Grupo 5, comprobado aplicando el test t-Student sin el outlier. En las categorías Administración de Bien Tipo y Administración de Plan se observaron 3 universos y 2 universos respectivamente; aquí también se aplicó el test t-Student para comparar grupos de datos independientes y obtener el nivel de significación de que sean universos diferentes [18].

**Tabla 1.** Cantidad de Símbolos detectados por Grupo y por Categoría.

Categoría	1-Adhesión		2-Adjudicación		3-Gestión de Cuotas		4-Administración de Bien Tipo		5-Administración de Plan		Total del Grupo	
	PRE	POST	PRE	POST	PRE	POST	PRE	POST	PRE	POST	PRE	POST
<b>Grupo 1</b>	7	6	13	11	13	13	17	12	4	4	54	46
<b>Grupo 2</b>	4	4	6	6	9	9	6	6	4	3	29	28
<b>Grupo 3</b>	8	6	6	5	8	7	4	4	5	2	31	24
<b>Grupo 4</b>	6	6	7	7	7	7	9	9	6	6	35	35
<b>Grupo 5</b>	8	7	10	9	21	20	15	15	13	11	67	62
<b>Grupo 6</b>	8	4	7	4	3	1	4	3	4	2	26	14
<b>Grupo 7</b>	8	7	10	11	10	11	9	9	9	6	46	44
<b>Grupo 8</b>	4	4	7	5	5	5	6	6	2	2	24	22
<b>Grupo 9</b>	7	7	12	10	11	11	12	11	13	12	55	51
Total categoría	<b>60</b>	<b>51</b>	<b>78</b>	<b>68</b>	<b>87</b>	<b>84</b>	<b>82</b>	<b>75</b>	<b>60</b>	<b>48</b>		
Total sin repetir	<b>20</b>	<b>13</b>	<b>26</b>	<b>20</b>	<b>35</b>	<b>31</b>	<b>27</b>	<b>19</b>	<b>21</b>	<b>14</b>	<b>129</b>	<b>97</b>
Media	6,67	5,67	8,67	7,56	9,67	9,33	9,11	8,33	6,67	5,33		
Varianza	2,75	1,75	7,00	7,53	27,25	29,00	22,11	15,50	16,50	14,75		
Desviación Std	1,66	1,32	2,65	2,74	5,22	5,39	4,70	3,94	4,06	3,84		
Lím confianza+	8,32	6,99	11,31	10,30	14,89	14,72	13,81	12,27	10,73	9,17		
Lím confianza -	5,01	4,34	6,02	4,81	4,45	3,95	4,41	4,40	2,60	1,49		

El trabajo completo se detalla en [19] y los primeros resultados se presentaron en [20].

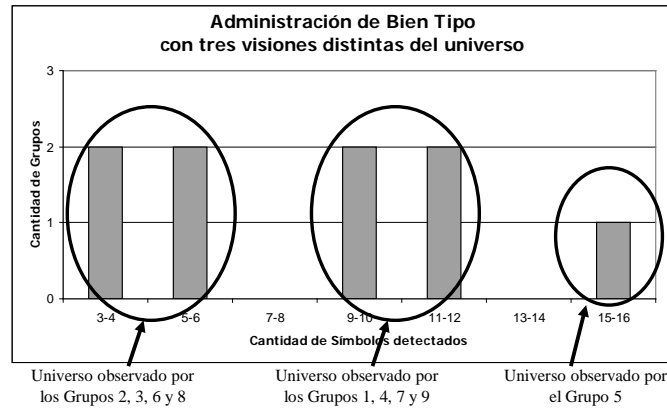
**Tabla 2.** Administración de Bien Tipo: resultados del análisis semántico.

Administración de Bien Tipo	Total de Símbolos (Estadísticos)	Total de Símbolos Relevantes (Semánticos)	Símbolos No Relevantes	Nombres de Símbolos Inexistentes	Símbolos Únicos	Símbolos Únicos Relevantes
<b>UdeD<sub>1</sub></b>						
Grupo 2	6	6	0	0	0	0
Grupo 3	4	4	0	0	0	0
Grupo 6	4	3	2	3	2	0
Grupo 8	6	6	0	0	0	0
<b>UdeD<sub>2</sub></b>						
Grupo 1	17	12	6	0	8	2
Grupo 4	9	9	0	0	0	0
Grupo 7	9	9	0	0	0	0
Grupo 9	12	11	0	0	1	1
<b>UdeD<sub>3</sub></b>						
Grupo 5	15	15	0	0	2	2
<b>Total</b>	<b>82</b>	<b>75</b>	<b>8</b>	<b>3</b>	<b>13</b>	<b>5</b>

<b>Total de Símbolos Sin Repetición</b>	<b>27</b>
<b>Total de Símbolos Sin Repetición No Relevantes</b>	<b>8</b>
<b>Total de Símbolos Sin Repetición Relevantes</b>	<b>19</b>

En la Tabla 2 se presentan los datos estadísticos obtenidos para la categoría Administración de Bien Tipo, donde se muestran los resultados luego del análisis semántico de las denotaciones y las connotaciones de los símbolos no exactamente iguales y de los símbolos considerados no relevantes.

Se analizó la distribución por rangos de símbolos para esta categoría, que se presenta en la Fig. 1, donde se observan tres universos distintos: quienes hallaron de 3 a 6 símbolos del LEL, denominado UdeD<sub>1</sub>; quienes hallaron de 9 a 12 símbolos, denominado UdeD<sub>2</sub>; y quienes hallaron 15 símbolos, denominado UdeD<sub>3</sub>.



**Fig. 1.** Distribución de grupos para Administración de Bien Tipo.

Se aplicó el test t de Student para comparar dos medias independientes utilizando la fórmula (1) [18], y se obtuvieron los resultados de la Tabla 3.

t-Student del UdeD<sub>k</sub> respecto al UdeD<sub>k+1</sub>:

$$t_{jk} = \frac{\overline{Q}_{j,k+1} - \overline{Q}_{j,k}}{\sqrt{\frac{(n_k - 1)\sigma_{j,k}^2 + (n_{k+1} - 1)\sigma_{j,k+1}^2}{n_k + n_{k+1} - 2}} \sqrt{\frac{1}{n_k} + \frac{1}{n_{k+1}}}} \quad (1)$$

Siendo: k: UdeD; j: categoría; n<sub>k</sub>: cantidad de grupos del UdeD<sub>k</sub>;  $\overline{Q}_{jk}$ : media del UdeD<sub>k</sub> en la categoría j;  $\sigma_{jk}^2$ : varianza del UdeD<sub>k</sub> en la categoría j.

**Tabla 3.** Administración de Bien Tipo: Probabilidad de diferencia estructural.

k	Grupos	UdeD <sub>k</sub>	n <sub>k</sub>	Media	Varianza	Desviación Estándar	t-Student	p-valor
1	2, 3, 6, 8	{6, 4, 3, 6}	4	4,75	2,250	1,500	5,185	0,003
2	1, 4, 7, 9	{12, 9, 9, 11}	4	10,25	2,250	1,500	2,832	0,0008
3	5	{15}	1	15	0	0		

En base a estos resultados, se confirma la existencia de tres universos en esta categoría, dado que se observa un nivel de significación de 0,05. Es decir, se rechaza

la igualdad entre el UdeD<sub>1</sub> y el UdeD<sub>2</sub>: ambos universos presentan una diferencia estructural, pues  $p\text{-valor} = 0,003 < 0,05$  (probabilidad de diferencia estructural 99,60%). Del mismo modo, el UdeD<sub>2</sub> y el UdeD<sub>3</sub> presentan también una diferencia estructural, lo cual fue observado con similar nivel de significación, pues  $p\text{-valor} = 0,0008 < 0,05$  (probabilidad 99,63%).

## 4 Observaciones

En base al análisis semántico-estadístico realizado [19], se ha observado que:

- Hubo 3 grupos que elicitaron con un mismo nivel de detalle todas las categorías.
- Un solo grupo observó débilmente todas las categorías, pues no aportó ningún símbolo único relevante y en todas las categorías siempre identificó menos símbolos que el resto. Solo detectó 14 símbolos relevantes de los 26 que elicitó e identificó 7 símbolos cuyos nombres no existían en el UdeD. Se puede decir que tuvo una visión muy diferente del UdeD dada una captura de símbolos distorsionada frente a nuestro conocimiento de este universo.
- En el otro extremo, el Grupo 5 fue el que identificó notoriamente más símbolos relevantes en todas las categorías, excepto en una, aunque elicitó una cantidad por encima de la media. Solo 5 símbolos fueron calificados como no relevantes del total de 67 que hallaron. Este grupo encontró el 62% de símbolos del total identificados (relevantes) entre todos los grupos.
- Aún cuando este grupo puede considerarse el mejor por detectar la mayor cantidad de símbolos distintos relevantes, le han faltado encontrar una cantidad interesante de símbolos reconocidos por los otros grupos. Esto podría indicar un nivel importante de incompletitud según los símbolos reconocidos.
- Las categorías Adhesión y Adjudicación fueron vistas con un nivel de detalle homogéneo por todos los grupos, es decir, todos observaron el mismo UdeD.
- Para las otras 3 categorías, se confirmó que los grupos observaron distintos UdeD en cada sub-problema. Esto está en concordancia con la hipótesis planteada.
- En Adhesión y en Adjudicación hubo mucha dispersión en cuanto a la cantidad de símbolos detectados por los grupos. Más del 60% de los símbolos fueron encontrados solo por 1 o 2 grupos, mientras que 30% de símbolos fueron comunes a la mayoría de los grupos. O sea, hubo un fuerte núcleo común de símbolos.
- Del total de símbolos distintos relevantes, hubo casi un 20% de símbolos comunes detectados entre 7 y 9 grupos, valor que puede considerarse relativamente bajo.
- Se comprobó semánticamente la existencia de 10 símbolos cuyos nombres no pertenecían al UdeD, aunque el significado de 4 de ellos era relevante y, por lo tanto, no debería omitirse. Efectivamente, se observó que esa información estaba contenida en otros símbolos de otros grupos.
- Se identificaron 32 símbolos no relevantes de un total de 129 símbolos distintos identificados entre todos los grupos. Es decir, se descartó un 25% de símbolos para realizar el estudio estadístico sobre la visión de los universos.

Luego del análisis semántico-estadístico se confirmó la hipótesis de trabajo planteada, pues se ha podido constatar que varios grupos observaron distintos UdeD para una misma categoría. Esto se dio en tres de las cinco categorías establecidas.



## 5 Conclusiones

Sobre la base de los resultados estadísticos obtenidos por Doorn y Ridao [12] al aplicar DPM [14] al modelo LEL creado en forma independiente por grupos de ingenieros, se elaboró una hipótesis para dar una explicación a tan bajo nivel de completitud alcanzado. Esta hipótesis plantea que no hay un exceso de omisiones en el modelo construido sino que los grupos tuvieron visiones muy disímiles de ese UdeD. Esas diferentes visiones podrían ocurrir por manejar distintos límites en el mundo real o por estudiarlo con distinto grado de detalle.

Para comprobar esta hipótesis, se estudió estadísticamente la visión que tenía cada grupo de elicitadores haciendo una posterior corrección semántica para comparar las distintas visiones. Entonces, inicialmente se aplicó un estudio netamente estadístico considerando exclusivamente los nombres de los símbolos y los sinónimos expresamente identificados por nombres idénticos o gramaticalmente idénticos (flexiones verbales, plurales, formas sustantivas para verbos). Este fue el mismo criterio aplicado en [12] para identificar el conjunto de símbolos distintos detectados por todos los grupos. Luego, se estudiaron las denotaciones y connotaciones de cada símbolo, para establecer efectivamente los sinónimos y los símbolos realmente importantes del UdeD. De la cantidad inicial de 129 símbolos diferentes identificados en conjunto por los nueve grupos, se determinó que se habían identificado en total 97 símbolos diferentes relevantes. La diferencia de 32 símbolos correspondía a términos cuyo nombre solo sería necesario mencionar en otros símbolos o cuyo contenido podría incluirse en otros símbolos, y un pequeño número correspondía a nombres de términos inexistentes en el UdeD.

Se pudo establecer estadísticamente, analizando adicionalmente la semántica de los elementos observados, que los grupos tuvieron distintas percepciones del dominio del problema según las categorías.

Además, se observó que aplicando el método de captura y recaptura para estimar el grado de completitud de un modelo de requisitos escrito en lenguaje natural, sin evaluar inicialmente la semántica de los elementos del modelo, puede provocar una distorsión importante en los resultados obtenidos. Pues en el caso del modelo LEL estudiado, luego del análisis semántico, se descartaron el 25% de los elementos observados (términos léxicos) para realizar el estudio estadístico.

Queda por analizar y comprender cuales de los enfoques diferentes es el correcto para el caso de estudio en cuestión, por ejemplo en el caso del sub-problema "Administración de Bien Tipo". La pregunta a responder es si se trata de que algunos grupos trabajaron en exceso o de que otros omitieron aspectos que pueden comprometer el éxito del futuro sistema. El objetivo es proponer heurísticas que permitan guiar el alcance del trabajo a realizar en una etapa temprana del proyecto.

Se espera refinar el análisis semántico realizado sobre el modelo léxico, elaborando con mayor precisión criterios semánticos de comparación. Esto permitirá definir algunas heurísticas para mejorar la construcción del LEL, principalmente cómo determinar si un símbolo es parte del UdeD o no, es decir, esto llevará a precisar mejor los límites del UdeD.

**Agradecimientos.** Los autores agradecen de manera especial al Ing. Aldo Sacerdoti (DIIT - UNLaM) por su asesoramiento e invaluable comentarios.

## Referencias

1. Faulk, S., et al.: The Core Method for Real-Time Requirements. *IEEE Software*, vol. 9, n° 5, pp.22--33 (1992)
2. Lutz, R.: Analyzing Software Requirements Errors in Safety-Critical Embedded Systems. En: 1st IEEE Intl Symposium on Requirements Engineering (RE'93), IEEE Computer Society Press, Los Alamitos, CA, pp.126--133 (1993)
3. Leveson, N.G.: *Safeware, System Safety and Computers*. Addison-Wesley (1995)
4. Kotonya, G., Sommerville, I.: *Requirements Engineering: Process and Techniques*. John Wiley & Sons (1998)
5. Loucopoulos, P., Karakostas, V.: *System Requirements Engineering*. McGraw-Hill, Londres (1995)
6. Firesmith, D.: Are Your Requirements Complete? *Journal of Object Technology*, vol. 4, n° 1, pp. 27--43 (2005)
7. Basili, V.R., Weiss, D.: Evaluation of a Software Requirements Document by Analysis of Change Data. En: 5th International Conference on Software Engineering, IEEE Computer Society Press, Los Alamitos, CA, pp.314--323 (1981)
8. Leite, J.C.S.P.: Gerenciando a Qualidade de Software com Base em Requisitos. *Qualidade de Software: Teoria e Prática*. Prentice-Hall, Rocha A, Maldonado J, Weber K (eds), cap. 17, pp 238--246 (2001)
9. Pitts, M.G., Browne, G.J.: Stopping behavior of systems analysts during information requirements elicitation. *Journal of Management Information Systems*, vol. 21, n° 1, pp. 203--226 (2004)
10. Carrizo, D., Dieste, O., Juristo, N.: Study of Elicitation Techniques Adequacy. En: 11th Workshop on Requirements Engineering (WER'08), España, pp. 104--114 (2008)
11. Mauco, V., Ridao, M., del Fresno, M., Rivero, L., Doorn, J.H.: Ingeniería de Requisitos, Proyecto: Sistema de Planes de Ahorro. Reporte Técnico, UNICEN, Tandil (1997)
12. Doorn, J.H., Ridao, M.: Completitud de Glosarios: Un Estudio Experimental. En: VI Workshop on Requirements Engineering (WER'03), Brasil, pp. 317--328 (2003)
13. Ridao, M., Doorn, J.H.: Estimación de Completitud en Modelos de Requisitos Basados en Lenguaje Natural. En: IX Workshop on Requirements Engineering (WER'06), Brasil, pp. 151--158 (2006)
14. Wohlin, C., Runeson, P.: Defect content estimations from Review Data. En: 20th Intl Conference on Software Engineering, Japón, pp. 400--409 (1982)
15. Leite, J.C.S.P., Franco, A.P.M.: A Strategy for Conceptual Model Acquisition. En: IEEE 1st Intl Symposium on Requirements Engineering (RE'93), IEEE Computer Society Press, EEUU, pp 243--246 (1993)
16. Jackson, M.: *Software Requirements & Specifications. A lexicon of practice, principles and prejudices*. Addison-Wesley, Reading, MA/ACM Press, Nueva York (1995)
17. Cavanos, G.: *Probabilidades y Estadística. Aplicaciones y métodos*, 1° edición. Mc. Graw Hill (1988)
18. Pértega Díaz, S., Pita Fernández, S.: Métodos paramétricos para la comparación de dos medias t de Student. *Unidad de Epidemiología Clínica y Bioestadística. Complejo Hospitalario Universitario de Coruña, España. Cad Aten Primaria*, vol. 8, pp.37--41 (2001)
19. Litvak, C.S., Hadad, G.D.S., Doorn, J.H.: Estudio semántico de modelos construidos por elicitadores independientes observando el mismo problema, Proyecto: Completitud de Modelos de Requisitos, Reporte Técnico, FTI, Universidad de Belgrano, 30 pág. (2012)
20. Hadad, G.D.S., Litvak, C.S., Doorn, J.H.: Agregando semántica a técnicas de predicción de completitud en modelos de requisitos. En: XIV Workshop de Investigadores en Ciencias de la Computación (WICC 2012), Universidad Nacional de Misiones, Posadas (2012)