

# Elicitación de requerimientos con *Grounded Theory*

Alejandro Oliveros, Ana Piccin, Diana Cicinelli

Facultad de Ingeniería y Tecnología Informática/ Universidad de Belgrano

Av. F. Lacroze 1947, CABA

aoliveros@ub.edu.ar, {ana.piccin; diana.cicinelli}@comunidad.ub.edu.ar

## Resumen

La *Grounded Theory* es una metodología utilizada en las ciencias sociales. Este trabajo propone determinar si esta metodología resulta efectiva para la práctica de la elicitación de requerimientos, si el modelo que emergerá de su aplicación será válido para este propósito. Surgen de este modo dos objetivos. En primer lugar, aplicar la GT a los datos básicos de los cuales se extraerá el modelo de requerimientos. En segundo lugar, validar el modelo obtenido. El proyecto se encuentra actualmente en desarrollo y se está en vías de la construcción del modelo. Se describen los pasos seguidos y las conclusiones obtenidas hasta el momento.

**Palabras clave:** Elicitación de requerimientos, Lenguaje natural, *Grounded Theory*.

## Contexto

### Elicitación de requerimientos

Mediante la Elicitación de Requerimientos se obtiene el conocimiento relevante para producir un modelo de requerimientos de un determinado dominio [1]. Parte significativa de este conocimiento se encuentra en Lenguaje Natural (LN), la mayoría de los documentos disponibles para el análisis de requerimientos son

provistos por el usuario o resultan de entrevistas [2]. Los enfoques de elicitación con LN se pueden clasificar en aquellos que interactúan directamente con el usuario y los que lo hacen con documentos [1]. La interacción con el usuario genera nuevamente documentos en LN en los que se registra esa interacción. En documentos de los sistemas heredados, en general formulados en LN [3], como manual del usuario o especificaciones de requerimientos, hay información relevante para la construcción de un nuevo sistema.

Pese a aspectos tales como la ambigüedad o la informalidad, el LN resulta esencial y de difícil reemplazo en el proceso de capturar el conocimiento necesario para producir la Especificación de Requerimientos de Software (SRS). En determinados momentos en los que aún no está consolidado el conocimiento del dominio, mantener una cierta ambigüedad puede resultar útil para avanzar en el proceso de elicitación y a los *stakeholders* puede resultarles conveniente un modo de expresión informal. El conocimiento del significado de los términos no específicos utilizados y de sus sintaxis son aspectos que favorecen la utilización del LN. En general el LN supera a los lenguajes formales o semi-formales en el alcance de su capacidad de descripción.

A pesar de que se dispone de una importante gama de técnicas de

elicitación de requerimientos y de métodos que proveen guías para su utilización [4], se necesitan técnicas que faciliten el tratamiento de descripciones inconsistentes e incompletas [4].

El proceso de elicitación de requerimientos de un sistema software utiliza una variedad de fuentes que debe procesar para producir la SRS. La elicitación produce conocimiento y ese conocimiento es el principal input del proceso que produce la SRS. No es lo mismo la información del mundo real (minutas, documentos, actas, especificaciones, etc.) que el conocimiento que se elabora en el proceso de elicitación. Establecida la diferencia cabe preguntarse: ¿dónde reside ese conocimiento?, ¿qué artefactos soportan ese conocimiento? Existen opciones como la construcción del Léxico Extendido del Lenguaje [4] a partir del Universo del Discurso, que permite elaborar Escenarios [5] o Casos de Uso [6].

### **Enfoques y técnicas provenientes de las ciencias sociales**

La Ingeniería de Requerimientos (RE, por sus iniciales en inglés) se desarrolla en un contexto de actividad humana organizada en términos sociales. Una parte importante de sus *stakeholders* son seres humanos, de allí que se haya apelado a fundamentos y técnicas provenientes de disciplinas como la psicología cognitiva, antropología, sociología y lingüística [7]. La RE ha incorporado una variedad de enfoques provenientes del campo de las ciencias sociales, en especial en el terreno de la elicitación. Una pequeña lista informal incluye: *Activity Theory*, *Brainstorming*, *Etnografía*, *Focus group*, *Lingüística*.

Esta lista, no exhaustiva, permite observar que los recursos de las ciencias

sociales que se han importado al campo de RE actúan en el terreno de la práctica específica del usuario (*Activity Theory*), la interacción directa con el usuario (*Brain Storming* y *Focus Group*), el entendimiento del usuario y del contexto en el cual se desempeña (*Etnografía*), y la comprensión de los documentos del proceso de desarrollo (*Lingüística*).

### **Introducción**

*Grounded Theory* (GT) se desarrolló en el campo de la sociología con el fin de descubrir teorías a partir de un análisis particular de los datos [8]. Ha sido utilizada como recurso de investigación en el campo de la ingeniería de software en general y de la RE en particular. En [9] se detalla una lista de este tipo de investigaciones en ambos campos, al que denominan, de *uso metodológico*.

Se dispone de un reducido grupo de trabajos iniciales que han utilizado GT en la RE, no como un recurso metodológico de investigación sino como una herramienta de desarrollo de software [9]: integración de etnometodología y GT (para obtención de datos y de análisis respectivamente); elicitación y análisis de requerimientos para la construcción del Modelo del Dominio; obtención de Clases, Diagramas de Flujo de Datos y Diagramas Entidad Relación; elaboración de Casos de Uso; producción de escenarios; extracción de los requerimientos funcionales y no funcionales de la Arquitectura de la Empresa y de la descripción del sistema.

En términos generales la GT provee un conjunto de herramientas que ayudan a la modelización y descripción del sistema en construcción.

### **Líneas de Investigación, Desarrollo e Innovación**

El trabajo se enfoca en la utilización de GT como una herramienta de desarrollo de software, en el ámbito de RE. Dentro de este contexto se desarrollan tareas de apropiación de la metodología, para lo que se tienen en cuenta dos ejes que se desarrollan simultáneamente: el metodológico y el específico del dominio.

El primero incluye la transposición metodológica del dominio de las ciencias sociales al de la elicitación de requerimientos. El segundo guía la construcción del modelo de requerimientos desde el dominio de la Ingeniería de Software.

La inclusión de los *stakeholders*, como objeto de estudio de la Ingeniería de Software, ha alentado a utilizar metodologías de investigación de orden cualitativo-interpretativo, propias del dominio de las ciencias sociales. La aplicación de la GT a otros dominios, en este caso el de la elicitación de requerimientos, exige un estudio profundo que asegure la validez de los resultados.

La aplicación de la GT tiene en cuenta las siguientes instancias: Recolección de datos, Codificación, Muestreo, Comentarios (memos), Clasificación y Redacción. El orden de las instancias no es estricto, algunas de ellas pueden desarrollarse simultáneamente. Se debe aplicar la “comparación constante”, la que obliga a la revisión permanente del trabajo realizado. [15]

La GT, aplicada rigurosamente, permite que emerjan hipótesis de los datos. Para el caso de su aplicación a la ingeniería de requerimientos esto exige, por un lado, el esfuerzo de despojarse de la inercia de los métodos de investigación tradicionales del dominio. Por el otro, y a lo largo de todo el análisis, mantener la atención en no importar hipótesis de otros entornos.

Para asegurar la validez del estudio fue necesario: 1) capacitarse el grupo de trabajo para desarrollar criterios que permitieran determinar la escuela que guiaría el estudio: ya sea la clásica elaborada por Glaser y Strauss y continuada por el primero [8] o la alternativa desarrollada por Strauss y Corbin [10], optándose por la primera; 2) leer críticamente y en profundidad la bibliografía fundamental de la GT para desarrollar las competencias necesarias para su aplicación a los fines de este estudio; 3) revisiones y contrastaciones de términos e interpretaciones a lo largo de toda la aplicación de la metodología.

La elicitación de requerimientos se basó en el estudio del documento que constituye un contrato de auto-ahorro. La aplicación sistemática de la GT permitió que, a través de las distintas instancias de abstracción, surgieran las categorías sobre las que, finalmente, se organizará el modelo.

La pregunta de investigación de mayor alcance es: “¿es posible obtener el conocimiento necesario para producir la SRS mediante la utilización de GT?”. La hipótesis se puede formular como: “El conocimiento que se obtiene del proceso de elicitación de requerimientos puede alcanzarse aplicando GT”.

## Resultados y Objetivos

Se dispone de un trabajo exhaustivo para la elicitación de los requerimientos en un dominio dado mediante la elaboración de LEL. Se trata del análisis del mismo universo: el sistema de Planes de Ahorro Previo para la Adquisición de Vehículos 0km sobre la base del documento “Solicitud de adhesión. Condiciones generales de contratación”, que es un contrato tipo para un sistema de autoahorro [11]. Sobre este se realizó un trabajo de aplicación de GT para luego

cotejar los resultados obtenidos con. A continuación se describen las actividades realizadas.

**Objetivo.** Mediante la aplicación de GT obtener el conocimiento suficiente del dominio para especificar los requerimientos.

**Selección de las herramientas informáticas.** Se utilizó una combinación de procesador de texto y hojas de cálculo. Ello permitió conocer en profundidad los requerimientos de una herramienta automatizada de soporte que se aspira a diseñar e implementar.

**Organización de la información.** Se utilizó una tabla con múltiples columnas (Texto original; Código; Concepto; Categoría) para relacionar las sucesivas codificaciones y anotaciones de cada fragmento de texto.

El proyecto se encuentra en desarrollo y, hasta la fecha, se han logrado los siguientes objetivos de análisis:

**Recolección de datos:** Se identificaron en el texto original los fragmentos a ser analizados y se organizaron en la columna “Párrafo”. En el proceso estos fragmentos fueron subdivididos y/o fusionados

**Codificación:** A cada fragmento se le asignó (“emergió”) un *código*, entendido como término que refleja el significado del texto. La ejecución de la *comparación constante*, propia de la metodología, provocó que los códigos fueran revisados reiteradamente para mantener su consistencia.

**Conceptos.** Mediante el análisis de los códigos emergieron conceptos que, comparados por igualdad o diferencias, permitieron obtener las *categorías*.

**Muestreo teórico.** Ocurre durante la redacción final del modelo. Garantiza el vínculo entre la teoría en construcción y los datos originales. Aún no se ha alcanzado esta etapa en el trabajo.

**Comentarios.** Durante el proceso de comparación se registraron los

comentarios de los investigadores que hacen a las razones y comprensión de los códigos, conceptos y categorías. Dado que el estudio está dirigido a la elicitación de requerimientos, la redacción de los comentarios se orientó a desarrollar textos y fórmulas que pudieran ser utilizados en la formulación de los requerimientos. En GT, al concluir el proceso de análisis de los documentos, estos memos son el material de trabajo de los investigadores.

**Clasificación.** Durante el proceso de comparación constante, a partir de los códigos, emergieron categorías. A medida que las categorías iban surgiendo, se comenzó a percibir el modelo.

Un recurso que se utiliza a lo largo de todo el proceso es la *comparación constante*. La identificación de códigos es un proceso de abstracción sobre el texto original. A medida que se progresa en la lectura del texto y se identifican nuevos códigos, surge la necesidad de rever los que ya han sido definidos. Esto lleva a la modificación de algunos términos para asegurar la lógica de abstracción a lo largo del proceso. De este modo, la relación entre código y párrafo tendrá la misma calidad a lo largo de todo el análisis y se facilitarán las necesarias iteraciones, o reconstrucciones de las interpretaciones, requeridas por la metodología. El mismo enfoque se aplica a la revisión de conceptos, categorías y memos.

## Formación de Recursos Humanos

Se encuentra en desarrollo el trabajo final de carrera “Requerimientos funcionales de las aplicaciones educativas”, que se desarrollará mediante la aplicación de GT.

**Equipo de trabajo.** Lic. Alejandro Oliveros, director; Lic. Ana M. Piccin,

profesora investigadora; Lic. Diana Cicinelli, profesora investigadora.

## Referencias

- [1] P. Loucopoulos and V. Karakostas, *Systems Requirements Engineering*. McGraw-Hill, 1995.
- [2] D. M. Berry and E. Kamsties, “Ambiguity in Requirements Specification,” in *Perspectives on Software Requirements*, Springer US, 2004, pp. 7 – 44.
- [3] I. John and J. Dörr, “Elicitation of requirements from user documentation,” in *REFSQ’03 - PRE-PROCEEDINGS. 9th International Workshop on REQUIREMENTS ENGINEERING - FOUNDATION FOR SOFTWARE QUALITY*, Klagenfurt/Velden, Austria, 2003, pp. 3 – 12.
- [4] J. Leite and A. Franco, “A strategy for conceptual model acquisition,” in *Proceedings of the First IEEE International Symposium on Requirements Engineering*, San Diego, California, 1993, pp. 243 – 246.
- [5] J. C. S. do P. Leite, G. Rossi, F. Balaguer, V. Maiorana, G. Kaplan, G. Hadad, and A. Oliveros, “Enhancing a Requirements Baseline with Scenarios,” *Requirements Engineering Journal*, no. 2, pp. 184 – 198, 1997.
- [6] L. Antonelli, J. C. S. do P. Leite, G. Rossi, and A. Oliveros, “Deriving requirements specifications from the application domain language captured by Language Extended Lexicon,” in *Proceedings of Workshop on Requirements Engineering*, Buenos Aires, Argentina, 2012.
- [7] B. Nuseibeh and S. Easterbrook, “Requirements Engineering: A Roadmap,” in *ICSE ’00 Proceedings of the Conference on The Future of Software Engineering*, Limerick, Ireland, 2000, pp. 35 – 46.
- [8] Glaser, Barney G. and A. L. Strauss, *The Discovery of Grounded Theory: Strategies for Qualitative Research*. New Jersey, USA: Aldine Transaction, 2008.
- [9] D. M. Berry, M. W. Godfrey, R. Holt, C. J. Kapsner, and I. Ramos, “Requirements Specifications and Recovered Architectures as Grounded Theories,” *The Grounded Theory Review*, vol. 12, no. 1, pp. 56 – 65, Jun. 2013.
- [10] Strauss, Anselm L. and Corbin, Juliet M., *Basics of Qualitative Research : Techniques and Procedures for Developing Grounded Theory*. Sage Publications, Inc., 1998.
- [11] C. Litvak, G. Hadad, and J. Doorn, “Correcciones semánticas en métodos de estimación de completitud de modelos en lenguaje natural,” presented at the WER 2013 – 16th Workshop on Requirements Engineering, Montevideo, Uruguay, 2013, pp. 105 – 117.