

HACIA LA ADAPTACIÓN DE SCRUM PARA INCORPORAR CALIDAD DE DATOS EN ÁMBITO DEL DESARROLLO ÁGIL

Carrizo Claudio*, Javier Saldarini*, Gonzalo Castillo*, Sofía Bovo*, Rocío Cortese*, Angélica Caro#, Carlos Salgado+, Alberto Sanchez+, Mario Peralta+

*Facultad Regional San Francisco – Universidad Tecnológica Nacional

Av. de la Universidad 501 - San Francisco - Córdoba - Tel. 03564-421147

{cjcarrizo77, saldarinijavier, gonzalocastilloar, sofibovo501, rociocortese24}@gmail.com

#Departamento de Ciencias de la Computación y Tecnologías de la Información - Facultad de Ciencias Empresariales, Universidad del Bío-Bío

Casilla 447, 3780000, Chillán, Chile

{mcaro}@ubiobio.cl

+Departamento de Informática - Facultad de Ciencias Físico-Matemáticas y

Naturales Universidad Nacional de San Luis

Ejército de los Andes 950 – C.P. 5700 – San Luis – Argentina

{csalgado, alfanego, mperalta}@unsl.edu.ar

RESUMEN

En la actualidad, la mayoría de las organizaciones utilizan Sistemas de Información (SI) para llevar adelante sus procesos de negocio y toma de decisiones. Los sistemas informáticos son parte de los SI, y uno de sus elementos principales son los datos. Estos comúnmente son utilizados en diferentes ámbitos, como por ejemplo, procesos de negocio, gestión del conocimiento, toma de decisiones, explotación de la información, analítica de datos, etc. Por este motivo, resulta lógico pensar que los datos deben tener un nivel de calidad adecuado, de manera que las organizaciones puedan cumplir sus objetivos estratégicos. En el ámbito del desarrollo ágil, no es común definir, especificar e implementar requisitos de calidad en relación a los datos, y la metodología ágil Scrum no es la excepción. El objetivo de esta línea de investigación consiste en realizar una propuesta de adaptación del Framework de Scrum, para que se considere la definición, especificación e implementación de requisitos de calidad de datos, usando

como referencia la Norma ISO/IEC 25000. La principal contribución de esta investigación consiste en ayudar a garantizar la calidad de los datos en contexto del desarrollo ágil, particularmente cuando se usa Scrum.

Palabras clave: Requisitos de Calidad de Datos – ISO 25000 - Desarrollo Ágil – SCRUM -

CONTEXTO

El presente trabajo se enmarca dentro del siguiente proyecto de Investigación:

- Ingeniería de Software: Estrategias de Desarrollo, Mantenimiento y Migración de Sistemas en la Nube – Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis. Proyecto N° P-03-2020. Dicho proyecto es la continuación de diferentes proyectos de investigación a través de los cuales se ha logrado un importante vínculo con distintas universidades a nivel nacional e internacional. Además, se encuentra reconocido por el programa de Incentivos.

1. INTRODUCCIÓN

Los Sistemas de Información (SI) forman parte de un sistema mayor denominado “Sistema de la Empresa” (ISO/IEC 25030:2007), por lo tanto, la mayoría de las organizaciones hace uso de dichos SI para llevar adelante sus procesos de negocio y la toma de decisiones. Según (ISO/IEC 25030:2007), en un modelo de sistema, los sistemas informáticos son parte de un SI. En (M. Piattini et. al, 2011) se menciona no solo la relevancia que tienen los SI para las organizaciones, sino que también se hace hincapié acerca de la importancia que tiene la calidad de los sistemas informáticos, para que las organizaciones puedan cumplir sus objetivos estratégicos. En (ISO/IEC 25030:2007) también se menciona que los sistemas informáticos están compuestos por los siguientes 4 elementos: hardware, sistemas operativos, software de aplicación y datos. Este último, es un elemento de gran importancia, ya que es fundamental para la toma de decisiones objetivas en todos los niveles de una organización (Javed, B. & Hussain, S., 2003; Levy S., 2004; Naveh, E. & Halevy, A., 2000). Las organizaciones modernas consideran a los datos como el activo más importante, y como uno de sus recursos estratégicos (Olson, J., 2002; Tayi, G. & Ballou, D., 1998), debido a que a partir de los mismos, se pueden llevar adelante proyectos relacionados a la generación de conocimiento (Ponjuán, Gloria, 2006), proyectos de integración (Crm, 2002), dar soporte a la toma de decisiones y explotación de la información (Inmon, W.H., 1996), afrontar proyectos de analítica de datos (V. M. Schönberger & K. Cukier, 2013), etc.

Según lo expuesto anteriormente, es fundamental que los datos que gestiona un SI tengan un nivel de calidad adecuado,

esto significa, que estén libres de errores y que posean características relevantes, para que el dato sea de utilidad para quienes lo consuman. Estas características permiten establecer un lenguaje común con el propósito de focalizar problemas de calidad de datos y oportunidades de mejora (Javed, B. & Hussain, S., 2003; Naveh, E. & Halevy, A., 2000).

En (ISO/IEC 25012: 2008) se define la calidad de datos como el “grado en que las características de los datos satisfacen necesidades implícitas y establecidas cuando son usados en condiciones específicas”.

En la literatura existen varias propuestas que abordan la calidad de los datos desde la perspectiva de características (Richard Y. Wang, D.M.S., 1996; ISO/IEC 25012: 2008), una de ellas corresponde a una reconocida Familia de Normas denominada ISO/IEC 25000, también conocida como SQuaRE (Systems and software engineering-Systems and software Quality Requirements and Evaluation), la cual propone un marco para la definición de requerimientos de la calidad de software/datos y evaluación de la calidad del software/datos, apoyados por un proceso de medición de la calidad de software/datos. En lo que respecta a la calidad de datos, SQuaRE incluye la Norma ISO/IEC 25012, la cual proporciona un modelo general de características de calidad de datos, que están contextualizadas desde 2 puntos de vista: inherente y dependiente del sistema. Por otro lado, también incluye la Norma ISO/IEC 25024, la cual brinda un conjunto de medidas asociadas a las características del modelo de calidad de datos.

En cuanto al abordaje de la calidad de datos en el ámbito del desarrollo de software, en (R. Wang, 1998) se destaca la

importancia de incorporar aspectos de calidad de datos en fases tempranas del desarrollo, específicamente en fase de desarrollo de requisitos. Si bien existen trabajos que abordan la incorporación de aspectos de calidad de datos en etapas tempranas del proceso de desarrollo de software (A. Caro et. al, 2013; A. Rodríguez y A. Caro, 2013; C. Guerra-García et. al, 2016; Y. Fernanda Torres Alonso, 2020), se han encontrado escasas propuestas en el ámbito del desarrollo de Software, a través del uso de Metodologías Ágiles. En la industria del software, existen varias metodologías de este tipo, entre ellas Scrum (Ken Schwaber & Jeff Sutherland, 2020), Extreme Programming y Kanban, siendo Scrum la metodología ágil más adoptada y utilizada a nivel mundial. En el contexto de Scrum, también existen escasas propuestas que aborden la incorporación de aspectos de calidad asociados a datos, desde la perspectiva de la especificación temprana e implementación de requisitos de calidad de datos, en un escenario de ambiente controlado de desarrollo.

Lo descripto anteriormente, nos motivó a iniciar una línea de investigación cuyo objetivo principal es proponer una adaptación del Framework de Scrum, a fin de incorporar aspectos de calidad de datos, desde la perspectiva de la especificación e implementación de requisitos de calidad de datos. Estos requisitos se obtienen a partir de un proceso de análisis y definición, por medio del uso de las Normas ISO/IEC 25030, ISO/IEC 25012, e ISO/IEC 25024.

La adaptación propuesta incluirá 3 aspectos principales:

1. La creación de un nuevo rol dentro de Scrum, el cual tendrá la tarea de gestionar la calidad de los datos.

2. La definición de un método o técnica que permita guiar el trabajo de especificación y verificación de cumplimiento de implementación de requisitos de calidad de datos.
3. El desarrollo de herramientas que faciliten la aplicación del método o técnica propuesta.

El aporte principal de esta línea de investigación consiste en potenciar el marco de trabajo de Scrum, de manera que se pueda abordar la calidad de datos desde el momento en que se inicia el desarrollo de un producto de software.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Los principales ejes de trabajo de esta línea de investigación se detallan a continuación:

- Realizar una revisión sistemática de la literatura acerca de la temática calidad de datos en contexto del desarrollo de software en general, y en el ámbito del uso de la metodología ágil Scrum en particular.
- Realizar un estudio de la Familia de Normas ISO/IEC 25000, en particular de las Normas ISO/IEC 25030, ISO/IEC 25012, e ISO/IEC 25024.
- Realizar un estudio de la última versión del manifiesto ágil de Scrum, para comprender como funciona la metodología ágil.
- Definir un rol dentro de Scrum, el cual será responsable de velar por la calidad de los datos, durante el desarrollo del producto de software.
- Definir métodos o técnicas que permitan guiar la especificación y

verificación de implementación de requisitos de calidad de datos, basados en SQuaRE.

- Realizar pruebas de concepto para realizar los ajustes que sean necesarios.
- Desarrollar herramientas que permitan facilitar el uso y/o aplicación del método o la técnica definida.

3. RESULTADOS OBTENIDOS/ESPERADOS

Resultados Obtenidos

- Informe que contiene el resultado de la revisión sistemática de la literatura.
- Informe que contiene el resultado del estudio realizado sobre la Familia de Normas ISO/IEC 25000.
- Conjunto de requisitos de calidad de datos, obtenidos a partir de un proceso de análisis y definición, a través de ISO/IEC 25012 e ISO/IEC 25024.
- Informe del estudio realizado al manifiesto ágil de Scrum.

Resultados Esperados

- Método o técnica definida y ajustada, luego de la prueba de concepto realizada.
- Herramientas de uso libre que permitan automatizar el método o la técnica propuesta.
- Nueva versión del Framework de Scrum, que integre el método o técnica propuesta, y las herramientas.

FORMACIÓN DE RECURSOS HUMANOS

Esta línea de investigación se trabaja desde el Grupo de Investigación

denominado "Calidad de Software y Datos", el cual pertenece a la UTN Facultad Regional San Francisco, Córdoba. El equipo de trabajo está constituido por: 1 Director de Proyecto, 1 Co-Director, 4 docentes investigadores, 1 graduado y 1 estudiante, todos de la especialidad Sistemas de Información. Cabe destacar que también se están llevando a cabo algunas tesis de grado y trabajos finales, en el marco de las Ingenierías de las distintas universidades.

5. BIBLIOGRAFÍA

A. Caro, A. Fuentes, A. Soto. "Desarrollando sistemas de información centrados en la calidad de datos", Revista Chilena de Ingeniería, vol.21, pp. 54-69, 2013.

C. Guerra-García, I. Caballero, M. Cardenas-Juarez, R. Samano-Robles, "Adding quality in the user requirements specification: A first approach", 2016 Eleventh International Conference on Digital Information Management (ICDIM). IEEE.

CrM. Como Mejorar Las Relaciones Con Los Clientes. Pearson Educación, 2002.

Extreme Programming. Disponible en <https://www.agilealliance.org/glossary/xp>.

Inmon, W.H. "The data warehouse and data mining." Communications of the ACM, vol. 39, no. 11, Nov. 1996, pp.

ISO/IEC 25000 Systems and software engineering- Systems and software Quality Requirements and Evaluation (SQuaRE) - System and software quality models.

ISO/IEC 25030:2007, Software engineering — Software product Quality Requirements and Evaluation (SQuaRE) — Quality requirements., ISO, 2007.

ISO/IEC 25012: 2008. Software engineering -- Software product Quality Requirements and Evaluation (SQuaRE) -- Data quality model.

- ISO/IEC 25024:2015. Systems and software engineering - Systems and software Quality Requirements and Evaluation (SQuaRE). Measurement of data quality.
- Javed, B. & Hussain, S. (2003). Data quality – A problem and an approach. Wipro Technologies.
- Kanban. Redacción APD. Disponible en <https://www.apd.es/metodologia-kanban>.
- Ken Schwaber & Jeff Sutherland, 2020. La Guía Definitiva de Scrum: Las Reglas del Juego <https://scrumguides.org/docs/scrumguide/v2020/2020-Scrum-Guide-Spanish-European.pdf>.
- Levy, S. (2004). Model Documents and forms for Organizing and Maintaining a Data Quality Program.
- M. Piattini, F. O. G. I. García y F. Pino, Calidad de sistemas de información, 9 ed., España: RaMa, 2011.
- Naveh, E. & Halevy, A. (2000). A hierarchical framework for a quality information system. Total Quality Management, Vol.11, No. 1, p 87-111.
- Olson, J. (2002). Data Profiling: The Data Quality Analyst's Best Tool. DM Direct, December. DMReview.com.
- Ponjuán, Gloria. (2006). Gestión de Información en las Organizaciones. Editorial Félix Varela.
- R. Wang, "A Product Perspective on Total data Quality Management". Communications of the ACM. Vol. 41, Issue 2, pp. 58-65. 1998.
- Richard Y. Wang, D.M.S., Beyond accuracy: what data quality means to data consumers. Journal of Management Information Systems, 1996. 12(4): p. 5-33.
- Tayi, G. & Ballou, D. (1998).Examining Data Quality. Communications of the ACM. Vol. 41, No. 2.
- Viktor Mayer-Schönberger, Kenneth Cukier. Big data: la revolución de los datos masivos. Turner, 2013 - 320 páginas.
- Y. Fernanda Torres Alonso, "Especificación de requerimientos de software con un enfoque de calidad de datos", Univesridad Autónoma de San Luis, Potosí, 2020.