

# Propuesta Metodológica de Enfoque “Híbrido” para la Gestión de Proyectos de Minería de Datos

Patricia Cristaldo, Cristhian Richard, Ramiro Rivera,  
Esteban Schab, Anabella De Battista

Grupo de Investigación en Bases de Datos, Departamento Ingeniería en Sistemas de Información, Fac. Reg. Concepción del Uruguay, Universidad Tecnológica Nacional

Entre Ríos, Argentina

{cristaldop, richardc, riverar, schabe, debattistaa}@frcu.utn.edu.ar

**Norma Edith Herrera**

Departamento de Informática, Universidad Nacional de San Luis, San Luis, Argentina

nherrera@unsl.edu.ar

***Resumen.** Este trabajo propone una metodología de enfoque “híbrido” para la gestión de proyectos de minería de datos, enmarcada en los aspectos relevantes de la gestión tradicional y ágil, que toma como punto de partida los factores críticos de éxito de la gestión de proyectos de minería de datos y de la gestión ágil de proyectos, de forma tal de lograr un conjunto unificado de criterios que debería cumplir una metodología de gestión de proyectos en este entorno. Se especifican los criterios para orientar las decisiones, las etapas definidas en la metodología, los resultados obtenidos y las conclusiones alcanzadas.*

## 1. Introducción

En los últimos años, las metodologías de gestión de proyectos han adquirido popularidad en el ámbito de la gestión de proyectos, sobre todo en la gestión de proyectos de software. Estas guías de buenas prácticas y metodologías de gestión de proyectos, según su filosofía de desarrollo se pueden clasificar en dos enfoques: *tradicionales*, que se basan en una fuerte planificación durante todo el desarrollo y, *ágiles*, en las que el desarrollo es incremental, cooperativo, sencillo y adaptado. Entre las metodologías y guías de buenas prácticas, aplicables a proyectos de minería de datos podemos citar: CRISP-DM [KDnuggets 2014], [Chapman y otros 2010], SEMMA [Olson 2008], ATERN [Richards 2013] y SCRUM [Schwaber, K. y Sutherland, J. 2011], [Schwaber, K. y Beedle, M. 2006], entre otras.

Sin embargo, la gestión de proyectos sigue siendo un esfuerzo altamente problemático [Varajão 2016] y siguen mostrando exiguos niveles de éxito [Colomo-Palacios, y otros 2014], [Varajão y otros 2014] debido a varias razones [Montequin y otros 2016], [Chow y Chao 2016], [Lehtinen, y otros 2014], [Ramos y Mota 2014], [Elkadi 2013], [IEmam y Koru 2008].

En estudios recientes se reconoce que la mayoría de los proyectos de software utilizan una combinación de los métodos tradicionales y ágiles [Cristaldo y otros 2015] [Spundak 2014] [Conforto y Amaral 2008]. En 2014, Nogueira propone una metodología ágil para el desarrollo de proyectos de *data mining*, denominada Scrum-

DM [Nogueira 2014]. Este fue el punto de partida para el diseño de una propuesta metodológica de enfoque “híbrido” a partir de la fusión de las guías de buenas prácticas: PMBOK [PMBOK Guide 2013] de enfoque tradicional y ATERN [Richards 2013] de enfoque ágil; y las metodologías: CRISP-DM [KDnuggets 2014] de enfoque tradicional y SCRUM [Schwaber, K. y Sutherland, J. 2011], [Schwaber, K. y Beedle, M. 2006] de enfoque ágil.

En el presente trabajo se propone una metodología de gestión de proyectos de minería de datos a partir de la fusión de metodologías de enfoque ágil y de enfoque tradicional, permitiendo así reunir en una misma metodología un conjunto de factores de éxito para la gestión de proyectos de minería de datos, con el objetivo de contribuir a mejorar la gestión de este tipo de proyectos.

El resto de este documento está organizado de la siguiente manera. La Sección 2 analiza los trabajos relacionados a la luz de las distintas variantes que proponen, particularmente enfocándose en la fusión de metodologías de enfoque tradicional y metodologías de enfoque ágil. En la Sección 3 se describen las metodologías y guías de gestión de proyectos más utilizadas en la industria y que se consideran en el presente estudio a los fines de generar una propuesta metodológica de enfoque “híbrido” sólida. La Sección 4 muestra los factores críticos de éxito para un proyecto de minería de datos. La Sección 5 describe la propuesta metodológica para la gestión de proyectos de minería de datos, a partir de la fusión de metodologías y guías utilizadas en el mercado. La Sección 6 puntualiza la validación de dicha propuesta a un caso de estudio real. Finalmente, en la Sección 7 se presentan las conclusiones y trabajos futuros.

## 2. Trabajos Relacionados

En la literatura existen varias propuestas de metodologías de gestión de proyectos y particularmente de gestión de proyectos vinculados con la minería de datos.

La metodología ágil para el desarrollo de proyectos de data mining Scrum-DM propuesta por [Nogueira 2014] consiste en la asociación de la metodología ágil Scrum con la metodología tradicional Crisp-DM. Scrum-DM proporciona un marco donde el desarrollo del proyecto de data mining sigue los lineamientos de Crisp-DM y la gestión del trabajo sigue el esquema de la metodología Scrum. El proyecto se inicia a partir de la fase de *Comprensión del Negocio* de Crisp-DM donde se realiza el análisis de los objetivos y del negocio y finaliza en la fase de *Implantación* donde se realiza la integración de los resultados de la minería de datos. El desarrollo se realiza en la fase intermedia, de *Sprint*, donde se contemplan las demás fases de Crisp-DM.

Otra metodología propuesta es ASD-DM (Adaptative Software Development- Data Mining) [Alnoukari y otros, 2008]. Es una solución basada en principios ágiles para aplicaciones predictivas de minería de datos. Se basa en la metodología ASD propuesta originalmente por el mismo autor, basada fuertemente en un principio iterativo de prueba y error. Este marco asegura continuo aprendizaje e intensa colaboración entre desarrolladores, testers y clientes de minería de datos, reduciendo el tiempo de desarrollo en proyectos predictivos de minería de datos.

Otro trabajo propuesto es [Santana do Nascimento y otros, 2013], que consiste en una adaptación del KDD clásico que toma conceptos de CRISP-DM y del manifiesto ágil. Es una propuesta basada en establecer metas iterativas y ciclos más cortos entre los

entregables. El autor en este trabajo establece los puntos en común entre las metodologías de minería de datos y el manifiesto ágil.

En [Guo, 2013] se analiza el proceso de minería de datos y los desafíos técnicos que se enfrenta un “*data scientist*” durante el proyecto, logrando un ciclo de trabajo iterativo muy similar al descrito en el manifiesto ágil.

Durante los últimos años IBM ha sido una de los grandes promotores de la metodología CRISP-DM, y en el año 2015 realizaron una revisión de la metodología CRISP-DM obteniendo como resultado a la metodología ASUM-DM, que incorpora principios ágiles, extendiendo y refinando la metodología CRISP, con enfoque hacia el sector empresarial.

En los últimos años varios autores han propuesto metodologías alternativas a CRISP-DM, no obstante la mayoría difiere muy poco, limitándose a una mera reinterpretación de algunas de las fases que componen la misma con el fin de facilitar su adaptación a un campo u organización particular, a saber: la adecuación de la metodología SEMMA para gestionar un proyecto ligado al sector bancario y de inversiones y con un fuerte énfasis en la fase de implantación [Way 2013]; la adaptación de CRISP-DM a la disciplina de la medicina, proponiendo una nueva descomposición de la metodología [Niaukso, 2008].

Por otro lado autores como Mariscal (2013), Marchx (2014) y Topuszynski (2016) han realizado comparaciones entre la filosofía planteada en la literatura relacionada a las metodologías tradicionales de minería de datos y los principios del manifiesto ágil, concluyendo en que las metodologías tradicionales se beneficiarían ampliamente con la adopción de los principios ágiles.

### **3. Metodologías de Gestión de Proyectos**

En esta sección se describen las metodologías y guías de buenas prácticas que hemos considerado analizar en este trabajo, para la gestión de proyectos de minería de datos.

#### **1.1. PMBOK**

La Guía PMBOK es el estándar de gestión de proyectos del PMI (Project Management Institute) [PMBOK Guide 2013]. Su primera versión se publicó en el año 1987 y recientemente se ha publicado la quinta edición, cuya finalidad es documentar y estandarizar la información y prácticas de gestión de proyectos ampliamente aceptadas. Es un conjunto de buenas prácticas, que sirve de referencia a todos los interesados en la gestión de proyectos y que un jefe de proyecto debe entender y ser capaz de practicar, suministrando un léxico común y una estructura consistente. Comprende dos grandes secciones: la primera, sobre los procesos y contextos de un proyecto. La segunda, sobre las áreas de conocimiento específico de su gestión. Cada proceso es definido dentro de estas áreas de conocimiento y se describe en términos de sus insumos, herramientas y técnicas, y las correspondientes salidas. Es una guía descriptiva que proporciona herramientas y técnicas que facilitan la dirección de los proyectos con un enfoque de gestión tradicional.

#### **1.2. ATERN**

Es un marco de trabajo basado en un conjunto de buenas prácticas aprendidas por los miembros del Consorcio DSDM (Método de Desarrollo de Sistemas Dinámicos)<sup>1</sup>, cuya última versión fue lanzada en el año 2007. Es similar a PMBOK, pero focalizado a la gestión ágil de proyectos. El manual de ATERN presenta: principios, componentes, técnicas de modelado iterativo y la estructura de roles y responsabilidades de cada persona involucrada en el proyecto.

Abarca todo el ciclo de vida de desarrollo de un proyecto, estableciendo claramente los roles del equipo de proyecto y las partes interesadas, donde clientes y proveedores trabajan muy estrechamente con el fin de romper las barreras de comunicación. Este marco de trabajo se utiliza principalmente para el desarrollo de proyectos de software, con un enfoque gradual que abarca los principios del desarrollo ágil, incluyendo al usuario en forma continua.

### 1.3. CRISP-DM

CRISP-DM (Cross-Industry Standard Process for Data Mining) [KDnuggets 2014] es una metodología de enfoque tradicional especialmente desarrollada para proyectos de Data Mining. Actualmente está considerada como un estándar “de facto” en el mundo empresarial. Propone un conjunto de actividades que hay que llevar a cabo en el desarrollo de un proyecto de data mining y cada una de las actividades se divide en tareas, para las que se indica qué salidas produce y qué entradas necesita.

### 1.4. SCRUM

SCRUM es una metodología de gestión de proyectos iterativa e incremental que involucra buenas prácticas ingenieriles, para controlar el riesgo y optimizar la previsibilidad de un proyecto. Además, es muy utilizado dentro del área de comunicaciones y relaciones humanas [Schwaber, K. y Sutherland, J. 2011], donde establecen que los miembros del equipo trabajen juntos y de forma eficiente obteniendo productos complejos y sofisticados, guiándose por su conocimiento y experiencia más que por planes de proyecto formalmente definidos y por consiguiente la satisfacción de todos los que participan en el desarrollo, fomentando la cooperación a través de la auto-organización [Schwaber y Beedle, 2006].

## 4. Factores críticos de éxito para la gestión de proyectos de minería de datos

La primera etapa de la investigación fue identificar los factores críticos de éxito para la gestión de proyectos de minería de datos y la gestión de proyectos del enfoque ágil, para elaborar un andamiaje sólido a la propuesta de gestión de proyectos de enfoque “híbrido”. La bibliografía sugiere que la combinación de ambos enfoques, tradicional y ágil, es la solución que han encontrado varias organizaciones para la gestión de sus proyectos [Batra y otros 2010] e incluso pueden ambos enfoques complementarse en un mismo proyecto [Stare 2014], [Fernandez y Fernandez 2008] [Karlström y Runeson 2005]. En la Tabla 1 se detallan los factores de éxito propuestos por algunos autores resultados de la experiencia o en base a encuestas realizadas a profesionales para la gestión de proyectos de minería de datos [Mérida Sanchez 2017], [Ranjan y Bhatnagar 2008] [Sim 2003]. Por otro lado, Chow y Cao (2008), distinguen factores de éxito en los proyectos que se gestan utilizando un enfoque ágil, y los clasifican en cinco

---

<sup>1</sup> <http://www.dsdm.org/dig-deeper/book/dsdm-atern-handbook>

dimensiones: organización, personas, procesos, técnicas y proyecto, los cuales se muestran en la Tabla 2. Estos datos fueron recogidos de un estudio realizado a más de cien proyectos de diversas organizaciones ubicadas en diferentes lugares geográficos. Sus autores concluyen que los factores más relevantes para el éxito de los proyectos son: correcta estrategia de entrega, buen ejercicio de técnicas de ingeniería ágiles y equipo de gestión de alta capacidad.

**Tabla 1. Factores de éxito en la gestión de proyectos de minería de datos**

Dimensión	Factores de Éxito
Relaciones Personales	Colocar en primer lugar los requerimientos del cliente El Cliente debe estar comprometido con el proyecto Comunicación constante entre el cliente y el equipo de trabajo Comunicación dentro del equipo de trabajo
Planificación	Establecer objetivos claros antes de comenzar el proyecto Selección adecuada del equipo de trabajo Planificación adecuada del proyecto Definición correcta de plazos y costos
Minería y análisis de datos	Selección adecuada de los datos Entendimiento adecuado del área de negocio Uso de tecnología adecuada Interpretación correcta de los resultados
Implantación	Feedback constante Asegurar la privacidad de los datos Implementación por pasos
Control/Monitorización	Los resultados pueden no ser inmediatos Asegurar la privacidad de los datos

**Tabla 2. Factores de éxito. Enfoque ágil (Chow y Cao, 2008).**

Dimensión	Factores de Éxito
Organización	Fuerte apoyo ejecutivo. Administrador o patrocinador comprometido. Cultura de la organización cooperativa, en lugar de jerárquica. Cultura de comunicación oral o cara a cara. Establecimiento de entorno de trabajo de estilo ágil.
Personas	Miembros del equipo competentes y con experiencia. Miembros del equipo con gran motivación. Coherencia y trabajo en equipo auto-organizado. Buenas relaciones con los clientes.
Proceso	Agilidad en la gestión de requerimientos. Agilidad en la gestión de proyectos. Fuerte comunicación diaria cara a cara. Fuerte presencia y compromiso con el cliente. El cliente tiene plena autoridad.
Técnica	Normas de codificación definidas previamente. Diseño simple. Cantidad correcta de documentación. Adecuada formación técnica del equipo.
Proyecto	Proyecto dinámicos, agenda acelerada. Proyectos con equipos pequeños. Evaluación de costos por anticipado. Evaluación de riesgos por anticipado.

A partir de los datos provistos por ambas tablas, se concluye que los aspectos referidos a las dimensiones: *Minería y Análisis de Datos*, *Implantación* y

*Control/Monitorización/Entrenamiento* de la Tabla 1, se pueden incluir en la dimensión *Técnica* de la Tabla 2. De esta forma quedarían bien establecidos los factores críticos de éxito para gestionar proyectos de minería de datos.

Por otra parte, estos factores de éxito nos permiten analizar y comparar las características y restricciones de las metodologías y guías de gestión de proyectos de minería de datos existentes, de enfoque tradicional y ágil.

A los fines de reorganizar los factores críticos de éxito que muestran las tablas precedentes, para orientar las decisiones, se han formulado cuatro perspectivas. Estas perspectivas tienen como objetivos proporcionar un marco y una estructura que se espera que cumpla una metodología que gestione proyectos de minería de datos. Las perspectivas son:

**4.1 Formación y Crecimiento:** esta perspectiva incluye todos aquellos aspectos relacionados con la gestión de los recursos humanos necesarios para poder ejecutar un proyecto de minería de datos, habilidades, motivaciones, etc. Los requisitos involucrados en esta perspectiva son los inductores necesarios para el desarrollo de equipos de trabajo.

**4.2 Procesos internos:** esta perspectiva incluye todos aquellos aspectos relacionados con la gestión de los procesos del proyecto, que el jefe del mismo deberá acentuar de manera tal de lograr la satisfacción de los interesados. Las autoridades y el jefe de proyecto, deben identificar los procesos críticos en los que se debe sobresalir con excelencia si es que quieren satisfacer los objetivos que ellos mismos se propongan y los objetivos de los demás interesados en el proyecto.

**4.3 Interesados:** esta perspectiva incluye todos aquellos aspectos vinculados con la satisfacción de los interesados. Se trata de identificar claramente los interesados potenciales y la satisfacción de sus necesidades. Los requisitos involucrados en esta perspectiva representan el grado de apoyo y compromiso de las autoridades e interesados seleccionados valoran y eligen la propuesta de valor que ellos mismos entregaran al proyecto.

**4.4 Técnicas:** esta perspectiva incluye todos aquellos aspectos relacionados con las técnicas de ingeniería vinculadas con el proyecto, tales como: diseño, documentación, pruebas, tecnología, etc.

La Tabla 3, resume los principales aportes de cada metodología o guía de gestión de proyectos, en relación a los factores críticos de éxito contenidos en las perspectivas consideradas precedentemente. Los factores críticos de éxito en la gestión de proyectos de minería de datos se utilizan para realizar la comparación de características y restricciones de las metodologías y guías de gestión de proyectos.

**Tabla 3. Grado de adecuación de los factores de éxito y las metodologías y guías de gestión de proyectos de minería de datos**

Dimensión	Factores de éxito	Enfoques	
		Tradicional	Ágil

		PMBOK	CRISP-DM	SCRUM	ATERN
Formación y Crecimiento	Dinámicas de equipo en la gestión de proyecto	X		X	X
	Compromiso, colaboración y disponibilidad de los miembros del equipo e interesados	X	X		X
	Conductas de apoyo emocional entre los miembros del equipo	X		X	
	Recursos especializados según el rol que ocupa cada miembro del equipo de proyecto	X	X		X
	Gestión eficaz del director del proyecto	X	X		
	Coordinación y competencias de los miembros del equipo	X	X	X	
Procesos Internos	Definición de requerimientos, alcance y limitaciones	X	X		X
	Objetivos claros de negocios	X	X	X	X
	Políticas y procedimientos de gestión de proyectos	X		X	X
	Planificación adecuada	X	X		X
	Prácticas ágiles en los procesos	X	X	X	X
Interesados	Apoyo de la organización	X	X		
	Participación de otros interesados	X	X		
	Igual percepción entre los interesados y los profesionales de SI	X			X
	Cultura de la comunicación	X		X	X
	Definición clara de los interesados y su vinculación con el proyecto	X	X	X	
	Contratistas o patrocinadores comprometidos	X	X	X	
Técnicas	Gestión de la comunicación		X		
	Selección adecuada de los datos		X		
	Interpretación correcta de los resultados		X		
	Seguimiento y control del proyecto	X			
	Gestión de riesgos	X			X
	Gestión de la contratación	X			
	Tecnología y herramientas adecuadas	X			
	Documentación	X		X	

La Tabla 3, muestra que si bien, la mayoría de los factores críticos de éxito son considerados por las metodologías y guías analizadas, los mismos no están todos presentes en una única metodología de gestión de proyectos. A partir del análisis realizado, puede concluirse que una metodología no es suficiente para gestionar un proyecto de minería de datos, considerando los factores críticos de éxito. Por otro lado, no debería existir la aplicación de varias metodologías dentro del contexto de un proyecto, ya que conviene que todos los interesados estén alineados a una misma metodología. Como consecuencia debido a las limitaciones de los enfoques analizados,

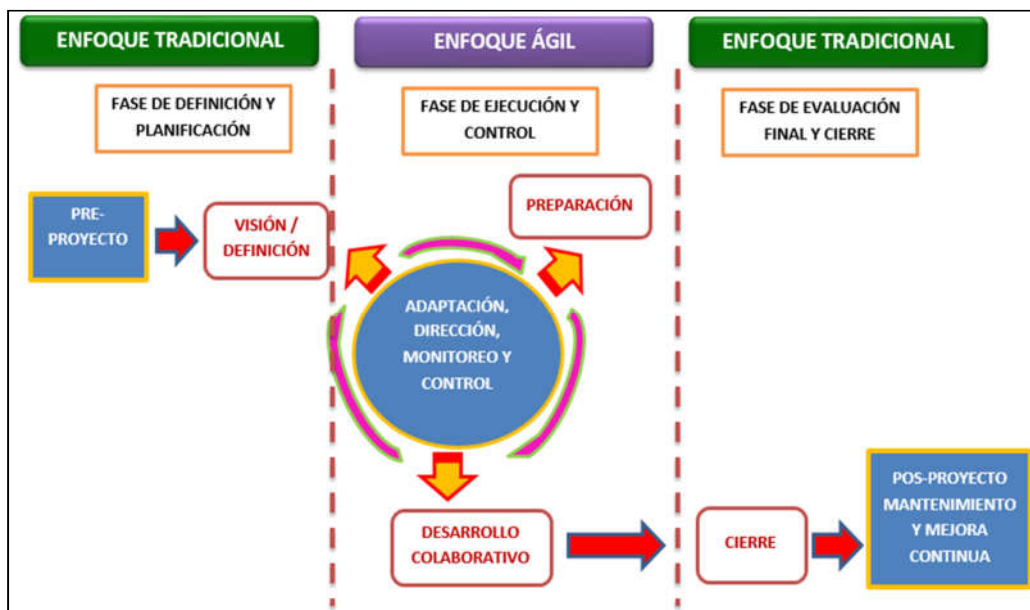


la integración de ambos enfoques, tradicional y ágil, a través del desarrollo de un enfoque híbrido.

## 5. Propuesta metodológica de enfoque “híbrido”

La propuesta metodológica de enfoque “híbrido” para la gestión de proyectos de minería de datos, se basa en el framework de una propuesta diseñada y probada en otro contexto [Cristaldo y otros 2015]. La propuesta metodológica está constituida por fases y componentes. PMI divide un proyecto en tres fases: iniciación, intermedia y de cierre, por otra parte considera en su PMBOK (2013) cinco grupos de procesos: iniciación, planificación, ejecución, control y cierre. [Charvat 2003] establece que cada metodología de proyecto debe contener fases, y aunque varíen según el proyecto o la industria, en general las fases incluyen: concepto, desarrollo, implementación y apoyo. En consecuencia, la metodología propuesta se divide en tres fases: definición y planificación, ejecución y control y, evaluación final y cierre, basados en el ciclo de vida de PMI (2013). El ciclo de vida de la metodología propuesta se muestra en la Figura 1.

Figura 1. Núcleo de la Propuesta Metodológica de enfoque “híbrido”



**FASE 1** – involucra la definición y planificación del proyecto, y se prioriza un enfoque tradicional, ya que se siguen los lineamientos establecidos por el PMBok y la metodología CRISP-DM. Esta fase se divide en dos componentes. La primera, Pre-Proyecto, donde el Jefe de Proyecto debe preparar un documento formal que indique objetivos, alcance y restricciones que permita justificar o no la realización del proyecto y definir las partes interesadas, internas y externas al sector estrechamente vinculado con el proyecto, de forma de lograr comprometer, involucrar y recibir apoyo de todos los interesados. En esta etapa se vincula estrechamente la Gestión de Alcance del PMBok y Comprensión del Negocio de CRISP-DM. En la segunda componente,



Visión/Definición, el Jefe de Proyecto debe determinar planes dentro del proyecto y para cada uno: aprobar y certificar el grado de compromiso de todos los interesados, para con el proyecto; comunicar los planes a todos los interesados; y generar espacio para el intercambio de experiencias. En esta segunda componente se involucra la gestión del tiempo, costos, calidad, riesgos, del PMBOK y la Comprensión del Negocio y la Comprensión de los datos de CRISP-DM.

**FASE 2** – Involucra la ejecución y control del proyecto, y se prioriza un enfoque ágil porque tienen predominio la dinámica y adaptabilidad, características particulares de los proyectos de minería de datos. La presente fase presenta sus componentes solapados entre la metodología SCRUM y las fases de Preparación de los datos, Modelado, Evaluación e Implantación de la metodología CRISP-DM. Esto se debe a que al aplicar SCRUM como principal metodología ágil de gestión, los componentes Preparación y Desarrollo Colaborativo están fuertemente enlazados. En esta fase, el Jefe de Proyecto debe vigilar parte de los requisitos que se corresponden con la formación y crecimiento, tales como: grado de cooperación y comunicación, crecimiento del equipo y desarrollo de habilidades; y la mayoría de los que se corresponden a los procesos internos. Esta fase se divide en tres componentes. La primera, Preparación, donde el Jefe de Proyecto supervisa las etapas de Preparación de los Datos de CRIP-DM y luego definir los planes graduales de lanzamientos correspondientes a cada requerimiento que tendrá como resultado un modelo entregable. La segunda es Desarrollo Colaborativo; la razón del nombre es que la palabra colaboración implica trabajo en equipo, comunicación constante y permanente participación de las partes interesadas, principalmente del o los usuarios representativos. Esto es muy importante para garantizar tanto la culminación de cada modelo entregable como del proyecto en su conjunto. Al cumplir el tiempo asignado para terminar el entregable, se procede a evaluar y probar el mismo (Fase de Evaluación de CRISP-DM). Mientras se ejecuta al *Plan de Lanzamiento*, se efectúan *Revisiones* que tienen como objetivo demostrar concretamente y claramente el progreso del equipo y, recibir retroalimentación de los usuarios sobre el entregable generado (SCRUM).

Al finalizar cada *Plan de Lanzamiento* se genera un reunión de *Retrospectiva* que tiene como objetivos: escuchar distintos puntos de vista dentro del Equipo de Proyecto, identificar colaborativamente las causas de los principales problemas del equipo durante la ejecución del *Plan de Lanzamiento* e idear, consensuar y seleccionar acciones de mejoras concretas que pueda ejecutar el equipo en el próximo *Plan de Lanzamiento* (SCRUM). Recopilar lecciones aprendidas a partir de las reuniones diarias (Ceremonias de SCRUM) donde el *Equipo de Proyecto* le responde al *Director de Proyecto* las tres preguntas clásicas utilizadas en SCRUM: *¿Cuál fue mi avance desde la última reunión diaria?*, *¿En cuáles tareas me comprometo a trabajar hasta la próxima reunión diaria?* y, *¿Qué problemas tengo que me frenan o bloquean?* La tercera componente, Adaptación, Dirección, Monitoreo y Control, debe integrar las componentes, visión/definición, preparación y desarrollo colaborativo. El objetivo de esta componente es garantizar que no se establecerán cambios incontrolados, es decir todos los cambios que se establezcan durante la gestión del proyecto deben tener como objetivo mantener la integridad del proyecto a lo largo de todo el desarrollo del mismo, garantizando que no se realizarán cambios incontrolados.

**FASE 3** – involucra la evaluación final y cierre, y si bien, no hay un preeminencia de un enfoque sobre otro, ya que ambos, tradicional y ágil tienen como objetivos formular

un final claro del proyecto, se plantea el cierre teniendo en cuenta los lineamientos del PMBOK. Esta fase se divide en dos componentes. El cierre, cuyo objetivo es formular un final claro del proyecto, firmando la aceptación del mismo y proporcionando toda la documentación necesaria. La segunda componente, Pos-Proyecto, Mantenimiento y Mejora Continua, tiene como objetivo asegurar que la solución sigue aportando valor para la organización. Para el caso de una mejora significativa o una actualización, la propuesta metodológica de gestión de proyectos con enfoque “híbrido”, vuelve a iniciarse a la Fase 1: Definición y Planificación. La propuesta obedece al requisito de mantener la documentación justa y definir técnicas de difusión de políticas de documentación, organización y coordinación del equipo. Por cada proyecto, el Jefe de Proyecto, crea un Repositorio con el nombre del proyecto, donde se guardan todos los documentos vinculados al proyecto y tienen acceso al mismo todas las partes interesadas. De esta forma, se logra una amplia difusión de la documentación.

## **6. Validación de la propuesta a un caso de estudio real**

Con el fin de validar la propuesta metodológica, se ha desarrollado un proyecto perteneciente al GIBD (Grupo de Investigación de Bases de Datos) siguiendo la misma. Dicho proyecto se encuentra en el marco de una línea de investigación sobre Agenda Setting. El término Agenda Setting hace referencia a la influencia que tienen los medios de comunicación en la fijación de temas en la opinión pública. En dicha línea de investigación se busca realizar un análisis de dominio cruzado para medir los efectos de la instalación de asuntos en la agenda pública tomando como base artículos escritos sobre diferentes temáticas en medios digitales de relevancia para determinar los tópicos que tratan y luego analizar su difusión en redes sociales empleando técnicas de minería de textos y procesamiento de lenguaje natural. Como red social a estudiar se eligió Twitter, por lo que es necesario recolectar y preprocesar tweets en función de distintos parámetros en relación a las noticias de los diarios. El proyecto consiste en el diseño y desarrollo de una aplicación web para la recolección y preprocesamiento de tweets de forma automática según determinados parámetros que funcionan como filtros, los cuales pueden aplicarse al momento de la captura o durante un proceso posterior.

Las tareas se planificaron y ejecutaron siguiendo las fases de la propuesta metodológica.

### **FASE 1 – Definición y planificación**

#### **Pre-Proyecto**

Luego de las reuniones informales en las cuales se acordó el desarrollo del proyecto, se formalizó el mismo en un documento que indica objetivos, alcance y restricciones. Dicho documento justifica la realización del proyecto y define las partes interesadas.

#### **Visión/Definición**

Se realizó un estudio del negocio y de los datos, lo que permitió comprender con mayor detalle el objetivo del proyecto, puntualmente que datos se necesitan capturar y el preprocesamiento que se le debe aplicar a los mismos.

Se determinó el plan del proyecto teniendo en cuenta la metodología a seguir. El diseño del plan del proyecto involucró la planificación de tiempo, costos, recursos humanos y riesgos. Se realizó la comunicación y aprobación del plan con todos los interesados.

También se definieron las herramientas a utilizar y la interacción entre las mismas, haciendo énfasis en la escalabilidad y eficiencia del sistema.

## **FASE 2 – Ejecución y control**

### **Preparación**

El Jefe de Proyecto definió en conjunto con el equipo de trabajo los planes graduales de lanzamientos correspondientes a cada requerimiento, que tienen como resultado un modelo entregable.

### **Desarrollo Colaborativo**

Se realizó el desarrollo de los planes graduales, trabajando en equipo, con comunicación constante y permanente participación de las partes interesadas, principalmente del desarrollador y de los usuarios representativos.

Se mantuvieron reuniones diarias para analizar el avance logrado, las próximas tareas y los problemas actuales. Estas reuniones permitieron recopilar lecciones aprendidas.

Al llegar a cada entregable, se fueron realizando las pruebas y evaluaciones de los mismos, junto con revisiones del trabajo realizado para demostrar el progreso del equipo y recibir retroalimentación de los usuarios sobre el entregable generado.

### **Adaptación, Dirección, Monitoreo y Control**

Como se mencionó anteriormente, se realizó un proceso de dirección y control constante. Esto permitió realizar las adaptaciones necesarias durante el transcurso del proyecto de forma controlada.

## **FASE 3 – Evaluación final y cierre**

### **El cierre**

Se realizó el cierre formal de la primera etapa del proyecto. El proyecto continúa para realizar mejoras a la aplicación desarrollada.

### **Pos-Proyecto, Mantenimiento y Mejora Continua**

Actividades en desarrollo para asegurar que la solución siga aportando valor a la organización. En esta misma línea, se pactó volver a la Fase 1 para plantear una actualización de la aplicación.

## **Documentación, organización y coordinación del equipo**

Se creó un repositorio para almacenar el código y la documentación del proyecto, dando acceso al mismo todas las partes interesadas. La organización del equipo y las actividades se realizó por medio de Trello<sup>2</sup> y Slack<sup>3</sup>. Esto permitió tener los objetivos claros, una correcta gestión del tiempo y una comunicación fluida entre los miembros del equipo.

---

<sup>2</sup> <https://trello.com/>

<sup>3</sup> <https://slack.com/>

## 7. Conclusiones

Los resultados obtenidos permiten afirmar que la propuesta metodológica de enfoque “híbrido” ha cumplido con los objetivos propuestos. Si bien dicha propuesta fue validada solamente para un proyecto de minería de datos en particular, el principal aporte de esta investigación surge a partir de la necesidad de mejorar la gestión de este tipo de proyectos, a través de la colaboración entre las partes interesadas, logrando integrar áreas vinculadas con el proyecto, a través de la definición y ejecución en forma conjunta. Es decir se busca la integración de todo el proyecto en su conjunto. Adicionalmente utiliza pocos recursos, no requiere formación específica por la sencillez que presenta, y permite la generación de sólo la documentación necesaria y la información requerida para asegurar la comunicación y no volver a caer en los mismos errores a partir de las lecciones aprendidas.

## 8. Líneas de trabajo futuras

Se considera como trabajo futuro ampliar la validación de la propuesta metodológica a más proyectos de minería de datos, a fin de incorporar nuevas contribuciones.

## 9. Referencias

KDnuggets, (2014), “Crisp-dm, still the top methodology for analytics, data mining, or data science projects” <http://www.kdnuggets.com/2014/10/crisp-dm-top-methodology-analytics-data-mining-data-science-projects.html>

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C. y Wirth, R. (2010), “CRISP-DM 1.0: Step-by-step data mining guide,” <https://www.the-modeling-agency.com/crisp-dm.pdf>

Olson, D., Delen, D. (2008), “Schematic of SEMMA”. Data mining Techniques, 19. Softcover

Richards, K. (2013), Agile project management: running PRINCE2 projects with DSDM Altern. Norwich: The Stationery Office. 98 p. ISBN-10: 0113310587, ISBN-13: 978-0113310586.

Schwaber, K.; Beedle, M. (2006), Agile Software Development with SCRUM. In: Conchango ISBN: 0130676349.

Schwaber, K.; Sutherland, J. (2011), The Scrum Guide, the Definitive Guide to scrum: The Rules of the Game. [http://www.scrum.org/Portals/0/Documents/ScrumGuides/Scrum\\_Guide.pdf](http://www.scrum.org/Portals/0/Documents/ScrumGuides/Scrum_Guide.pdf), (acceso 14.02.17).

Varajão, J. (2016). Success Management as a PM knowledge area – work-in-progress. Procedia Computer Science. Vol. 100, pp.1095 – 1102.

Colomo-Palacios, R., Casado-Lumbreras, C. Soto-Acosta, P., García-Peñalvo, F.J., Tovar-Caro, E. (2014), Project managers in global software development teams: a study of the effects on productivity and performance, Software Quality. 22 (1)3–19.

Varajão, J. Dominguez, C. Ribeiro, P. Paiva, A. (2014), Critical success aspects in Project management: similarities and differences between the construction and software industry, Tech. Gazette 21 (3) 583–589.

- Lehtinen, T.; Mäntylä, M.; Vanhanen, J.; Itkonen, J.; Lassenius, C. (2014), Perceived causes of software project failures – An analysis of their relationships. *Information and Software Technology* 56, pp. 623–643.
- Ramos, P.; Mota, C. (2014), Perceptions of success and failure factors in information technology projects: a study from Brazilian companies. *Procedia - Social and Behavioral Sciences* 119, pp. 349 – 357.
- Montequin, S.; Fernandez, C.; Fernandez, O.; Balsera, J.V. (2016), Analysis of the Success Factors and Failure Causes in Projects: Comparison of the Spanish Information y Communication Technology (ICT) Sector. *Journal Information Technology Project Management*, Vol. 7(1), pp. 18-31.
- Chow, T.; Chao, D. (2008), A survey of critical success factors in agile software projects. *Journal of systems and software*, Jun. Available: Science Direct, 81(6):961–971.
- Elkadi, H. (2013), Success and failure factors for e-government projects: A case from Egypt. *Egyptian Informatics Journal*, 14, pp. 165–173.
- ElEmam, K.; Koru, A.G. (2008), A replicated survey of IT software project failures Software. *IEEE Software*, vol. 25, pp. 84–90.
- Cristaldo, P.; Ballejos, L.; Alé, M. (2015). Un enfoque híbrido de gestión de proyectos de TICs en el sector público. 44 Jornadas Argentinas de Informática. Universidad Nacional de Rosario, Facultad de Ingeniería, Argentina, ISBN.
- Spundak, M. (2014). Mixed agile/traditional project management methodology – reality or illusion? *Procedia - Social and Behavioral Sciences* 119, pp. 939 – 948.
- Conforto, E. C.; Amaral, D. C. (2008).Evaluating an agile method for planning and controlling innovative projects. *Project Management Journal*, 33(4), pp. 4–14.
- Nogueira, D. (2014). Agile Data Mining: Uma metodologia ágil para o desenvolvimento de projetos de data mining. Faculdade de Engenharia da Universidade do Porto.
- PMBOK Guide. (2013). Project Management Institute. A Guide to the Project Management Body of Knowledge. 5 Ed. NewtownSquare, Pennsylvania 19073-3299 USA. PE: Project Management Institute Inc .ISBN: 978-1-935589-67-9.
- Alnoukari, M.; Alzoabi, Z. Hanna, A. (2008), Applying Adaptive Software Development (ASD) Agile Modeling on Predictive Data mining Applications: ASD-DM Methodology. *Information Technology, ITSIM 2008. International Symposium on (Volume 2)*. doi: 10.1109/ITSIM.2008.4631695
- Santana Do Nascimento, G.; Aparecida de Oliveira, A. (2013), Agile KDD An Agile Knowledge Discovery in Databases Process Model. Federal University of Sergipe São Cristóvão, Brazil gsnascimento@petrobras.com.br Adicinéia Federal University of Sergipe São Cristóvão, Brazil [adicineia@ufs.br](mailto:adicineia@ufs.br)
- Guo, P. (2013), Data Science Workflow: Overview and Challenges. Acceso [acm.org](http://acm.org).
- Ibm (2015), Have you seen ASUM-DM? <https://developer.ibm.com/predictiveanalytics/2015/10/16/have-you-seen-asum-dm/>.
- Way, R. (2013), Model deployment: the moment of truth, Corios Redpaper.

- Mariscal, G.; Marbán, O.; Segovia, J. (2013), Un enfoque Ágil para el Desarrollo de Proyectos de Data Minig, Researchgate,
- Marchx, G. (2014). Big Data and Agile: The Perfect Marriage. Head of Agile Competency Center, EPAM Canadá.
- Lopuszynski, M (2016), Agile Approach to Data Minig Projects. Warsaw Data Science Meetup.
- Batra, D.; Xia, W., VanderMeer, D., y Dutta, K. (2010), Balancing Agile and Structured Development Approaches to Successfully Manage Large Distributed Software Projects: A Case Study from the Cruise Line Industry, Communications of the Association for Information Systems, Vol. 27, Article 21, p. 379-394.
- Fernandez, D. y Fernandez, J. (2008), Agile Project Management: Agilism Versus Traditional Approaches. Journal of Computer Information Systems, Vol. 49, Nro. 2, p. 10-17.
- Karlström, D. y Runeson, P. (2005), Combining Agile Methods with State-Gate Project Management. IEEE Software, Vol. 22, Nro. 3, p. 43-49.
- Stare, A. (2014), Agile Project Management in Product Development Projects. Procedia - Social and Behavioral Sciences 119, 295 – 304
- Sim, J. (2003), Critical Success Factors in Data Mining Project. Tesis University of North Texas.
- Ranjan, J. y Bhatnagar, V (2008), Critical Success Factors for Implementing CRM Using Data Mining. Journal of Knowledge Management Practice, vol9, N3.
- Mérida Sanchez, J.C. (2017), Adaptación de Estándares de Dirección de Proyectos Particularizados para la Minería de Datos. Tesis Universidad de Oviedo.
- Charvat, J. (2003) Project management methodologies: selecting, implementing and supporting methodologies and processes for projects. Hoboken, NJ: Wiley. 264 p