

Propuesta de Ciclo de Vida y Mapa de Actividades para Proyectos de Explotación de Información

Hernán Arboleya

Laboratorio de Investigación y Desarrollo en Ingeniería de Explotación de Información

Grupo Investigación en Sistemas de Información

Departamento Desarrollo Productivo y Tecnológico. Universidad Nacional de Lanús.

Remedios de Escalada, Buenos Aires, Argentina.

arboleya.hernan@yahoo.com.ar, rgm1960@yahoo.com.ar, djhr_1977@yahoo.com.ar

Resumen— La Explotación de la Información es una sub-disciplina informática que le aporta herramientas a la Inteligencia del Negocio [18] para poder transformar la información en conocimiento [15] que sirva de base para llevar a cabo la toma de decisiones dentro de las organizaciones. Dado que este tipo de proyectos se diferencian a los proyectos convencionales de construcción de software y que la metodologías existente para Proyectos de Explotación de Información no contemplan ni tiene en cuenta los riesgos que dentro de los mismos pueden ocurrir, en esta investigación se propone el diseño de un Modelo de Ciclo de Vida específico que tenga un análisis de los riesgos. Posteriormente se plantea la construcción de un Mapa de Actividades teniendo en cuenta el Modelo de Ciclo de Vida desarrollado y los procesos del Modelo de Procesos seleccionado [25].

Palabras Claves – Explotación de Información, Modelo de Ciclo de Vida, Modelo de Procesos, Mapa de Actividades

I. INTRODUCCION

Aquí se presenta una introducción al trabajo de investigación que se ha realizado: se mencionan objetivos (sección A), luego se define el alcance que va a tener (sección B), se evalúan los fundamentos que dan soporte al mismo (sección C) y se presenta la estructura que presenta la investigación (sección D).

A. Objetivos de la Investigación

Esta Investigación tiene como objetivo general la construcción de un Modelo de Ciclo de Vida que contemple los riesgos existentes en cualquier Proyecto de Explotación de Información (E.I.), teniendo como base la metodología CRISP-DM [6] y, por sobre todo, las mejoras que le ha hecho Vanrell a dicha metodología con la presentación de su Modelo de Procesos, agregándole las actividades pertenecientes a los procesos de Administración de Proyectos [25].

Además, se plantea como objetivo general lograr un Mapa de Actividades teniendo como ejes el Modelo de Ciclo de Vida propuesto y el Modelo de Procesos seleccionado en este trabajo [25].

Por otro lado, esta Investigación posee como objetivos específicos:

- Intentar describir a los Modelos de Ciclo de Vida existentes para desarrollo de software convencional;
- Poder describir la metodología CRISP-DM [6] existente para Proyectos de E.I., mencionar sus alcances, ventajas y sus fases;

- Identificar y describir los procesos del Modelo de Procesos seleccionado [25];
- Seleccionar las tareas adecuadas de la metodología CRISP-DM [6] y de las fases convencionales de cualquier sistema software tradicional para la construcción del Modelo de Ciclo de Vida;
- Construir el Modelo de Ciclo de Vida que tenga en cuenta los riesgos y sirva como base para el diseño de estos particulares estilos de proyectos; y
- Diseñar a partir del Modelo desarrollado y del Modelo de Proceso seleccionado [25] y descripto un Mapa de Actividades donde se marcan que actividades del proceso software se va a ejecutar en un determinado tiempo.

B. Alcance

El alcance de este Trabajo de Investigación es establecer las falencias que posee CRISP-DM [6] para la realización de Proyectos de Explotación de Información justificando la creación de un Modelo de Ciclo de Vida que contemple los riesgos que dentro de los proyectos pueden llegar a ocurrir.

Adicionalmente se lleva a cabo la descripción de los Modelos de Ciclo de Vida existentes tanto para Proyectos Convencionales de Desarrollo de Software como así también para Proyectos de Explotación de Información.

Por otro lado se describen y analizan los Modelos de Procesos de interés para este artículo de investigación.

Al mismo tiempo se incluye la construcción de un Mapa de Actividades que tenga como ejes el Modelo de Ciclo de Vida propuesto y el Modelo de Procesos seleccionado dentro del Trabajo.

C. Fundamentos de la Investigación

Dentro de los distintos proyectos que son llevados a cabo por empresas dedicadas al área de tecnologías de la información se encuentra un conjunto denominado Proyectos de Explotación de Información (E.I.). Estos proyectos son especiales, poseen características únicas que lo diferencian de los proyectos comunes de sistemas.

Para los desarrollos de este estilo de proyectos se cuentan actualmente con distintos tipos de técnicas y metodologías que soportan la puesta en marcha de los mismos. A pesar de ello y por la gran importancia que tiene la información y la utilización de la misma para proveer de conocimiento a las empresas, es necesario tener en cuenta los riesgos que en este tipo de proyectos pueden llegar a ocurrir, para lograr estabilidad y que el proyecto no sufra contratiempos.

Es por ello que se detecta la necesidad de desarrollar un Modelo de Ciclo de Vida para este tipo de proyectos, el cual incluya un análisis y trato de los riesgos existentes y resulte eficaz para la empresa que vaya a llevar a cabo este tipo de Proyectos particulares. Además, resulta necesario la construcción de un Mapa de Actividades que determine qué actividades se llevan a cabo y cuales se descartan, y que tenga como ejes el Modelo de Ciclo de Vida a proponer y el Modelo de Procesos para Proyectos de Explotación de Información seleccionado [25]

D. Estructura de la Investigación

En la sección 1 (uno) se presenta una Introducción al trabajo de investigación que se lleva a cabo.

En la sección 2 (dos) se describe brevemente en que constan los Proyectos de E.I., al mismo tiempo se presentan los distintos tipos de Modelos de Ciclo de Vida que se encuentran para Proyectos de Desarrollo Tradicional de Software. Posteriormente se selecciona un Modelo que sirve de base para el desarrollo del Modelo de Ciclo de Vida para Proyectos de E.I. y, se analiza y describe en detalla cada fase de la metodología CRISP-DM [6] que se utiliza para desarrollar Proyectos de este estilo. Al mismo tiempo se analiza COMPETISOFT [19] el cual sirve de base para la construcción del Modelo de Procesos para Proyectos de Explotación de Información que se utiliza en este trabajo [25].

En la sección 3 (tres) se presenta el problema de la presente investigación al poseer la necesidad de tener un Modelo de Ciclo de Vida para estos Proyectos de E.I, el cual sirva para el desarrollo de los mismos pero al mismo tiempo contemple y trate los riesgos.

En la sección 4 (cuatro) se propone una solución para el problema anteriormente descrito. Por lo tanto se lleva a cabo la construcción de un Modelo de Ciclo de Vida que contemple y tenga en cuenta los riesgos existentes en cualquier Proyecto de E.I.. Seguido a esto se construye un Mapa de Actividades que tenga como ejes en primer lugar el Modelo de Ciclo de Vida propuesto y en segundo lugar los procesos presentes en el Modelo de Procesos seleccionado en este trabajo [25].

En el capítulo 5 (cinco) se describen las conclusiones obtenidas a partir del desarrollo de esta Investigación, dando respuesta a los interrogantes de investigación y describiendo las futuras líneas de investigación a desarrollar a partir del mismo.

II. ESTADO DE LA CUESTIÓN

La Explotación de Información consiste en la extracción de conocimiento no-trivial que reside de manera implícita en los datos disponibles en distintas fuentes de información [22]. Dicho conocimiento es previamente desconocido y puede resultar útil para algún proceso de decisión [23]. La información es un patrimonio muy importante para todas las organizaciones y, la utilización y aprovechamiento de la misma de forma correcta y oportuna, le puede producir infinitos beneficios ya que por medios de procesos de explotación de información los miembros de la organización se puede dotar de conocimiento que resulta indispensable para la toma de decisiones, logrando certeza y eficiencia en las mismas.

En la sección A se describen y analizan los Modelos de Ciclo de Vida que son de interés para esta Investigación (tanto el Modelo de Ciclo de Vida en Espiral de Boehm [1]

como la metodología CRISP-DM [6]). Además, en la sección B, se presentan los Modelos de Procesos que colaboran en esta Investigación: en primer lugar se lleva describe el Modelo de Procesos Competisoft [19] (subsección B.1) y en segundo lugar se analiza y describe el Modelo de Procesos seleccionado por este trabajo [25] (subsección B.2.)

A. Modelos de Ciclo de Vida de interés

En esta sección se presentan los Modelos de Ciclo de Vida disponibles para el Desarrollo de Software tradicional y además se describe la metodología CRISP-DM [6] disponible para Proyectos de E.I.

A.1) Modelos de Ciclo de Vida para Desarrollo de Software tradicional

El primer modelo concebido fue el de Desarrollo en Cascada. Este modelo se representa como un simple modelo con forma de cascada de las etapas del software. En este modelo la evaluación del producto software procede a través de una secuencia ordenada de transiciones de una fase a la siguiente según un orden lineal.

Otro Modelo de Ciclo de Vida concebido es el de Prototipado, el cual se utiliza cuando el cliente no tiene una idea muy detallada de lo que necesita, o cuando el ingeniero de software o está muy seguro de la viabilidad de la solución que tiene en mente. Este modelo tiene como objetivo contrarrestar la congelación de requisitos mal comprendidos.

El desarrollo en Espiral es un modelo de ciclo de vida del software definido por primera vez por Barry Boehm en 1986 [6]. Las actividades de este modelo se conforman en una espiral, en la que cada bucle o iteración representa un conjunto de actividades. Este Modelo de Ciclo de Vida tiene en cuenta fuertemente el riesgo que aparece a la hora de desarrollar software. Para ello, se comienza mirando las posibles alternativas de desarrollo, se opta por la de riesgo más asumible y se hace un ciclo de la espiral. Este sistema es muy utilizado en proyectos grandes y complejos como puede ser, por ejemplo, la creación de un Sistema Operativo. Al ser un modelo de Ciclo de Vida orientado a la gestión de riesgo se dice que uno de los aspectos fundamentales de su éxito radica en que el equipo que lo aplique tenga la necesaria experiencia y habilidad para detectar y catalogar correctamente los riesgos. Las principales ventajas encontradas en este modelo de ciclo de vida en comparación con otros modelos son que Su rango de opciones permiten utilizar los modelos de proceso de construcción de software tradicionales, su orientación al riesgo evita muchas dificultades, se presta atención a las opciones que permiten la reutilización de software existente, se centra en la eliminación de errores y alternativas poco atractivas, no establece una diferenciación entre desarrollo de software y mantenimiento del sistema, proporciona un marco estable para desarrollos integrados hardware-software, incorpora objetivos de calidad e, integra el desarrollo con el mantenimiento.

Entre las desventajas se puede mencionar que genera mucho tiempo en el desarrollo del sistema, es un modelo costoso y, requiere experiencia en la identificación de riesgos.

A.2.) Metodología CRISP-DM

Esta metodología define un ciclo de vida de los **proyectos de explotación de Información** que define las principales fases de un proyecto de este tipo. Estas fases son: Entendimiento de Negocios, Entendimiento de los Datos, Preparación de los Datos, Modelado, Evaluación y

Despliegue. A continuación se analizará cada una de esas fases, indicando las actividades y propósitos generales de cada una de las mismas [6]:

- Comprensión del negocio (Objetivos y requerimientos desde una perspectiva no técnica)
 - Establecimiento de los objetivos del negocio (Contexto inicial, objetivos, criterios de éxito)
 - Evaluación de la situación (Inventario de recursos, requerimientos, supuestos, terminologías propias del negocio,...)
 - Establecimiento de los objetivos de la minería de datos (objetivos y criterios de éxito)
 - Generación del plan del proyecto (plan, herramientas, equipo y técnicas)
- Comprensión de los datos (Familiarizarse con los datos teniendo presente los objetivos del negocio)
 - Recopilación inicial de datos
 - Descripción de los datos
 - Exploración de los datos
 - Verificación de calidad de datos
- Preparación de los datos (Obtener la vista minable o dataset)
 - Selección de los datos
 - Limpieza de datos
 - Construcción de datos
 - Integración de datos
 - Formateo de datos
- Modelado (Aplicar las técnicas de minería de datos a los dataset)
 - Selección de la técnica de modelado
 - Diseño de la evaluación
 - Construcción del modelo
 - Evaluación del modelo
- Evaluación (De los modelos de la fase anteriores para determinar si son útiles a las necesidades del negocio)
 - Evaluación de resultados
 - Revisar el proceso
 - Establecimiento de los siguientes pasos o acciones
- Despliegue (Explotar utilidad de los modelos, integrándolos en las tareas de toma de decisiones de la organización)
 - Planificación de despliegue
 - Planificación de la monitorización y del mantenimiento
 - Generación de informe final
 - Revisión del proyecto

Esta metodología para proyectos de minería de datos es muy útil para comprender esta tecnología o extraer ideas para diseñar o revisar métodos de trabajo para proyectos de similares características.

A partir del trabajo realizado por Britos [3], se proponen cinco procesos de explotación de información que pueden ser considerados por CRISP-DM dentro de la etapa de Modelado.

Los procesos de explotación de información definidos son los siguientes:

- Descubrimiento de Reglas de Comportamiento
- Descubrimiento de Grupos
- Ponderación de Interdependencia de Atributos
- Descubrimiento de Reglas de Pertenencia a Grupos

- Ponderación de Reglas de Comportamiento o de la Pertenencia a Grupos

El proceso de Descubrimiento de Reglas se utiliza al querer identificar condiciones para obtener resultados del dominio del problema. Puede ser utilizado para descubrir las características del local más visitado por los clientes o establecer las características de los clientes con alto grado de fidelidad a la marca.

El proceso de Descubrimiento de Grupos es útil en los casos en que necesitemos identificar una partición dentro de la información disponible dentro del dominio de un problema.

El proceso de Ponderación de Interdependencia de Atributos se utiliza cuando se desea identificar los factores con mayor incidencia sobre un determinado resultado de un problema. Son ejemplos aplicables a este proceso la determinación de factores que poseen incidencia sobre las ventas o la individualización de atributos clave que convierten en vendible a un determinado producto.

El proceso de Descubrimiento de Reglas de Pertenencia a Grupos es utilizado cuando se necesita identificar las condiciones de pertenencia a cada una de las clases en una partición desconocida pero que se encuentra presente en la masa de información disponible sobre el dominio del problema. Este tipo de proceso puede ser utilizado para la segmentación etaria de estudiantes y el comportamiento de cada segmento o la determinación de las clases de las llamadas telefónicas en una región y caracterización de cada clase.

Finalmente, el proceso de Ponderación de Reglas de Comportamiento de la Pertenencia a Grupos se utiliza cuando se requiere identificar las condiciones con mayor incidencia sobre la obtención de un determinado resultado en el dominio del problema, ya sea por la mayor medida en la que inciden sobre su comportamiento o las que mejor definen la pertenencia a un grupo.

Como ejemplos de este tipo de proceso se puede citar la identificación del factor dominante que incide en el alza de ventas de un producto dado o el rasgo con mayor presencia en los clientes con alto grado de fidelidad a la marca.

B. Modelos de Procesos de Interés

En esta sección se describen los distintos Modelos de Procesos que son de interés para esta investigación: en primer lugar en la subsección B.1), se analiza y describe al Modelo de Procesos CompetiSoft [19], y posteriormente en la subsección B.2) se presenta el Modelo de Procesos propuesto por Vanrell [25] en su Trabajo de Tesis.

B.1.) Competisofit

Se basa inicialmente en MoProSoft [20] el cual es un modelo de procesos para el desarrollo de software creado por encargo de la Secretaría de Economía Mexicana para servir de base a la norma Mexicana para la Industria de Desarrollo y Mantenimiento de Software, e incorporando el método de evaluación EvalProSoft [21] y una división en etapas de madurez de los procesos similar a las identificadas en CMM o CMMI, entre otras pequeñas modificaciones.

Su propósito es fomentar la estandarización de las operaciones de pequeñas y medianas empresas o departamentos internos de desarrollo, a través de la incorporación de las mejores prácticas en gestión e ingeniería de software, esperando “elevar la capacidad de las organizaciones para ofrecer servicios con calidad y alcanzar niveles internacionales de competitividad” [19].

El modelo busca ser fácil de entender, fácil de aprender, no costoso en su adopción y ser la base para alcanzar evaluaciones exitosas con otros modelos o normas como ISO/IEC 15504 o CMMI.

Este modelo puede ser utilizado tanto por organizaciones que no cuenten con procesos establecidos, ajustándolo de acuerdo a sus necesidades, como por organizaciones que ya poseen procesos establecidos que pueden utilizarlo como punto de referencia para identificar los elementos que les hace falta cubrir.

La estructura del modelo se encuentra dividida en tres categorías [19]: Alta Dirección (DIR), Gerencia (GER) y Operaciones (OPE) reflejando la estructura de una organización. Estas categorías contienen los procesos de gestión de negocio (DIR), gestión de procesos, gestión de proyectos y gestión de recursos (GER) y administración de un proyecto específico, desarrollo de software y mantenimiento de software (OPE).

B.2.) Modelo de Procesos

En su trabajo *Modelo de Proceso de Operación para Proyectos de Explotación de Información* [24], Vanrell propone un marco teórico para la creación de modelo de procesos de operación para proyectos de explotación de información para PYMEs, siguiendo los lineamientos del modelo de procesos para la industria de software (COMPETISOFT [19]).

Así, propone mantener los procesos definidos en COMPETISOFT [19] como Administración de Proyectos y Desarrollo de Proyectos readecuándolos a los proyectos de explotación de información. Para esto elimina las tareas del proceso de desarrollo y las reemplaza por las fases de CRISP-DM [6] que se relacionan con el desarrollo del proyecto.

La reestructuración de estos dos procesos en el desarrollo de proyectos de Explotación de Información brinda mayor claridad a las tareas de administración ya que, como se menciona anteriormente, las tareas de administración que se encuentran mencionadas en las metodologías evaluadas se encuentran dentro de un mismo proceso de desarrollo.

Esta división no se limita a separar las tareas existentes en las metodologías sino que se construye el nuevo proceso basándose en el proceso de administración utilizado para el desarrollo de software clásico, específicamente el definido en el modelo Competisoft, lo cual enriquece el proceso de administración de proyectos [25]. A este proceso se lo adecuó para responder a las exigencias de los proyectos de Explotación de Información.

Las actividades de administración de proyectos son actividades que deben llevarse en paralelo en procesos separados.

Es por esto que el autor decidió separar algunas de las tareas definidas en el ciclo de desarrollo de CRISP-DM [6] para incluirlas en un ciclo diferente que pasará a formar parte del proceso de administración de proyectos.

III. PROBLEMA

En esta sección se incluye en primer lugar el contexto del problema de la presente investigación (sección A), luego se describe el problema en cuestión (sección B) y por último se llevan a cabo preguntas de investigación (sección C).

A. Contexto del Problema

Al intentar llevar adelante diferentes Proyectos de Explotación de Información se utilizan diferentes metodologías y herramientas para poder desarrollarlos. En la actualidad más allá de proyectos de software tradicional cada día más empresas desarrollan proyectos de explotación de información que les permitirá desarrollar conocimiento el cual beneficiará al momento de la toma de decisiones, por lo que es necesario metodologías y herramientas que puedan estimarlos, analizando riesgos existentes en los mismos y mejorando su calidad. Este artículo de investigación apunta a eso.

B. Descripción del Problema

Teniendo en cuenta y, aceptando que los proyectos de Explotación de Información poseen características particulares y distintas a los proyectos convencionales y tradicionales de software, es necesario contar con un Modelo de Ciclo de Vida para este tipo de Proyectos que se adapte a las necesidades y requerimientos de los mismos.

Teniendo como base las metodologías y modelos de ciclos de vida para proyectos tradicionales de software y contando con la metodología CRISP-DM [6] para Proyectos de Explotación de Información, se ha encontrado que no hay ninguna herramienta o metodología que tenga en cuenta los posibles riesgos que pueden tener este tipo de proyectos.

Teniendo en cuenta este problema de no contar con un Modelo de Ciclo de Vida que sirva para Proyectos de Explotación de Información (teniendo como base las fases descritas en CRISP-DM [6] necesarias para llevar a cabo este tipo de Proyectos), que tenga en cuenta y analice los riesgos que pueden en ellos ocurrir; se detecta que es necesario construir un Modelo de Ciclo de Vida que tenga en cuenta los riesgos y sirva como base para el diseño de estos particulares estilos de proyectos y al mismo tiempo, diseñar a partir del Modelo desarrollado [25] y del Modelo de Proceso seleccionado y descripto, un Mapa de Actividades donde se marcan qué actividades del proceso software se va a ejecutar en un determinado tiempo.

C. Preguntas de Investigación

Teniendo en cuenta el problema planteado surgen los siguientes interrogantes que se intentan responder a lo largo de este trabajo:

- ¿Es posible construir un Modelo de Ciclo de Vida para Proyectos de Explotación de Información que además de las tareas necesarias para los mismos tenga en cuenta los riesgos?
- ¿Se podrá construir un Mapa de actividades teniendo como ejes el Modelo de Ciclo de Vida a proponer con el Modelo de Procesos descripto en la sección anterior que tenga en cuenta tanto los procesos de Administración como los de Desarrollo?

Luego de presentar estos interrogantes y teniendo en cuenta la problemática descrita en este capítulo, en el siguiente capítulo se presenta la propuesta de un Modelo de Ciclo de Vida adaptado a los Proyectos de Explotación de Información que solucione esta problemática existente y, posteriormente, se lleva a cabo la construcción de un Mapa de Actividades teniendo como ejes el Modelo de Procesos propuesto por Vanrell [25] y el Modelo de Ciclo de Vida desarrollado en el presente Trabajo.

IV. SOLUCIÓN

Esta sección está compuesta en primer lugar por la propuesta de un Modelo de Ciclo de Vida para el problema en cuestión (sección A) y, en segundo lugar, por la construcción de un Mapa de Actividades teniendo como ejes el Modelo de Ciclo de Vida propuesto y el Modelo de Procesos seleccionado (sección B).

A. Modelo de Ciclo de Vida

Como ya se ha mencionado, se ha tomado la base del Modelo de Ciclo en Espiral de Boehm readaptando las fases a las actividades descritas por el Modelo de Procesos desarrollado por Vanrell [25] para Proyectos de Explotación de Información, el cual posee un valor agregado en comparación a la metodología existente para este tipo de Proyectos como es CRIS-DM [6]. Estas actividades corresponden a las 6 fases de dicha metodología: Entendimiento del Negocio, Entendimiento de los Datos, Preparación de los Datos, Modelado, Evaluación y Despliegue.

Básicamente, en cada vuelta del espiral se realizan actividades correspondientes al Modelo de Procesos correspondiente para Proyectos de E.I. [25] luego se planea la próxima iteración, se lleva a cabo un análisis de riesgo y por último se realiza la presentación y desarrollo de un prototipo a partir de las fases de Desarrollo en sí del Modelo de Ciclo de Vida. Estas actividades se realizan hasta cubrir todas las fases de la metodología para Proyectos de Explotación de Información.

Entre los ejes Evaluación de alternativas, identificación y resolución de riesgos (análisis de riesgos) y el eje de Desarrollo, verificación del producto del próximo nivel se desarrollan distintos prototipos y simulaciones, modelos y benchmark, tanto para ir viendo con el cliente si lo entendido por el grupo de desarrollo es lo correcto y se está avanzando adecuadamente en el transcurso del proyecto o si hay que modificar algo, como también para ir analizando el rendimiento del sistema y/o de los componentes desarrollados hasta el momento. Teniendo en cuenta lo dicho, en la figura 1 se puede observar el Modelo de Ciclo de Vida desarrollado para cubrir las necesidades planteadas en la sección 3. Este modelo planteado incluye más vueltas al espiral para abarcar todos los procesos del Modelo de Procesos que derivan de las fases de CRISP-DM y para garantizar la no aparición de riesgos y el desarrollo tranquilo del Proyecto.

Se ha determinado que es necesario ilustrar en el Modelo de Ciclo de Vida en Espiral propuesto las tareas y subprocesos correspondientes a la Administración del Proyecto [25]. Estos subprocesos abarcan todo el largo del desarrollo del Proyecto y dan soporte a las actividades de desarrollo propiamente dichas, manteniendo “los Procesos de Desarrollo en movimiento y corregir aquellos desvíos que se produzcan con el fin de lograr una finalización exitosa del proyecto” [25]. Este proceso de Administración que se ubica a lo largo del Modelo de Ciclo de Vida propuesto tiene la función de recolectar información que le sea útil a los procesos de Desarrollo para aumentar la calidad de los mismos permitiendo realizar ajustes y mantener un estándar en la realización de distintos proyectos.

En la *primera vuelta* del espiral propuesto, en el eje de desarrollo (verificación del producto del próximo nivel), se lleva a cabo la tarea de determinar los Objetivos del Negocio.

Mediante esta actividad, se describen los objetivos primarios del cliente, desde una perspectiva del negocio. Además de los objetivos del negocio primario, allí hay típicamente otras preguntas de negocio relacionadas con lo que al cliente le gustaría administrar. Por otro lado, se describen los criterios del éxito del negocio, los resultados útiles que se desean para el proyecto y bajo qué circunstancias se va a dar como aprobado el desarrollo del proyecto a realizar. Pasando al próximo eje del espiral, se realiza la Planificación de Requisitos y Planificación del Ciclo de Vida del Proyecto de Explotación de Información a realizar.

Teniendo en cuenta la primera iteración, se lleva a cabo un análisis de riesgo (AR1), donde se lleva a cabo el estudio y previsión de futuros riesgos que puedan llegar a ocurrir teniendo en cuenta lo descrito, modelado y desarrollado hasta el momento.

En la *segunda vuelta* del espiral, situándonos en el eje de Desarrollo, Adaptación y Pruebas, se continúan con las actividades de la fase de **Entendimiento del Negocio** de la metodología CRISP-DM [6] y tomadas en los Procesos de Desarrollo de Proyectos del Modelo de Procesos de Vanrell [25] para Proyectos de Explotación de Información. En primer lugar se lleva a cabo la Evaluación de la Situación, en la cual se realiza una investigación más detallada sobre los recursos, restricciones, presunciones, y otros factores que deberían ser considerados en la determinación del objetivo de análisis de datos y el plan de proyecto.

Las salidas de esta actividad serán Inventarios de recursos, Requerimientos, presunciones y restricciones, Riesgos y contingencias, Terminología y Costos y beneficios.

Posteriormente a esta actividad y siguiendo con el análisis de la segunda vuelta del espiral propuesto se tiene la Determinación de los Objetivos del Proyecto de Explotación de Información en sí. A diferencia de los Objetivos del Negocio planteados en la primera vuelta, en esta actividad se declaran los objetivos del proyecto en términos técnicos.

Como salida a esta actividad se obtienen los Objetivos del Proyecto de Explotación de Información (descripción de las salidas intencionadas del proyecto que permiten el logro de los objetivos de negocio) y los Criterios de Éxito del Proyecto de Explotación de Información (definición de los criterios de un resultado exitoso para el proyecto en términos técnicos).

Luego de la descripción de los Objetivos del Proyecto de Explotación de Información se debe desarrollar un Plan de Proyecto, en el cual se especifican los pasos a ser realizados durante el resto del proyecto, incluyendo la selección inicial de herramientas y técnicas.

Las salidas de esta actividad son el Plan de Proyecto en sí, donde se listan las etapas a ser ejecutadas en el proyecto, juntos con su duración, recursos requeridos, entradas, salidas, y dependencias; y además la Evaluación de Herramientas y Técnicas.

Finalizando con las tareas de la segunda vuelta del espiral y siguiendo situados en el eje de Desarrollo, Adaptación y Pruebas, se realiza al final una Validación de los Objetivos, donde se analizan los objetivos planteados para el Proyecto de Explotación de Información y se determina si cumple técnica y funcionalmente con lo que el cliente ha solicitado. Superando esta actividad, se pasa al eje de Planeamiento de la Siguiete Fase, donde se realiza la Planificación de la Selección de Datos.

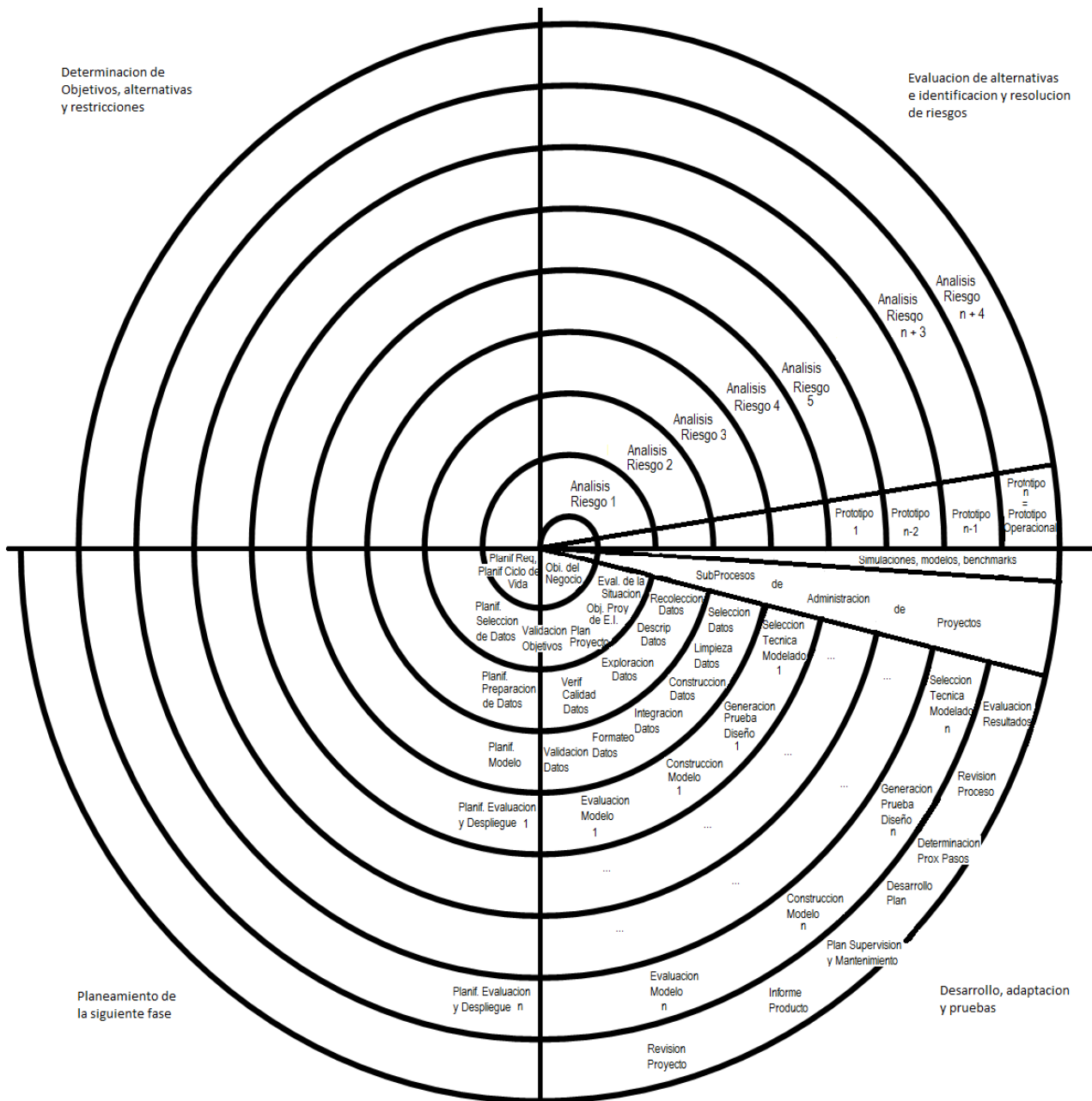


Fig. 1. Modelo de Ciclo de Vida en Espiral propuesto

Teniendo en cuenta la segunda iteración, se lleva a cabo un análisis de riesgo (AR2), donde se lleva a cabo el estudio y previsión de futuros riesgos que puedan llegar a ocurrir teniendo en cuenta lo descrito, modelado y desarrollado hasta el momento.

Luego de la segunda iteración, del Análisis de Riesgo 2 (AR2) y teniendo como referencia los procesos que corren en paralelo y a lo largo de todo el Modelo de Ciclo de Vida como son las Simulaciones, los Modelos y los Benchmarks, como así también los Subprocesos de Administración de Proyectos [25] se pasa a la *tercera iteración* partiendo del Eje de Desarrollo, Adaptación y Pruebas.

Durante esta iteración se llevan a cabo tareas correspondientes a la Fase de **Entendimiento de los Datos** de la Metodología CRISP-DM [6] y de los Procesos de Desarrollo de Proyectos del Modelo de Procesos [25] para Proyectos de Explotación de Información.

En primer lugar se lleva a cabo la Recolección de Datos primarios o iniciales, donde se accede a los datos listados en los Recursos del Proyecto y donde se incluye carga de datos, si es necesario para la comprensión de los datos.

Como salida perteneciente a esta actividad se obtiene el Informe de Colección de Datos Inicial, donde quedan listados el conjunto de datos obtenido junto con sus posiciones, los métodos usados para adquirirlos, algunos de los problemas encontrados y algunas de las resoluciones alcanzadas.

Posteriormente se realiza la actividad de Descripción de los Datos donde se examina las propiedades "gruesas" o "superficiales" de los datos e informe adquiridos en los resultados.

La salida de esta actividad es un Informe de descripción de datos donde se describen los datos que han sido adquiridos, incluyendo el formato de los datos, la cantidad de datos (por ejemplo, el número de registros y campos en cada tabla), los

identificadores de los campos, y cualquier otro rasgo superficial que ha sido descubierto.

En tercer lugar se realiza la Exploración de los Datos donde se dirigen interrogantes de minería de datos usando preguntas, visualización, y técnicas de reporte incluyendo la distribución de atributos claves relacionados entre pares o pequeños números de atributos, los resultados de simples agregaciones, las propiedades de las subpoblaciones significativas, y análisis estadísticos simples. Como salida a esta actividad se obtiene un Informe de Exploración de Datos donde se describen los resultados de esta tarea, incluyendo primeras conclusiones o hipótesis iniciales y su impacto sobre el resto del proyecto.

Por último se lleva a cabo una Verificación de la Calidad de los Datos, donde se examina la calidad de los datos, realizando preguntas como: ¿Los datos están completos? ¿Son correctos, o estos contienen errores y, si hay errores, que tan comunes son estos? ¿Hay valores omitidos en los datos? Si es así, ¿cómo se representan estos, donde ocurre esto, y que tan comunes son estos? [6]

Al finalizar las actividades de la tercera vuelta en el espiral posicionados en el eje de Desarrollo, Adaptación y Pruebas, se pasa al eje de Planeamiento de la siguiente fase llevando a cabo la Planificación de Preparación de los Datos.

Teniendo en cuenta la tercera iteración y situados en el eje de Evaluación de Alternativas e identificación y resolución de riesgos, se lleva a cabo un análisis de riesgo (AR3), donde se lleva a cabo el estudio y previsión de futuros riesgos que puedan llegar a ocurrir teniendo en cuenta lo descripto, modelado y desarrollado hasta el momento

En la *cuarta vuelta* del espiral propuesto se llevan a cabo las tareas pertenecientes a la fase de **Preparación de los Datos** de la metodología CRISP-DM [6] y de los Procesos de Desarrollo del Modelo de Procesos [25] para Proyectos de Explotación de Información.

En primer lugar se realiza la Selección de Datos, donde se decide que datos serán usados para el análisis. Los criterios incluyen la importancia a los objetivos de la minería de datos, la calidad, y las restricciones técnicas como límites sobre el volumen de datos o los tipos de datos.

Como salida correspondiente a dicha tarea se obtiene un razonamiento para inclusión o exclusión de datos y los motivos de esa determinación [6].

Posteriormente se lleva a cabo la Limpieza de Datos, que implica la selección de los subconjuntos de datos limpios, la inserción de datos por defectos adecuados, o técnicas más ambiciosas tales como la estimación de datos faltantes mediante modelado.

La salida asociada a esta tarea es un Informe de la Limpieza de Datos, donde se describe que decisiones y acciones fueron tomadas para dirigir los problemas de calidad de datos informados.

A continuación se construyen las operaciones de preparación de datos tales como la producción de atributos derivados o el ingreso de nuevos registros, o la transformación de valores para atributos existentes en la actividad Construcción de Datos.

Las salidas correspondientes a esta actividad son Atributos derivados (atributos nuevos que son construidos de uno o más atributos existentes en el mismo registro) y los Registros generados (creación de registros completamente nuevos).

Luego, en la actividad Integración de Datos, se definen los métodos por el cual la información es combinada de múltiples tablas o registros para crear nuevos registros o valores. Esta actividad tiene como salida la Combinación de Datos, en la cual se realiza la unión simultánea de dos o más tablas que tienen información diferente sobre el mismo objeto.

En cuarto lugar, se lleva a cabo el Formateo de Datos, en el cual se realizan modificaciones principalmente sintácticas hechas a los datos que no cambian su significado, pero podría ser requerido por la herramienta de modelado. A partir de dicha actividad se obtienen los Datos formateados.

Por último y para finalizar las actividades pertenecientes a la fase de Preparación de los Datos, se lleva a cabo la Validación de los mismos para determinar si los mismos son realmente correctos y servirán para el desarrollo del Proyecto de Explotación de Información.

Pasando al siguiente eje, se realiza la Planificación del Modelo y se lleva a cabo un análisis de riesgo (AR4), donde se lleva a cabo el estudio y previsión de futuros riesgos que puedan llegar a ocurrir teniendo en cuenta lo descripto, modelado y desarrollado hasta el momento

A partir de aquí, en la *quinta vuelta* del espiral, comienza las actividades pertenecientes a la fase de **Modelado** correspondiente a la metodología CRISP-DM [6] y de los Procesos de Desarrollo del Modelo de Procesos [25] para Proyectos de Explotación de Información. En esta fase se producen iteraciones, ya que se construirán n cantidad de modelos y prototipos hasta alcanzar el modelo que más se adecue a las pretensiones y requerimientos que el cliente ha brindado.

En primer lugar se lleva a cabo la Selección de Técnica de Modelado 1, donde se elige la técnica de modelado específico, por ejemplo, un árbol decisión construido con C4.5. Como salida a esta actividad se obtiene la Técnica de Modelado seleccionada y Presunciones del Modelado.

A continuación de la Selección de la Técnica de Modelado 1 se realiza la Generación de Prueba de Diseño 1, donde se genera un procedimiento o el mecanismo para probar la calidad y validez del modelo. A partir de dicha actividad se genera la Prueba de Diseño, donde se describe el plan intencionado para el entrenamiento, la prueba, y la evaluación de los modelos.

Posteriormente, se realiza la primer Construcción del Modelo, donde se ejecuta la herramienta de modelado sobre el conjunto de datos preparados para crear uno o más modelos. Al finalizar dicha actividad se obtienen los parámetros de ajustes (Lista de los parámetros y sus valores escogidos, también con el razonamiento para elegir los parámetros de ajustes), el/los modelo/s en sí (modelos reales producidos por la herramienta de modelado, no un informe) y la descripción del/os modelo/s presentado/s (descripción de los modelos obtenidos e informe sobre la interpretación de los modelos y documentar cualquier dificultad encontrada con sus significados).

Al finalizar las 3 actividades mencionadas anteriormente, se lleva a cabo la primer Evaluación del Modelo (Evaluación del Modelo 1) donde se interpreta los modelos según el conocimiento de dominio, los criterios de éxitos de minería de datos, y el diseño de prueba deseado, se juzga el éxito de la aplicación del modelado y se descubre técnicas más adecuadas. Se evalúa los modelos según los criterios de evaluación. La salida es la Evaluación del/os Modelo/s en sí

donde se resumen los resultados de esta tarea, se enlista las calidades de los modelos generados y se clasifica su calidad en relación con cada otro; y se obtienen los parámetros de ajustes revisados, donde se revisan los parámetros de ajuste para la siguiente corrida en la tarea de Construcción del Modelo.

Al finalizar este cuarto de vuelta, se realiza la Planificación de Evaluación y Despliegue 1 y en el eje de Evaluación de alternativas e identificación y resolución de Riesgos se lleva a cabo el Análisis de Riesgo número 5 (AR5).

Luego del AR5, se realiza el Prototipo 1, donde se lleva a cabo la demostración del Modelo construido y se analiza si cumple o no con lo que el cliente solicita.

Dependiendo de la satisfacción o cumplimiento del Modelo, se realizan 2 vueltas más al espiral, con las mismas actividades descritas anteriormente, hasta que finalmente el Modelo construido y desarrollado sea el adecuado para cumplir con los Objetivos planteados en un primer momento en el Proyecto.

Esas vueltas en blanco son previas a la conformación del Modelo final y aprobado y cada una de ellas es acompañada por un Análisis de Riesgo y la construcción de un Prototipo. Además, los subprocesos de Administración de Proyectos [25] acompañan estas tareas en paralelo.

En la antepenúltima vuelta del espiral se lleva a cabo la última construcción del Modelo (Construcción Modelo n), donde n es la variable que indica cuantas veces se han llevado a cabo las actividades de la fase de Modelado hasta finalmente cumplir con los establecidos.

Por último se pasa al eje de Planeamiento de la Siguierte fase para realizar la Planificación y Despliegue final (n), se pasa al siguiente eje realizando el Análisis de Riesgo $n+4$ (ya que previo al Prototipo 1 y AR5 se han llevado a cabo 4 análisis de Riesgo) y se construye el Prototipo final o Prototipo Operacional, que muestra el Modelo final o definitivo.

En la *última vuelta* del espiral propuesto se llevan a cabo las actividades de las fases de **Evaluación y Despliegue** de la metodología CRISP-DM [6] y de los Procesos de Desarrollo del Modelo de Procesos [25] para Proyectos de Explotación de Información.

En primer lugar se Evalúan los Resultados, donde se analiza el grado al que el modelo responde a los objetivos de negocio y procura determinar si hay alguna decisión de negocio por el que este modelo es deficiente. Además verifica otros resultados generados por la minería de datos.

La salida a esta actividad es la Evaluación de los Resultados de la minería de datos en lo que concierne a criterios del éxito del negocio, y por otro lado, los Modelos finalmente aprobados.

Por otro lado, en segundo lugar, se realiza la Revisión del Proceso, donde los modelos resultantes pasan a ser satisfactorios y a satisfacer las necesidades de negocio y se lleva a cabo una revisión más cuidadosa de los compromisos de la minería de datos para determinar si hay cualquier factor importante o tarea que de algún modo ha sido pasada por alto. Como salida a esta actividad se tiene la Revisión del Proceso propiamente dicha, donde se resume la revisión de proceso y destaca las actividades que han sido omitidas y/o aquellas que deberían ser repetidas.

Luego de la Revisión del Proceso, se Determinan los Próximos Pasos, donde según los resultados de la evaluación y la revisión de proceso, el equipo de proyecto decide como

proceder. En esta instancia el equipo decide si hay que terminar este proyecto y tomar medidas sobre el desarrollo si es apropiado, tanto iniciar más iteraciones, o comenzar nuevos proyectos de minería de datos.

Las salidas a esta actividad es una Lista de Posibles Acciones y Decisión (se describe la decisión en cuanto a cómo proceder, junto con el razonamiento).

Pasando a las actividades de Despliegue, en primer lugar se lleva a cabo el Desarrollo del Plan, donde de acuerdo al desarrollo de los resultados de minería de datos en el negocio, esta tarea toma los resultados de la evaluación y determina una estrategia para el desarrollo. La salida de esta actividad es el Plan propiamente dicho.

Luego del Plan descrito anteriormente se realiza un Plan de Supervisión y Mantenimiento, ya que son cuestiones importantes si los resultados de minería de datos son parte del negocio cotidiano y de su ambiente.

Al finalizar esta tarea se obtiene el Plan propiamente dicho que resume la estrategia de supervisión y mantenimiento incluyendo los pasos necesarios y como realizarlos.

Posteriormente se pasa a la actividad de Producción de Informe Final, donde el líder del proyecto y su equipo sobrescribe un informe final. Según el plan de desarrollo, este informe puede ser sólo un resumen del proyecto y sus experiencias o esto puede ser una presentación final y comprensiva de los resultados de minería de datos.

La salida a esta tarea es el Informe definitivo y la Presentación final, donde puede haber una reunión en la conclusión del proyecto en el que los resultados son presentados verbalmente al cliente.

Por último, se lleva a cabo una Revisión General del Proyecto, donde se evalúa lo que fue correcto y lo que se equivocó, lo que fue bien hecho y lo que necesita para ser mejorado. La salida correspondiente es la Documentación de la experiencia de realización del Proyecto, donde se resumen las experiencias importantes ganadas durante el proyecto.

B. Mapa de Actividades

Se presentan en las figuras 2 y 3 el Mapa de Actividades construido y propuesto, teniendo como ejes en primer lugar el Modelo de Ciclo de Vida propuesto que tuvo como base el Modelo de Ciclo de Vida en Espiral que propuso Boehm y las fases del Modelo de Procesos desarrollado por Vanrell para Proyectos de Explotación de Información, agregándole procesos de Administración de Proyectos los cuales la metodología existente, CRISP-DM [6], no poseía, y en segundo lugar el Modelo de Procesos de Vanrell propiamente dicho [25].

Como se puede observar, se dividen las actividades de los Procesos de Administración de Proyectos y los Procesos de Desarrollo de Proyectos, tal como describe Vanrell en su trabajo [25]. Por otro lado, cada columna es una fase o actividad que se lleva a cabo en el Modelo de Ciclo de Vida propuesto en esta investigación, marcando la división por vuelta.

Las actividades de Administración de Proyectos no están ocupados del desarrollo propiamente dicho del Proyectos de Explotación de Información, sino que son procesos en paralelo y que tienen la función de recolectar información que le sea útil a los procesos de Desarrollo para aumentar la calidad de los mismos permitiendo realizar ajustes y mantener un estándar en la realización de distintos proyectos llevando a

cabo la conversión de procesos caóticos y aislados en procesos repetitivos y controlados [25]. Por ello mismo se puede ver que las actividades de Administración de Proyectos abarcan todo el Ciclo de Vida, soportando y manteniendo las actividades de Desarrollo de Proyectos del Modelo de Vanrell.

En la primera vuelta del espiral la única fase que se lleva a cabo es la de Objetivos del Negocio, donde se describen los objetivos primarios del cliente, desde una perspectiva del negocio. Teniendo en cuenta los procesos y tareas que se definen en el Modelo de Procesos para Proyectos de Explotación de Información propuesto por Vanrell [25] se puede decir que dentro de los Procesos de Administración de Proyectos se encuentran la Planificación/Entendimiento del Negocio, la Realización, la Evaluación y Control y el Cierre/Entrega.

En primer lugar, dentro de la Planificación /Entendimiento del Negocio, la única actividad que se realiza es la de Entendimiento del Negocio. Esto se debe a que la salida correspondiente a esta tarea es la Base de Conocimiento del Negocio, cuyas actividades previas son la de por medio de organigramas, redes de expertos e informes relacionados con el negocio, desarrollar gráficos organizacionales, identificar personas claves, sponsor interno, si existe comité de gobierno, las unidades del negocio que serán impactadas, etc. Por otro lado, otra salida correspondiente a esta tarea son los Objetivos del Negocio propiamente dicho, describiendo el problema en términos generales, evaluando el estado actual del proyecto, especificando los requerimientos y preguntas del negocio, determinando los beneficios esperados en términos de negocio. Por último, dentro de aquí se determinan los Criterios del Éxito del negocio y se identifican quienes determinan dichos criterios. Por otro lado, no se han marcado las actividades de Definir el proceso específico basado en la descripción del proyecto y el proceso de desarrollo y mantenimiento; Definir el protocolo de entrega con el cliente; Definir ciclos y actividades con base en la descripción del proyecto y en el proceso específico; Determinar tiempo estimado para cada actividad; Elaborar plan de adquisiciones y capacitación; Establecer el equipo de trabajo; Establecer el calendario de actividades; Calcular el costo estimado del proyecto; Evaluación de la situación; Producir un Plan de Proyecto; Producir un Plan de Desarrollo; y Formalizar el inicio de un nuevo ciclo del proyecto dado que se ha entendido que dichas actividades no corresponden al Negocio en sí, sino mas a los Objetivos del Proyecto de Explotación de Información en sí mismo.

En la tabla 1 se presentan las descripciones de los códigos de las fases del Modelo de Ciclo de Vida propuesto, utilizados en el Mapa de Actividades.

Por otro lado, dentro de la Realización, se ha definido que no se lleva a cabo ninguna actividad en la Fase de Objetivos del Negocio, ya que son actividades más relacionadas con el Proyecto de Explotación en sí, su realización y con el seguimiento y monitoreo que se realiza durante el desarrollo del Proyecto.

En tercer lugar, dentro de Evaluación y Control, se ha determinado que la única actividad que se hará en esta Fase es la de Generar el reporte de seguimiento del proyecto, dado que dentro de la misma se evalúan si los objetivos del negocio y del uso del modelo puede cambiar a lo largo del tiempo y se documenta lo más completamente posible el problema inicial que el modelo está intentando resolver. Por otro lado, las

tareas de Evaluar el cumplimiento del plan de proyecto y plan de desarrollo; y Analizar y controlar los riesgos no se contemplan en esta Fase, dado que son más ligados al Plan del Proyecto de Explotación de Información en sí que del Negocio.

TABLA I. CÓDIGOS DEL MAPA DE ACTIVIDADES

Código	Descripción
O.N.	Objetivos del Negocio
E.Sit.	Evaluación de la Situación
O.P.	Objetivos del Proyecto de E.I.
P.P.	Plan De Proyecto
Val.O.	Validación de Objetivos
R.D.	Recolección de Datos
D.D.	Descripción de Datos
E.D.	Exploración de Datos
Ver.CD.	Verificación Calidad de Datos
S.D.	Selección de Datos
L.D.	Limpieza de Datos
C.D.	Construcción de Datos
I.D.	Integración de Datos
F.D.	Formateo de Datos
Val.D.	Validación Datos
S.T.M.	Selección Técnica Modelado
G.P.D.	Generación Prueba de Diseño
C.M.	Construcción Modelo
E.M.	Evaluación Modelo
E.R.	Evaluación Resultados
R.P.	Revisión Proceso
D.P.P.	Determinación Próximos Pasos
D.P.	Desarrollo Plan
P.S.M.	Plan Supervisión y Mantenimiento
I.P.	Informe Producto
R.Py.	Revisión Proyecto

Por último, dentro del Cierre / Entrega, se definió que no se lleva ninguna tarea dentro de esta Fase, dado que estas actividades tienen más que ver con las últimas fases del Modelo de Ciclo de Vida, donde se lleva a cabo el Cierre y Entrega del Proyecto de Explotación de Información desarrollado.

Dentro de los Procesos de Desarrollo de Proyecto definido por Vanrell en su Modelo de Procesos [25], se encuentran Entendimiento del Negocio, Entendimiento de los Datos, Preparación de los Datos, Modelado, Evaluación y Entrega.

En primer lugar, dentro del subproceso Entendimiento del Negocio, se ha definido que su actividad Determinar las metas del Data Mining ha de llevarse en la Fase Objetivos del Negocio del Modelo de Ciclo de Vida, dado que se trasladan los interrogantes del negocio a metas del proyecto de Explotación de Información y se realiza informe indicando una traducción de los interrogantes a metas de negocio.

Por otro lado, dentro de los demás subprocesos correspondiente a los procesos de Desarrollo del Proyecto, se ha determinado que no se lleva a cabo ninguna tarea dentro de esta fase, dado que las mismas tienen más que ver con fases siguientes y más avanzadas del Modelo de Ciclo de Vida propuesto.

En la segunda vuelta del Modelo en Espiral cuenta con las siguientes fases: Evaluación de la Situación, Objetivos del Proyecto de Explotación de Información, Plan del Proyecto y Validación de Objetivos.

Empezando a analizar en detalle a la fase de Evaluación de la Situación, se puede determinar que dentro del subproceso Planificación/Entendimiento del Negocio correspondiente a los Procesos de Administración de Proyectos, las únicas actividades que se llevan a cabo son la de Entendimiento del

Negocio y Evaluación de la Situación, ya que se entiende que dentro de ellas tareas tales como evaluar el estado actual del proyecto, identificar necesidades y expectativas de los usuarios, identificar los recursos hardware disponible, establecer si los planes de mantenimiento de hardware entran en conflicto con la disponibilidad de hardware necesaria para el proyecto de Explotación de Información, identificar las fuentes de datos, chequear las disponibilidades de herramientas y técnicas, identificar analistas de mercado, entre otras, resultan necesarias realizarlas en dicha fase

Siguiendo con el análisis correspondiente a la fase de Evaluación de la Situación, se dictamina que las actividades pertenecientes tanto al SubProceso Realización, Evaluación y Control, y Cierre /Entrega, no se llevan a cabo dentro de esta fase puesto que la misma es prematura y dichas actividades corresponden cuando el Proyecto está más avanzado y se encuentra en etapas posteriores.

Dentro de los Procesos de Desarrollo de Proyectos, en el único Proceso que se llevan a cabo actividades en esta Fase es en el de Entendimiento del Negocio.

Dentro de este SubProceso se lleva a cabo la tarea de Determinar las metas del Proyecto de Explotación de Información. Tareas de SubProcesos siguientes tales como Entendimiento de los Datos, Preparación de los Datos, Modelado, Evaluación y Entrega no se llevan a cabo en esta Fase del Modelo de Ciclo de Vida ya que es un fase inicial y dichos procesos corresponden a tareas posteriores.

Siguiendo con el análisis de la segunda vuelta del Modelo de Ciclo de Vida en Espiral propuesto, la fase que se analiza a continuación es la de Objetivos del Proyecto de Explotación de Información.

Dentro de los Procesos de Administración de Proyectos, las únicas tareas que se llevan a cabo en el SubProceso Planificación/Entendimiento del Negocio son las de Entendimiento de Negocio y Definir el proceso específico basado en la descripción del proyecto y el proceso de desarrollo y mantenimiento. Se ha definido esto ya que tareas tales como evaluar la descripción del proyecto, evaluar el proceso de desarrollo y mantenimiento, describir el problema en términos generales, clasificar prerequisites del proyecto, especificar los beneficios esperados y definir el proceso específico a utilizar, son necesarias desarrollarlas en esta fase para dejar en claro los objetivos propios del Proyecto.

Por otro lado, las tareas correspondiente al SubProceso de Realización no se realiza ninguna salvo las de recibir y analizar las solicitudes de cambio del cliente y realizar reuniones con el equipo de trabajo y cliente para reportar avances y tomar acuerdos, ya que recibir solicitudes de cambio del cliente y las Realizaciones de reuniones con el cliente pueden afectar a los Objetivos establecidos para el Proyecto de Explotación de Información.

Dentro del SubProceso de Evaluación y Control, la actividad que se lleva a cabo es la de generar el reporte de seguimiento del cliente, ya que se analiza que cosas pueden cambiar a futuro. Por otro lado, dentro del SubProceso de Cierre/Entrega, no se lleva a cabo ninguna actividad ya que se entiende que corresponden a fases posteriores.

Siguiendo con el análisis de esta fase, dentro de los Procesos de Desarrollo de Proyecto, el único SubProceso que se relaciona con esta fase es la de Entendimiento del Negocio. Es así ya que determinar las metas del proyecto de Explotación de Información especifica los criterios de

valoración de los modelos, determina los benchmarks para la evaluación de criterios y especifica criterios a valoración de criterios subjetivos, entre otras tareas.

Las otras actividades que corresponden a los demás SubProcesos de Desarrollo de Proyectos, se entiende que no corresponden a esta fase ya que son actividades que se deben desarrollar fases posteriores.

La siguiente fase a analizar es la de Plan del Proyecto. Dentro de los Procesos de Administración del Proyecto, en esta fase, salvo en la tarea de Entendimiento del Negocio, se llevan a cabo todas las tareas tanto del SubProceso de Planificación /Entendimiento del Negocio como de Realización. Se lleva a cabo esta asunción ya que todas esas actividades influyen en el Plan de Proyecto desarrollado, tales como la generación de una base de conocimiento del negocio, la definición de los objetivos del negocio, definición de criterios del éxito, desarrollo de Plan de Entrega, Plan de Desarrollo, Plan de Adquisiciones, armado de Calendario de actividades, estimación de costos, generación de inventario de recursos, riesgos y contingencias, descripción de la terminología empleada, lista de técnicas, reporte de seguimiento y monitoreo.

Dentro de los Procesos de Administración de Proyectos, las únicas tareas que se llevan a cabo en el SubProceso Planificación/Entendimiento del Negocio son las de Entendimiento de Negocio y Definir el proceso específico basado en la descripción del proyecto y el proceso de desarrollo y mantenimiento. Se ha definido esto ya que tareas tales como evaluar la descripción del proyecto, evaluar el proceso de desarrollo y mantenimiento, describir el problema en términos generales, clasificar prerequisites del proyecto, especificar los beneficios esperados y definir el proceso específico a utilizar, son incluidas dentro de esta fase.

Por otro lado, las tareas correspondiente al SubProceso de Realización no se realiza ninguna salvo las de recibir y analizar las solicitudes de cambio del cliente y realizar reuniones con el equipo de trabajo y cliente para reportar avances y tomar acuerdos, ya que recibir solicitudes de cambio del cliente y las Realizaciones de reuniones con el cliente pueden afectar y es importante tener en cuenta los Objetivos establecidos para el Proyecto de Explotación de Información.

Dentro del SubProceso de Evaluación y Control, la actividad que se lleva a cabo es la de generar el reporte de seguimiento del cliente, ya que se analiza que cosas pueden cambiar a futuro. Por otro lado, dentro del SubProceso de Cierre/Entrega, no se lleva a cabo ninguna actividad ya que se entiende que corresponden a fases posteriores.

Siguiendo con el análisis de esta fase, dentro de los Procesos de Desarrollo de Proyecto, el único SubProceso que se relaciona con esta fase es la de Entendimiento del Negocio.

Es así ya que determinar las metas del proyecto de Explotación de Información especifica los criterios de valoración de los modelos, determina los benchmarks para la evaluación de criterios y especifica criterios a valoración de criterios subjetivos, entre otras tareas.

Las otras actividades que corresponden a los demás SubProcesos de Desarrollo de Proyectos, se entiende que no corresponden a esta fase ya que son actividades que se deben desarrollar fases posteriores.

La siguiente fase a analizar es la de Plan del Proyecto.

ACTIVIDADES DE LOS PROCESOS	O.N.	E.Sit.	O.P.	P.P.	Val.O.	R.D.	D.D.	E.D.	Ver.CO	S.D.	L.D.	C.D.	I.D.	F.D.	Val.D.	S.T.M.	G.P.D.	C.M.	E.M.	E.R.	R.P.	D.P.P.	D.P.	P.S.M.	I.P.	R.Py.	
Proceso de Administración de Proyectos																											
Planificación /Entendimiento del negocio																											
- Entendimiento del negocio	■	■	■																								
- Definir el proceso específico basado en la descripción del proyecto y el proceso de desarrollo y mantenimiento			■																								
- Definir el protocolo de entrega con el cliente				■																							
- Definir ciclos y actividades con base en la descripción del proyecto y en el proceso específico				■	■																						
- Determinar tiempo estimado para cada actividad				■	■																						
- Elaborar plan de adquisiciones y capacitación				■	■																						
- Establecer el equipo de trabajo				■	■																						
- Establecer el calendario de actividades				■	■																						
- Calcular el costo estimado del proyecto				■	■																						
- Evaluación de la situación		■		■	■	■	■	■	■	■	■	■	■	■	■												
- Producir un Plan de Proyecto				■	■										■												
- Producir un Plan de Desarrollo				■	■						■	■	■	■	■												
- Formalizar el inicio de un nuevo ciclo del proyecto				■	■										■												
Realización																											
- Acordar las tareas con el equipo de trabajo				■	■											■	■	■									
- Acordar la distribución de información				■	■												■	■									
- Revisar con el responsable la descripción del producto, el equipo de trabajo y el calendario				■	■													■	■								
- Revisar cumplimiento del plan de adquisiciones y capacitación				■	■														■								
- Administrar subcontratos				■	■														■								
- Recolectar reportes de actividades y mediciones y sugerencias de mejora y productos de trabajo				■	■			■							■				■								
- Registrar costo real del proyecto				■	■																						
- Revisar el registro de rastreo basado en los productos de trabajo recolectados				■	■														■								
- Revisar los productos terminados durante el proyecto				■	■														■								
- Recibir y analizar las solicitudes de cambio del cliente			■	■	■																			■	■	■	
- Realizar reuniones con el equipo de trabajo y cliente para reportar avances y tomar acuerdos			■	■	■																			■	■	■	
Evaluación y Control																											
- Evaluar el cumplimiento del plan de proyecto y plan de desarrollo																				■	■	■	■	■	■	■	■
- Analizar y controlar los riesgos																								■	■		■
- Generar el reporte de seguimiento del proyecto	■		■	■	■																			■	■	■	■
Cierre / Entrega																											
- Formalizar la terminación del proyecto o ciclo																								■	■	■	■
- Llevar a cabo el cierre del contrato con subcontratistas																								■	■	■	■
- Generar el reporte de mediciones y sugerencias de mejora																								■	■	■	■
- Planear la entrega																								■	■	■	■

Fig. 2. Mapa de Actividades Propuesto para Administración de Proyectos

ACTIVIDADES DE LOS PROCESOS	O.N.	E.Sit.	O.P.	P.P.	Val.O.	R.D.	D.D.	E.D.	Ver.CD	S.D.	L.D.	C.D.	I.D.	F.D.	Val.D.	S.T.M.	G.P.D.	C.M.	E.M.	E.R.	R.P.	D.P.P.	D.P.	P.S.M.	I.P.	R.Py.	
Proceso de Desarrollo de Proyectos																											
Entendimiento del negocio																											
- Determinar las metas del Data Mining	■	■	■	■	■																						
Entendimiento de los datos																											
- Reunir los datos iniciales						■																					
- Describir los datos							■																				
- Explorar los datos								■																			
- Verificar la calidad de los datos									■																		
Preparación de los datos																											
- Tareas preparatorias										■																	
- Seleccionar los datos										■																	
- Limpiar los datos											■																
- Construir los datos												■															
- Integrar los datos													■														
- Formatear los datos														■	■												
Modelado																											
- Seleccionar la técnica de modelado																■											
- Generar el diseño de test																	■										
- Construir el modelo																		■									
- Evaluar el modelo																			■								
Evaluación																											
- Evaluar resultados																					■						
- Revisar el proceso																						■					
- Determinar próximos pasos																							■				
Entrega																											
- Producir un reporte final																								■	■	■	■

Fig. 3. Mapa de Actividades Propuesto para Desarrollo de Proyectos

Dentro de los Procesos de Administración del Proyecto, en esta fase, salvo en la tarea de Entendimiento del Negocio, se llevan a cabo todas las tareas tanto del SubProceso de Planificación / Entendimiento del Negocio como de Realización. Se lleva a cabo esta asunción ya que todas esas actividades influyen en el Plan de Proyecto desarrollado, tales como la generación de una base de conocimiento del negocio, la definición de los objetivos del negocio, definición de criterios del éxito, desarrollo de Plan de Entrega, Plan de Desarrollo, Plan de Adquisiciones, armado de Calendario de actividades, estimación de costos, generación de inventario de recursos, riesgos y contingencias, descripción de la terminología empleada, lista de técnicas, reporte de seguimiento y monitoreo.

Dentro de los Procesos de Administración de Proyectos, las únicas tareas que se llevan a cabo en el SubProceso Planificación/Entendimiento del Negocio son las de Entendimiento de Negocio y Definir el proceso específico basado en la descripción del proyecto y el proceso de desarrollo y mantenimiento. Se ha definido esto ya que tareas tales como evaluar la descripción del proyecto, evaluar el proceso de desarrollo y mantenimiento, describir el problema en términos generales, clasificar prerrequisitos del proyecto, especificar los beneficios esperados y definir el proceso específico a utilizar, son incluidas dentro de esta fase.

Por otro lado, las tareas correspondiente al SubProceso de Realización no se realiza ninguna salvo las de recibir y analizar las solicitudes de cambio del cliente y realizar reuniones con el equipo de trabajo y cliente para reportar avances y tomar acuerdos, ya que recibir solicitudes de cambio del cliente y las Realizaciones de reuniones con el cliente pueden afectar y es importante tener en cuenta los Objetivos establecidos para el Proyecto de Explotación de Información.

Dentro del SubProceso de Evaluación y Control, la actividad que se lleva a cabo es la de generar el reporte de

seguimiento del cliente, ya que se analiza que cosas pueden cambiar a futuro. Por otro lado, dentro del SubProceso de Cierre/Entrega, no se lleva a cabo ninguna actividad ya que se entiende que corresponden a fases posteriores.

Siguiendo con el análisis de esta fase, dentro de los Procesos de Desarrollo de Proyecto, el único SubProceso que se relaciona con esta fase es la de Entendimiento del Negocio. Es así ya que determinar las metas del proyecto de Explotación de Información específica los criterios de valoración de los modelos, determina los benchmarks para la evaluación de criterios y especifica criterios a valoración de criterios subjetivos, entre otras tareas.

Las otras actividades que corresponden a los demás SubProcesos de Desarrollo de Proyectos, se entiende que no corresponden a esta fase ya que son actividades que se deben desarrollar fases posteriores.

La siguiente fase a analizar es la de Plan del Proyecto. Dentro de los Procesos de Administración del Proyecto, en esta fase, salvo en la tarea de Entendimiento del Negocio, se llevan a cabo todas las tareas tanto del SubProceso de Planificación /Entendimiento del Negocio como de Realización. Se lleva a cabo esta asunción ya que todas esas actividades influyen en el Plan de Proyecto desarrollado, tales como la generación de una base de conocimiento del negocio, la definición de los objetivos del negocio, definición de criterios del éxito, desarrollo de Plan de Entrega, Plan de Desarrollo, Plan de Adquisiciones, armado de Calendario de actividades, estimación de costos, generación de inventario de recursos, riesgos y contingencias, descripción de la terminología empleada, lista de técnicas, reporte de seguimiento y monitoreo.

Por otro lado, dentro de la Evaluación y Control, la actividad que se lleva a cabo es la de generar el reporte de seguimiento del proyecto, ya que afecta al Plan de Proyecto. Las actividades de cierre/entrega no se realizan en esta fase ya que no corresponde a esta altura.

Dentro de los Procesos de Desarrollo del Proyecto, en el único SubProceso donde se realiza una actividad en esta fase es en el de Entendimiento del Negocio, debido a que las metas del proyecto de Explotación de Información y los Criterios del Éxito del Proyecto de E.I. quedan registrados en la conformación del Plan de Proyecto.

Las demás actividades correspondientes a los demás SubProcesos de Desarrollo del Proyecto no se llevan a cabo en esta fase ya que son actividades posteriores a las configuraciones del Plan de Proyecto.

La última fase dentro de la segunda vuelta que se lleva a cabo en el Modelo de Ciclo de Vida en Espiral propuesto es la de Validación de Objetivos.

Dentro de los Procesos de Administración del Proyecto, en esta fase al igual que la fase anterior, salvo en la tarea de Entendimiento del Negocio y la de definir el proceso específico basado en la descripción del proyecto y el proceso de desarrollo y mantenimiento, se desarrollan todas las tareas tanto del SubProceso de Planificación /Entendimiento del Negocio como de Realización. Se asume esto puesto que todas esas actividades influyen en el Plan de Proyecto desarrollado, tales como la generación de una base de conocimiento del negocio, la definición de los objetivos del negocio, definición de criterios del éxito, desarrollo de Plan de Entrega, Plan de Desarrollo, Plan de Adquisiciones, armado de Calendario de actividades, estimación de costos, generación de inventario de recursos, riesgos y contingencias, descripción de la terminología empleada, lista de técnicas, reporte de seguimiento y monitoreo. Dentro de la Evaluación y Control se lleva a cabo la tarea de generar el reporte de seguimiento del cliente, ya que se evalúan si los objetivos pueden cambiar a lo largo del Proyecto.

En los Procesos de Desarrollo del Proyecto la única actividad que se realiza en esta fase es la de determinar las metas del Data Mining, correspondiente al SubProceso de Entendimiento del Negocio. Las demás actividades de los restantes 5 SubProcesos no corresponden a esta fase ya que como se ha mencionado en otras fases, esas tareas son ligadas a etapas posteriores.

La primera fase de la tercera vuelta del espiral es la de Recolección de Datos. Teniendo en cuenta los Procesos de Administración de Proyectos, dentro de la Planificación/Entendimiento del Negocio, la única tarea que se lleva a cabo es la de Evaluación de la Situación. Esto es así ya que tareas como identificar las fuentes de datos, identificar los distintos tipos de fuentes de datos, identificar fuentes de conocimiento, chequear la disponibilidad de técnicas y herramientas, identificar los administradores de sistemas, bases de datos, entre otras, se vinculan estrechamente con esta Fase.

Por otro lado, las actividades anteriores y posteriores de este SubProceso no se realizan en esta fase del Modelo ya que se entiende que deben hacerse en otras fases debido al tiempo de desarrollo del Proyecto. Las actividades pertenecientes a los SubProcesos de Realización, Evaluación y Control, y Cierre/Entrega, no se vinculan con esta fase puesto que son actividades que se llevan a cabo en fases posteriores y no se relacionan con la Recolección de Datos propiamente dicha.

Entrando en los Procesos de Desarrollo de Proyectos, se define que las actividades pertenecientes al SubProceso de Entendimiento del Negocio no se llevan a cabo en esta fase ya que son tareas propias de las 2 primeras vueltas del espiral.

Sin embargo, la tarea de Reunir los Datos Iniciales del SubProceso de Entendimiento de los Datos sí se realiza ya que se identifica cual es la información necesaria, se chequea si la información está disponible, se especifica el criterio de selección, se seleccionan datos dentro de tablas o archivos, se evalúa como se debe adquirir datos faltantes, se analiza si los datos contienen entradas de texto libre evaluar, si es necesario codificarla para el modelado o si es necesario agrupar entradas específicas, y se describe como se debe extraer los datos entre otras tareas.

Por otro lado, las demás actividades de este SubProceso y de los SubProcesos de Preparación de los Datos, Modelado, Evaluación y Entrega no se llevan a cabo ya que son tareas de fases siguientes.

La próxima fase a analizar es la de Descripción de Datos. Empezando por los Procesos de Administración de Proyectos, la única tarea que se lleva a cabo es la de Evaluación de la Situación, dentro del SubProceso de Planificación/Entendimiento del Negocio. Esto se define así ya que entre otras tareas se identifican las fuentes de datos, se describe el conocimiento (background), se listan las asunciones de la calidad de los datos, se chequean las restricciones generales en torno a los datos, se chequea la accesibilidad que los datos poseen, etc. Las actividades siguientes de este SubProceso no se asignan a esta fase ya que se entiende que tienen que ver más con el Negocio que con los Datos y su análisis en sí. Además, los demás SubProcesos no se relacionan con esta Fase ya que cubren más a otras que se llevan a cabo más adelante.

Considerando los Procesos de Desarrollo de Proyecto, las actividades del SubProceso de Entendimiento de Negocio no se relaciona con esta fase dado que son actividades de fases anteriores. En cambio, la actividad de describir los datos del SubProceso de Entendimiento de los datos sí, ya que se identifican los datos y métodos de captura, se acceden a las fuentes de datos, se usan análisis estadísticos, se chequea el volumen de datos, se chequea la disponibilidad y accesibilidad, se chequean los tipos de atributos, se analiza la correlación de atributos, entre otras tareas. Por otro lado, las demás actividades de este SubProceso y las actividades de Preparación de los Datos, Modelado, Evaluación y Entrega se realizan en fases posteriores, cuando el Proyecto esté más avanzado en cuanto a su desarrollo.

Siguiendo con las fases correspondientes a la tercera vuelta, se analiza la Exploración de Datos. En primer lugar, dentro de los Procesos de Administración de Proyectos, se lleva a cabo en el SubProceso de Planificación /Entendimiento del Negocio la tarea de Evaluación de la Situación, tal como las dos fases anteriores de esta vuelta del Modelo de Ciclo de Vida en Espiral, ya que se chequean los derechos de acceso a las fuentes de datos, se chequean las técnicas de acceso a los datos, si el conocimiento relevante es accesible, etc. Por otro lado, las actividades de los SubProcesos siguientes no se realizan en esta fase ya que dan soporte y se relacionan con fases más avanzadas.

Teniendo en cuenta los Procesos de Desarrollo de Proyectos, el SubProceso que lleva a cabo una actividad en esta fase es la de Entendimiento del Negocio. Dentro del mismo, se realiza la exploración de datos, o sea analizar propiedades de atributos interesantes en detalle, identificar características de sub-poblaciones, considerar y evaluar información y resultados en los reportes de descripción de

datos, formular hipótesis e identificar acciones, transformar hipótesis en metas de Explotación de Información si es posible, clarificar las metas de Explotación de Información o hacerlas más precisas y realizar análisis básicos para verificar las hipótesis. Las actividades de los SubProcesos siguientes se realizan en fases más adelantadas por su avance en el desarrollo del Proyecto de E.I.

La última fase a analizar en esta vuelta del Modelo de Ciclo de Vida en Espiral es la de Verificación Calidad de Datos. En primer lugar, teniendo en cuenta los Procesos de Administración de Proyectos, se lleva a cabo la tarea de Evaluación de la Situación del SubProceso de Planificación/Entendimiento del Negocio, ya que se listan las presunciones de calidad de los datos. Luego, dentro del SubProceso de Realización, se realiza la tarea de recolectar reportes de actividades y mediciones y sugerencias de mejora y productos de trabajo. Se informa que esta actividad se realiza en esta fase dado que se generan un informe con los datos recolectados, se realizan los ajustes necesarios, se promueven las sugerencias de mejora en caso de ser adecuadas, se recolectan productos de trabajo, entre otras. Estas actividades influyen en la verificación de la calidad de datos. Las actividades de los demás SubProcesos correspondientes a la Administración de Proyecto corresponden a fases más avanzadas.

Por otro lado y finalizando con el análisis de la tercera vuelta del Modelo de Ciclo de Vida en Espiral, la actividad que se desarrolla en esta fase dentro de los Procesos de Desarrollo de Proyectos, es la de verificar la calidad de datos propiamente dicho, correspondiente al SubProceso de Entendimiento de los Datos. Dentro de esta actividad se incluyen tareas tales como identificar valores especiales y catalogar su significado, chequear claves, chequear cubrimientos, identificar atributos faltantes, chequear desviaciones, dar significado a la falta de datos, planear como lidiar con los ruidos, detectar tipos de ruidos, entre otros. Este tipo de actividad es propia de esta fase y de ninguna otra. Las demás actividades de los SubProcesos continuos corresponden a fases siguientes del Modelo de Ciclo de Vida propuesto.

La primera fase en analizar de la cuarta vuelta del espiral es la de Selección de Datos. Teniendo en cuenta los Procesos de Administración de Proyectos, en el único SubProceso donde se lleva a cabo una actividad es en la de Planificación/Entendimiento del Negocio. La actividad que se realiza en esta fase es la de Evaluación de la Situación, dado que registran las fuentes de datos, fuentes de conocimientos, se describen los datos, se identifican tipos de fuentes de conocimientos, entre otras. Las actividades de los demás SubProcesos no han de llevarse a cabo en dicha fase dado que no se relacionan directamente con la selección de datos propiamente dicha y además, dan soporte a fases posteriores en el desarrollo del proyecto.

Teniendo en cuenta los Procesos de Desarrollo de Proyectos, esta fase realiza las siguientes actividades pertenecientes al SubProceso de Preparación de los Datos: Tareas Preparatorias y Selección de Datos. Con relación a la primer actividad mencionada, se ha de llevar a cabo la reunión de datos producidos por la fase de preparación de datos y, con respecto a la selección de datos propiamente dicha, se recolectan datos adicionales apropiados, se realizan test de significancia, se reconsideran los criterios de selección de datos, se seleccionan diferentes subconjuntos de datos, se

documentan razones de exclusión e inclusión, etc. Se ha definido estas actividades para esta fase ya que se relacionan con la selección de datos y no las demás ya que son propias de fases siguientes, como limpieza, construcción, integración y formateo.

En segundo lugar, analizando las actividades vinculadas con la fase de Limpieza de Datos, se define que tal como con la Selección de Datos, el único SubProceso de Administración de Proyecto que tiene actividades en esta fase es la de Planeamiento /Entendimiento del Negocio, ya que se realiza la Evaluación de la Situación y la Producción de un Plan de Desarrollo. La evaluación de la situación da soporte a la generación de limpieza de los datos obtenidos con las listas de asunciones y de riesgos producidas. Además, en la Producción de un Plan de Desarrollo quedan registrados los pasos críticos y técnicas de limpieza que se llevan a cabo.

Por otro lado, teniendo en cuenta los Procesos de Desarrollo de Proyectos, se realiza la actividad de limpiar los datos pertenecientes al SubProceso de Preparación de los Datos, dado que se reconsidera como lidiar con los tipos de ruidos, se corrigen, remueven e ignoran los ruidos, se decide como lidiar con valores especiales y se reconsidera la selección de datos con base a experiencia de limpieza. Como se ha mencionado en fases anteriores, las actividades tanto de SubProceso de Modelado y Evaluación están relacionadas con fases más avanzadas del espiral.

Siguiendo con el análisis de la cuarta vuelta del Modelo de Ciclo de Vida en Espiral propuesto en este capítulo, la siguiente fase en ser analizada es la de Construcción de Datos. Con respecto a los Procesos de Administración de Proyectos, se lleva a cabo dentro del SubProceso de Preparación /Entendimiento del Negocio la actividad de Evaluación de la Situación y la de Producción de un Plan de Desarrollo. Ambas actividades dan soporte a nivel administrativo a la construcción de datos y sus técnicas de transformaciones.

Teniendo en cuenta los Procesos de Desarrollo de Proyectos, se lleva a cabo la actividad de construir los datos, dentro del SubProceso de Preparación de los Datos, chequeando los mecanismos de construcción disponibles, decidiendo si es mejor realizar la construcción dentro o fuera de la herramienta, reconsiderando los criterios de selección de datos, decidiendo si hay que normalizar atributos, decidiendo como pueden ser construidos o imputados atributos faltantes, realizando pasos de transformación o chequeando técnicas disponibles de generación de registros si son necesarios.

La siguiente fase de esta vuelta del espiral es la de Integración de Datos. Entrando en los Procesos de Administración de Proyecto, las actividades que se realizan tal como en las fases anteriores son las de Evaluación de la Situación y la de Producción de un Plan de Desarrollo, del SubProceso de Preparación/Entendimiento del Negocio. Tal como se menciono en la fase anterior, estas actividades dejan sentadas a nivel administrativo las técnicas de integración de los datos. Las actividades posteriores a estos procesos corresponden a fases avanzadas del proyecto desarrollado.

Por otro lado, dentro de los Procesos de Desarrollo de Proyectos, específicamente en el SubProceso de Preparación de los Datos, se realiza la actividad de integrar los datos propiamente dicha. Esto es ya que se chequean si las facilidades de integración son capaces de integrar las fuentes de entrada como es requerido, se integran las fuentes y almacenan los resultados, y se reconsideran los criterios de

selección de los datos basándose en las experiencias de integración.

La quinta fase en ser analizada de esta vuelta del Modelo en Espiral es la de Formateo de Datos. Teniendo en cuenta los Procesos de Administración de Proyectos, se realiza la actividad de Evaluar la Situación y la de Producir un Plan de Desarrollo del SubProceso de Planeamiento / Entendimiento del Negocio.

Dentro de los Procesos de Desarrollo de Proyecto, se realiza la actividad de formatear los datos que se realiza en el SubProceso de Preparación de los Datos. Las tareas que se llevan a cabo son las de reordenar los atributos si es requerido por las herramientas utilizadas, reordenar los registros si es necesario, reformatear valores para satisfacer los requerimientos de las herramientas utilizadas, entre otras.

La última fase de esta vuelta del espiral es la de Validación de los Datos. Con respecto a los Procesos de Administración de Proyectos se realizan las actividades de evaluar la situación, producir un plan de proyecto, producir un plan de desarrollo y formalizar el inicio de un nuevo ciclo del proyecto. A nivel gerencial y administrativo se da soporte por medio de estas actividades a la validación de los datos obtenidos y seleccionados, quedando registrados que los datos son validos y cumplen con los requerimientos del Proyecto. Por medio de la validación exitosa de los datos, se da inicio al ciclo correspondiente del proyecto y se les comunica a los interesados.

Por otro lado, dentro de los Procesos de Desarrollo de Procesos, la actividad que se vincula con esta fase es la de formatear los datos, dado que es la última actividad del SubProceso de Preparación de los Datos, y por medio de esta ya quedan los datos listos para ser usados en la construcción del modelo y finalmente se validan para corroborar si realmente cumplen con los requerimiento establecidos en el proyecto.

En la quinta vuelta del Modelo de Ciclo de Vida en Espiral propuesto en este capítulo se encuentran las siguientes fases: Selección Técnica de Modelado, Generación Prueba de Diseño, Construcción Modelo y Evaluación del Modelo.

Comenzando por la fase de Selección Técnica de Modelado, se puede observar que dentro de los Procesos de Administración de Proyectos, específicamente en el SubProceso de Realización, posee la actividad de acordar las tareas con el equipo del trabajo, puesto que se revisan las tareas a realizar, se evalúan los roles y se distribuyen las tareas. Por medio de estas actividades se brinda soporte a las actividades de desarrollo propias como la selección de la técnica a utilizar para desarrollar el modelo.

Teniendo en cuenta los Procesos de Desarrollo de Proyectos y dentro del SubProceso Modelado, se realiza la actividad de seleccionar la técnica de modelado propiamente dicha, mediante la cual se decide sobre la técnica apropiada para ejercitar teniendo en mente la herramienta seleccionada, se define cualquier asunción hecha por la técnica acerca de los datos, se comparan las asunciones con las hechas en el reporte de descripción de dato, y se asegura que las asunciones se mantienen y regresar a la fase de preparación de datos si es necesario.

En segundo lugar, siguiendo con la fase de Generación Prueba de Diseño y dentro de los Procesos de Administración de Proyectos se realizan las actividades de acordar las tareas con el equipo del trabajo y acordar la distribución de

información, pertenecientes al SubProceso Realización. Por medio de estas tareas se distribuyen las tareas, se identifica la información requerida para cada tarea y cada plan de prueba del modelo a construir, y se distribuye la información a todo el equipo responsable del desarrollo del Proyecto. Las actividades siguientes de este SubProceso y de los SubProcesos siguientes corresponden a fases más avanzadas donde el Modelo ya fue desarrollado y lo que queda son tareas tales como evaluación y entrega al cliente.

Por otro lado, dentro de los Procesos de Desarrollo de Proyectos, se realiza la actividad de generar el diseño del test, en el SubProceso de Modelado, donde se chequean los diseños de test existentes para cada meta de E.I. de forma separada, se deciden los pasos necesarios y se preparan los datos requeridos para testear. Las siguientes actividades de estos Procesos corresponden a fases siguientes donde el Modelo ya está construido.

La tercera fase de esta vuelta del espiral es la de Construcción del Modelo. En los Procesos de Administración de Proyectos se llevan a cabo las actividades de acordar las tareas con el equipo de trabajo, acordar la distribución de la información y revisar con el responsable la Descripción del producto, el equipo de trabajo y el calendario, pertenecientes al SubProceso de Realización. Por medio de estas tareas, en esta fase donde se ejecuta la técnica y se construye el modelo, a nivel administrativo se distribuye la información a cada responsable, se reúne con el equipo responsable del desarrollo del modelo, se revisa la Descripción del producto, se revisa el cumplimiento del calendario y se realizan correcciones si son necesarios. Como se ha mencionado anteriormente en la explicación del Modelo de Ciclo de Vida propuesto, en esta vuelta del espiral se cicla, donde los distintos prototipos desarrollados y Evaluaciones hechas sobre el modelo juegan un papel importante.

Adicionalmente, en los Procesos de Desarrollo de Proyecto, se realiza la actividad de construir el modelo propiamente dicho, actividad situada en el SubProceso de Modelado. Mediante esta actividad se setean los parámetros iniciales, se documentan razones de selección, se ejecuta la técnica seleccionada, se procesan los resultados, se describen las características del modelo, se da una explicación detallada del mismo, y se brinda una interpretación del mismo. En esta fase ya se encuentra desarrollado y validado el modelo, por lo que las actividades siguientes tendrán que ver con la Evaluación y Entrega.

La última fase de esta vuelta del espiral es la de Evaluación del Modelo. Dentro de los Procesos de Administración de Proyectos, específicamente en el SubProceso de Realización, se realizan las siguientes actividades: revisar con el responsable la descripción del producto, el equipo de trabajo y el calendario, revisar cumplimiento del plan de adquisiciones y capacitación, administrar subcontratos, recolectar reportes de actividades y mediciones y sugerencias de mejora y productos de trabajo, registrar costo real del proyecto, revisar el registro de rastreo basado en los productos de trabajo recolectados y revisar los productos terminados durante el proyecto. Por medio de estas actividades y una vez desarrollado el Modelo, se realizan correcciones si son necesarias, se controla cumplimiento, se realizan ajustes necesarios del plan de adquisición, se revisan subcontratos, se recolectan reportes de actividades y mediciones, se recolectan los productos del trabajo, se revisa

el costo real del proyecto, se revisa el registro de rastreo de los productos, y finalmente se revisan los productos terminados durante el Proyecto y se registran los mismos.

Además, entrando en los Procesos de Desarrollo de Proyectos, se realiza la actividad de evaluar el modelo, perteneciente al SubProceso de Modelado, por la cual se evalúan los resultados con respecto a los criterios de evaluación, se testean los resultados de acuerdo a la estrategia de testeo, se comparan la evaluación de los resultados y su interpretación, se seleccionan los mejores modelos, se interpretan los resultados en términos de negocios, se chequea el impacto para las metas de E.I., se chequea la confiabilidad de los resultados, se valoran los resultados, y se ajustan los parámetros para obtener un mejor modelo en caso de ser necesario o se cree conveniente.

En la última vuelta del espiral, en la fase de Evaluación de Resultados y dentro de los Procesos de Administración de Proyectos, se realiza la actividad de evaluar el cumplimiento del Plan de Proyecto y Plan de Desarrollo, correspondiente al SubProceso de Evaluación y Control. Por medio de esta actividad se evalúa el cumplimiento del plan de proyecto, se ajusta el mismo, se generan informes, se evalúa el plan de plan de desarrollo y se ajusta el mismo. Estas actividades se desarrollan una vez construido el modelo, por lo que las actividades anteriores no pertenecen a esta fase y fases siguientes tienen más que ver con fases avanzadas de la Evaluación y de la Entrega/Cierre.

Por otro lado, en los Procesos de Desarrollo de Proyectos, se lleva a cabo la actividad de evaluar resultados, incluida en el SubProceso de Evaluación. En esta actividad se entienden los resultados del Proyecto de E.I., se interpretan los resultados, se evalúan y valoran los resultados, se compara la evaluación, se chequea si existen nuevos objetivos del negocio y finalmente se obtiene la aprobación de los modelos.

En segundo lugar, la fase de Revisión Proyecto posee incluida dentro de los Procesos de Administración de Proyecto, la actividad de evaluar el cumplimiento del Plan de Proyecto y Plan de Desarrollo, correspondiente al SubProceso de Evaluación y Control. En esta actividad se revisa y evalúa el Proyecto y los Planes de los mismos, tanto el de Desarrollo como el de Proyecto en sí.

Teniendo en cuenta los Procesos de Desarrollo de Proyectos, se realiza la actividad de revisar el proyecto del SubProceso de Evaluación, en la cual se da un examen del proceso de E.I. utilizado, se analiza el proceso de E.I., se evalúa como puede ser mejorado, se identifican fallas, se identifican pasos faltantes, se identifican alternativas y se revisan resultados del Proyecto de E.I. con respecto a los criterios del éxito del negocio.

La siguiente fase de esta vuelta del espiral es la de Determinación de los Próximos Pasos, en la cual por medio de los Procesos de Administración de Proyectos y más específicamente en el SubProceso de Evaluación y Control, lleva a cabo la tarea de evaluar el cumplimiento del plan de proyecto y plan de desarrollo, donde genera informes de ajustes del plan de desarrollo y plan de proyecto, y genera un Reporte de Seguimiento / Plan de Monitoreo y Mantenimiento, previo paso al Cierre y Entrega del Proyecto.

Por otro lado, en los Procesos de Desarrollo de Proyectos (SubProceso de Evaluación), se incluye la actividad de determinar próximos pasos, donde analiza el potencial de despliegue de cada resultado, se estima el potencial de mejora

del proceso actual, se chequea recursos sobrantes para determinar si permiten realizar iteraciones del proceso adicionales, se recomiendan alternativas de continuidad, se refina el plan de proceso, se ranquean las posibles acciones y se selecciona y documenta las razones de elección.

La cuarta fase correspondiente a esta última vuelta del Modelo de Ciclo de Vida en Espiral es la de Desarrollo Plan. Teniendo en cuenta los Procesos de Administración de Proyecto, se define que posee las actividades de evaluar el cumplimiento del plan de proyecto y plan de desarrollo, analizar y controlar los riesgos, generar el reporte de seguimiento del proyecto (SubProceso de Evaluación y Control) y formalizar la terminación del proyecto o ciclo, llevar a cabo el cierre del contrato con subcontratistas, generar el reporte de mediciones y sugerencias de mejora, y planear la entrega (SubProceso Cierre/Entrega). Estas actividades generar un Plan de Seguimiento y Plan de Monitoreo y Mantenimiento, documentos de aceptación, informes de subcontratistas, reportes de mediciones y sugerencias de mejoras.

En los Procesos de Desarrollo de Proyectos, se lleva a cabo la actividad de producir un reporte final, perteneciente al SubProceso de Entrega. En esta actividad se identifican que reportes son necesarios para finalizar el Proyecto en cuestión.

En quinto lugar se encuentra la fase de Plan Supervisión y Mantenimiento, donde por medio de los Procesos de Administración de Proyectos, lleva a cabo las tareas de recibir y analizar las solicitudes de cambio del cliente, realizar reuniones con el equipo de trabajo y cliente para reportar avances y tomar acuerdos (SubProceso de Realización), evaluar el cumplimiento del plan de proyecto y plan de desarrollo, analizar y controlar los riesgos, generar el reporte de seguimiento del proyecto (SubProceso de Evaluación y Control) y formalizar la terminación del proyecto o ciclo, llevar a cabo el cierre del contrato con subcontratistas, generar el reporte de mediciones y sugerencias de mejora, y planear la entrega (SubProceso Cierre/Entrega). Por medio de estas tareas se obtiene a nivel administrativo el Plan propiamente dicho que resume la estrategia de supervisión y mantenimiento incluyendo los pasos necesarios y como realizarlos.

Teniendo en cuenta los Procesos de Desarrollo de Proyectos, se lleva a cabo la tarea de producir un reporte final (SubProceso Entrega), donde se esboza la estructura del informe, se seleccionan los descubrimientos a ser incluidos en el reporte, se informa la estrategia de supervisión y mantenimiento y se escribe finalmente el reporte.

La siguiente fase del Modelo de Ciclo de Vida es la de Informe Producto, donde por medio de los Procesos de Administración de Proyectos lleva a cabo las siguientes actividades: recibir y analizar las solicitudes de cambio del cliente, realizar reuniones con el equipo de trabajo y cliente para reportar avances y tomar acuerdos (SubProceso de Realización), evaluar el cumplimiento del plan de proyecto y plan de desarrollo, analizar y controlar los riesgos, generar el reporte de seguimiento del proyecto (SubProceso de Evaluación y Control) y formalizar la terminación del proyecto o ciclo, llevar a cabo el cierre del contrato con subcontratistas, generar el reporte de mediciones y sugerencias de mejora, y planear la entrega (SubProceso Cierre/Entrega). Por medio de estas actividades se generan los Reportes de Seguimiento y Mantenimiento, se realiza Informe de SubContratistas, Reportes de Mediciones, se documenta el

proceso de E.I. específico, se resumen los detalles para futuros Proyectos similares, se realiza un informe de los contratos, se comunica la finalización a todos los interesados, entre otras tareas.

Teniendo en cuenta los Procesos de Desarrollo de Proyectos, se lleva a cabo la actividad de producir un reporte final perteneciente al SubProceso de Entrega. En esta actividad se escribe el reporte e informe final y se lleva a cabo la presentación final del Proyecto.

La última fase del Modelo de Ciclo de Vida en Espiral propuesto en este capítulo y que se analiza con respecto al Mapa de Actividades construido es la de Revisión Proyecto. Teniendo en cuenta los Procesos de Administración de Proyectos, se llevan a cabo las siguientes tareas: recibir y analizar las solicitudes de cambio del cliente, realizar reuniones con el equipo de trabajo y cliente para reportar avances y tomar acuerdos (SubProceso de Realización), evaluar el cumplimiento del plan de proyecto y plan de desarrollo, analizar y controlar los riesgos, generar el reporte de seguimiento del proyecto (SubProceso de Evaluación y Control) y formalizar la terminación del proyecto o ciclo, llevar a cabo el cierre del contrato con subcontratistas, generar el reporte de mediciones y sugerencias de mejora, y planear la entrega (SubProceso Cierre/Entrega). Por medio de estas actividades se generan los reportes e informes finales y se lleva a cabo el Plan de Entrega, donde se planifica la entrega de los productos generados.

Con respecto a los Procesos de Desarrollo de Proyectos, se realiza la actividad de producir un reporte final perteneciente al SubProceso de Entrega, donde se lleva a cabo la Presentación Final del Proyecto y donde se documenta todo el desarrollo del Proyecto para futuras experiencias y trabajos similares a realizar.

V. CONCLUSIÓN

En esta sección se detallan las conclusiones obtenidas de la presente investigación, mencionando en un principio los aportes brindados (sección A) para finalmente destacar las futuras líneas de investigación a partir del problema mencionado dentro de este artículo de investigación (sección B).

A. Aportes del artículo de Investigación

Se ha comprobado que si bien la metodología existente para Proyectos de Explotación de Información, CRISP-DM [6], es la más completa para desarrollo de este tipo de Proyectos, hay fases que no cubre tal como los Procesos de Administración. Por ello, se ha detectado necesaria la utilización del Modelo de Procesos desarrollado por Vanrell [25], el cual le brinda un valor agregado a la metodología CRISP-DM al incluirle los Procesos de Administración. Estos procesos de Administración tienen la función de recolectar información que le sea útil a los procesos de Desarrollo para aumentar la calidad de los mismos permitiendo realizar ajustes y mantener un estándar en la realización de distintos proyectos llevando a cabo la conversión de procesos caóticos y aislados en procesos repetitivos y controlados.

Además de esto, se detecta que la metodología CRISP-DM no posee un Análisis enfocado a los Riesgos. Tal como ocurre en los Desarrollos Convencionales de Software, en los Proyectos de Explotación de Información también pueden ocurrir Riesgos y es fundamental detectarlos y tratarlos

tempranamente para que no haya inconvenientes en el Desarrollo del Proyecto.

Finalmente, en este Artículo de Investigación se ha desarrollado un Modelo de Ciclo de Vida que sirve para llevar a cabo Proyectos de Explotación de Información, tomando la mejora que ha hecho Vanrell [25] a la metodología CRISP-DM [6] con los SubProcesos de Administración de Proyectos, y realizando un enfoque a los Riesgos tomando como base el Modelo de Ciclo de Vida en Espiral desarrollado por Boehm para Proyecto Convencionales [1].

Se puede dictaminar que el Modelo propuesto en este Artículo de Investigación es superior a la Metodología CRISP-DM [6] ya que tiene en cuenta los SubProcesos de Administración de Proyectos que se realizan a lo largo de todo el desarrollo del mismo, y principalmente le agrega un Análisis enfocado a los Riesgos que dentro de los mismos pueden ocurrir.

Por otro lado se ha construido un Mapa de Actividades, el cual se encuentra constituido por las actividades pertenecientes a los SubProcesos de Administración y Desarrollo de Proyecto correspondiente al Modelo de Procesos realizado por Vanrell [25], y por las fases que se han conformado en el Modelo de Ciclo de Vida en Espiral construido en este artículo.

A partir de allí, se definieron las actividades del Modelo de Procesos que se realizan en cada fase del Modelo de Ciclo de Vida, justificando su elección.

Por otro lado, las respuestas a los interrogantes planteados en la sección 3 pueden ser resumidas en las siguientes conclusiones:

- Se ha logrado llevar a cabo la construcción de un Modelo de Ciclo de Vida que es útil para el desarrollo de distintos Proyectos de Explotación de Información y que posee un Análisis de Riesgo, cuya característica no existía en las metodologías existentes hasta el momento.
- Finalmente, en este artículo se ha construido un Mapa de Actividades que posee como ejes en primer lugar las actividades del Modelo de Procesos desarrollado por Vanrell [25], tanto de Administración como de Desarrollo, y en segundo lugar las fases pertenecientes al Modelo de Ciclo de Vida propuesto en el presente Trabajo.

B. Futuras líneas de Investigación

Se plantea como futuras líneas de trabajos poder llevar a cabo el desarrollo de una cantidad más amplia y variada de Proyectos de Explotación de Información para continuar evaluando la viabilidad de las Propuestas desarrolladas en el presente Artículo de Investigación.

Si bien el Mapa de Actividades contempla todas las actividades pertenecientes al Modelo de Procesos desarrollado por Vanrell [25], resulta un punto de interés seguir analizándolo para evaluar las actividades aplicadas en cada fase del Modelo de Ciclo de Vida. Además, por medio de casos de prueba, se puede verificar la falta de alguna actividad dentro de alguna fase del Modelo de Ciclo de Vida construido.

II. REFERENCIAS

- [1] Boehm, B.W., C. Abts, A.W. Brown, S. Chulani, B.K.Clark, E. Horowitz, R. Madachy, D. Reifer, B. Steece.(2000). Software Cost Estimation with COCOMO II, Prentice-Hall, Englewood Cliffs Boehm, Barry W., TRW Defense Systems Group (1988), A Spiral Model of Software Development and Enhancement

- [2] Britos, P. (2008a). Procesos de Explotación de Información Basados en Sistemas Inteligentes. Tesis Doctoral Universidad Nacional de la Plata. La Plata. Buenos Aires. Argentina.
- [3] Britos, P., Dieste, O., García-Martínez, R. 2008. Requirements Elicitation in Data Mining for Business Intelligence Projects. En *Advances in Information Systems Research, Education and Practice*. David Avison, George M. Kasper, Barbara Pernici, Isabel Ramos, Dewald Roode Eds. (Boston: Springer), IFIP Series, 274: 139–150.
- [4] Britos, P., Grosser, H., Rodríguez, D., Garcia-Martinez, R. (2008b). Detecting Unusual Changes of Users Consumption. En *Artificial Intelligence and Practice II*, Max Bramer Ed. (Boston: Springer), IFIP Series, 276: 297-306.
- [5] Britos, P., Jiménez Rey, E., Garcia-Martínez, E. (2008c). Work in Progress: Programming Misunderstandings Discovering Process Based On Intelligent Data Mining Tools. *Proceedings 38th ASEE/IEEE Frontiers in Education Conference*.
- [6] Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C. y Wirth, R. (2000). CRISP-DM 1.0 Step-by-step Data Mining guide. U.S.A. <http://www.crisp-dm.org/>. Página web vigente al 1/09/09.
- [7] Cogliati, M., Britos, P., García-Martínez, R. (2006) Patterns in Temporal Series of Meteorological Variables Using SOM & TDIDT In: Bramer M (ed) *Artificial Intelligence in Theory and Practice*, Boston, Springer, IFIP Series 217:305-314
- [8] D’Atri, M., Rodriguez, D., Garcia-Martínez, R. 2009. Improving Pipeline Risk Models by Using Data Mining Techniques. 24th World Gas Conference Proceedings (CD). Paper 663.
- [9] Felgaer, P., Britos, P., and García-Martínez, R. (2006) Prediction in Health Domain Using Bayesian Network Optimization Based on Induction Learning Techniques. *Int. J. of Mod. Ph. C* 17(3): 447-455
- [10] Flores, D., Garcia-Martinez, R. Fernandez, E., Merlino, H., Rodriguez, D., Britos, P. (2009). Detección de Patrones para la Prevención de Daños y/o Averías en la Industria Automotriz. *Proceedings XV Congreso Argentino de Ciencias de la Computación. Workshop de Base de Datos y Minería de Datos*. Págs. 1021-1030. ISBN 978-897-24068-4-1.
- [11] García-Martínez, R., Britos, P., Pesado, P., Bertone, R., Pollo-Cattaneo, F., Rodríguez, D., Pytel, P., Vanrell, J. (2011b). Towards an Information Mining Engineering. En *Software Engineering: Methods, Modeling, and Teaching* (ed. C. Zapata Jaramillo). Editorial de la Universidad de Medellín (en prensa).
- [12] García-Martínez, R., Lelli, R., Merlino, H., Cornachia, L., Rodriguez, D., Pytel, P., Arboleya, H. (2011a). Ingeniería de Proyectos de Explotación de Información para PYMES. *Proceedings XIII Workshop de Investigadores en Ciencias de la Computación*. Pág. 253-257.
- [13] Grosser, H., Britos, P., García-Martínez, R. (2005) Detecting Fraud in Mobile Telephony Using Neural Networks. *LNAI* 3533:613-615.
- [14] Kuna, H., García Martínez, R. Villatoro, F. 2010. Pattern Discovery in University Students Desertion Based on Data Mining. *Advances and Applications in Statistical Sciences Journal*, 2(2): 275-286. ISSN 0974-6811
- [15] Langseth, J., Vivatrat, N. (2003). Why Proactive Business Intelligence is a Hallmark of the Real-Time Enterprise: Outward Bound. *Intelligent Enterprise* 5(18): 34-41.
- [16] Lemus, M. A. (2005). Modelo de Procesos para la Industria de Software. Secretaría de Economía de México. <http://www.comunidadmoprosoft.org.mx/>. Página web vigente al 1/09/09.
- [17] Lockheed-Martin, (1998). LLC PRICE Systems, PRICE S Reference Manual Version 3.0, Lockheed-Martin.
- [18] Negash, S., Gray, P. (2008). Business Intelligence. En *Handbook on Decision Support Systems 2*, ed. F. Burstein y C. Holsapple (Heidelberg, Springer), Pág. 175-193.
- [19] Oktaba, H., Piattini, M., Pino, F.J., Orozco, M.J. y Alquicira, C. (2008). *Competisoft, Mejora de Procesos Software para Pequeñas y Medianas Empresas y Proyectos*. Ra-Ma.
- [20] Oktaba, H., Alquicira Esquivel, C., Ramos, A. S., Martínez Martínez, A., Quintanilla Ozorio, G., RuvalcabaLópez, M., López Lira Hinojo, F., Rivera López, M. E., Orozco Mendoza, M. J., Fernández Ordoñez, Y. y Flores
- [21] Oktaba, H., Alquicira Esquivel, C., Ramos, A. S., Palacios Elizalde, J., Pérez Escobar, C. J. y López Lira Hinojo, F. (2004). Método de Evaluación de Procesos para la Industria de Software. Secretaría de Economía de México. <http://www.software.net.mx/NR/rdonlyres/ED7B3399-0CA4-412E-9FAC-EEB94F85C5F/1224/EvalProSoftv11.pdf>. Página web vigente al 1/09/09.
- [22] Schiefer, J., Jeng, J., Kapoor, S., Chowdhary, P. (2004). Process Information Factory: A Data Management Approach for Enhancing Business Process Intelligence. *Proceedings 2004 IEEE International Conference on E-Commerce Technology*. Pág. 162-169.
- [23] Stefanovic, N., Majstorovic, V., Stefanovic, D. (2006). Supply Chain Business Intelligence Model. *Proceedings 13th International Conference on Life Cycle Engineering*. Pág. 613-618.
- [24] Vanrell, J., Bertone, R., García Martínez, R. Modelo de Proceso de Operación para Proyectos de Explotacion de Información. *Anuales del XVI Congreso Argentino de Ciencias de la Computación*. Pág. 674-682. ISBN 978-950-9474-49-9.
- [25] Vanrell, J. A., García-Martínez, R., Bertone, R. (2010b). Un Modelo de Procesos de Explotación de Información. *Proceedings XII Workshop de Investigadores en Ciencias de la Computación*. Pág. 167-171. ISBN 978-950-34-0652-6.
- [26] Vanrell, J. A. (2009). Elementos para un Modelo de Procesos de Explotación de Información para PyMES. Trabajo de Especialidad en Ingeniería en Sistemas de Información. Facultad Regional Buenos Aires. Universidad Tecnológica Nacional. <http://posgrado.frba.utn.edu.ar/investigacion/especialidades/vanrelltrabajofinaldeespecialidad.pdf>. Página vigente al 24/07/10.



Hernan Arboleya. Es Analista Programador Universitario y Licenciado en Sistemas por la Universidad Nacional de Lanús. Es Asistente de Docencia en las Asignaturas Ingeniería de Software I y Proyecto de Software. Es Asistente de Investigación del Laboratorio de Investigación y Desarrollo en Ingeniería de Explotación de Información en el Grupo de Investigación en Sistemas de Información del Departamento de Desarrollo Productivo y Tecnológico de la Universidad Nacional de Lanús.