



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

MEJORA DEL PROCESO DE SELECCIÓN DE PROYECTOS DE INVESTIGACIÓN Y DESARROLLO EN MOLYMET

*PROYECTO DE GRADO PARA OPTAR AL GRADO DE MAGISTER EN
INGENIERÍA DE NEGOCIOS CON TECNOLOGÍAS DE INFORMACIÓN*

CHRISTIAN IGNACIO SOTO INFANTE

PROFESOR GUÍA

EZEQUIEL MUÑOZ KRSULOVIC

MIEMBROS DE LA COMISIÓN

JUAN JOSÉ SEGURA FARÍAS

LUCIANO VILLAROEL PARRA

JAIME CONTESSE MARROQUÍN

SANTIAGO DE CHILE
2018

RESUMEN EJECUTIVO

En las empresas del área de los *commodities* el crecimiento siempre está acotado a la demanda mundial del producto. Un crecimiento mayor solo puede lograrse a partir de la exploración de nuevos negocios y la empresa ha depositado su confianza en la innovación como el motor generador de proyectos conducentes a este objetivo.

Las ideas de proyectos conducentes a nuevos negocios deben provenir de información externa a la empresa, puesto que la información interna apunta solo a la mejora continua. En este sentido, se ha utilizado la vigilancia tecnológica como herramienta de análisis del entorno competitivo, aunque sin contar con un proceso bien definido y coherente dada la gran cantidad de información disponible. En el presente proyecto se propone el rediseño de este proceso, incluyendo una herramienta tecnológica que permita realizar un análisis global de toda la información disponible. El conocimiento generado se utilizará para apoyar la toma de decisiones respecto a la selección de proyectos de investigación. Una toma de decisiones informada permite además disminuir la incerteza y por lo tanto aumentar el índice de proyectos exitosos.

Para el desarrollo se utilizaron las metodologías de ingeniería de negocios de Barros y la de extracción de conocimiento CRISP-DM. Se utilizaron modelos de extracción de tópicos y se determinó que LDA es el modelo que funciona mejor dada la naturaleza del *cluster* objetivo, obteniéndose tópicos y subtópicos perfectamente interpretables. Además, se construyó una herramienta de visualización que facilita la selección del número óptimo de tópicos y que permite exportar la información en forma de reportes HTML, y un sistema de recomendación de documentos basado en citas.

La evaluación económica y el análisis de sensibilidad permiten concluir que el proyecto se justifica económicamente siempre que el aumento en el índice de proyectos exitosos sea mayor a 0.13%. Valor relativamente bajo si se considera que el índice de proyectos exitosos actual ronda el 11%.

Se espera que la aplicación de un proyecto de esta naturaleza signifique un incentivo para crear proyectos de mejora similares, que utilicen tecnologías de información como una alternativa a los proyectos de ingeniería tradicional.

Al recuerdo de su mirada, su abrazo confortable,
su abnegación, su terquedad y por sobre todo su sonrisa sincera
Violeta, la flor que no está pero que se niega a desaparecer

AGRADECIMIENTOS

Me gustaría agradecer a todos aquellos que se detuvieron un segundo a brindarme su apoyo y a ofrecerme un abrazo confortable sobre todo en los momentos de frustración. Aprovecho también de disculparme si alguien se sintió desplazado o dejado de lado; fueron muchas las horas dedicadas a esto, espero me comprendan.

Agradezco a Ezequiel Muñoz y a Constanza Contreras por su apoyo y guía en el desarrollo de este proyecto.

Agradezco a Molymet, y en especial a Manuel Guzmán, por apoyarme tanto económicamente como con tiempo laboral, cuando decidí estudiar este Magister. Agradezco a Juan José Segura por apoyar el proyecto y ayudarme sobre todo en la parte de validación de los resultados.

Agradezco a mi familia y a mis compañeros de vida, por constituir el pilar fundamental en donde se sustenta y fortalece mi estabilidad emocional y mi incentivo para lograr nuevas metas.

Finalmente dedico unas palabras a Loreto, el espejo que refleja lo que siempre necesito mirar. Me encanta como transforma el peor de los días en uno lleno de luz y alegría. Es imposible expresar en palabras todo mi agradecimiento. Gracias por el cariño más desinteresado y sincero del mundo, solo puedo decir que estoy eternamente en deuda contigo.

TABLA DE CONTENIDO

RESUMEN EJECUTIVO	ii
CAPÍTULO 1: INTRODUCCIÓN Y CONTEXTO	10
1.1 ANTECEDENTES DE LA INDUSTRIA.....	10
1.2 DESCRIPCIÓN GENERAL DE LA EMPRESA	12
1.3 PROBLEMA U OPORTUNIDAD IDENTIFICADA	16
1.4 OBJETIVOS Y RESULTADOS ESPERADOS DEL PROYECTO.....	18
1.4.1 <i>Objetivo General</i>	18
1.4.2 <i>Objetivos Específicos</i>	18
1.4.3 <i>Resultados Esperados</i>	19
1.5 ALCANCE	19
1.6 RIESGOS POTENCIALES	20
CAPÍTULO 2: MARCO TEÓRICO	22
2.1 METODOLOGÍA DE INGENIERÍA DE NEGOCIOS	22
2.2 MINERÍA DE DATOS	24
2.2.1 <i>Introducción a la minería de datos</i>	24
2.2.2 <i>El proceso de la minería de datos</i>	25
2.2.3 <i>Minería de textos</i>	26
2.2.4 <i>Procesos para la implementación de minería de datos</i>	29
CAPÍTULO 3: PLANTEAMIENTO ESTRATÉGICO Y MODELO DE NEGOCIOS	31
3.1 POSICIONAMIENTO ESTRATÉGICO.....	31
3.2 LINEAMIENTOS ESTRATÉGICOS	32
3.3 MODELO DE NEGOCIOS	34
CAPÍTULO 4: ANÁLISIS SITUACIÓN ACTUAL	35
4.1 ARQUITECTURA DE PROCESOS.....	35
4.2 MODELAMIENTO DETALLADO DE PROCESOS	36
4.3 DIAGNÓSTICO DE LA SITUACIÓN ACTUAL.....	40
4.4 CUANTIFICACIÓN DEL PROBLEMA U OPORTUNIDAD	41
CAPÍTULO 5: PROPUESTA DE DISEÑO DE PROCESOS	44
5.1 DIRECCIONES DE CAMBIO Y ALCANCE.....	44
5.2 DISEÑO DETALLADO DE PROCESOS TO BE	46

5.3	DISEÑO DE LÓGICA DE NEGOCIOS	49
5.3.1	<i>Diseño del piloto</i>	49
5.3.2	<i>Selección del modelo de clustering</i>	50
5.3.3	<i>Selección del número de tópicos</i>	52
5.3.4	<i>Caracterización de tópicos</i>	55
5.3.5	<i>Sistema de recomendación</i>	56
5.3.6	<i>Subtópicos</i>	57
CAPÍTULO 6: PROPUESTA DE APOYO TECNOLÓGICO		59
6.1	ESPECIFICACIÓN DE REQUERIMIENTOS	59
6.1.1	<i>Requerimientos funcionales</i>	59
6.1.2	<i>Requerimientos no funcionales</i>	60
6.2	ARQUITECTURA TECNOLÓGICA	60
6.3	DISEÑO DE LA APLICACIÓN	62
6.3.1	<i>Diagrama de casos de uso</i>	62
6.3.2	<i>Diagrama de secuencia</i>	63
6.4	PROTOTIPO FUNCIONAL DESARROLLADO	64
CAPÍTULO 7: GESTIÓN DEL CAMBIO		68
7.1	CONTEXTO DEL CAMBIO	68
7.2	OBSERVACIÓN DE LA IMPLEMENTACIÓN A REALIZAR	68
7.3	ANÁLISIS DE LOS PRINCIPIOS DE CAMBIO	69
7.4	CARACTERIZACIÓN DEL CAMBIO	70
7.5	FACTORES CRÍTICOS DE ÉXITO	70
7.6	PLAN DE GESTIÓN DEL CAMBIO	71
CAPÍTULO 8: EVALUACIÓN DEL PROYECTO		73
8.1	RESULTADOS Y ANÁLISIS DEL PILOTO	73
8.2	DEFINICIÓN DE BENEFICIOS Y COSTOS	75
8.3	FLUJO DE CAJA	78
8.4	ANÁLISIS DE SENSIBILIDAD	79
CAPÍTULO 9: CONCLUSIONES		81
CAPÍTULO 10: BIBLIOGRAFÍA		85

ÍNDICE DE FIGURAS

FIGURA 1: DISTRIBUCIÓN DE LAS RESERVAS DE MOLIBDENO AL AÑO 2016.....	11
FIGURA 2: EVOLUCIÓN DEL PRECIO DEL MOLIBDENO EN EL PERIODO 1975-2017	12
FIGURA 3: PANORÁMICA DE LAS INSTALACIONES DE MOLYMET EN NOS AL AÑO 2017.....	13
FIGURA 4: ORGANIGRAMA DE MOLYMET A DICIEMBRE DE 2017	15
FIGURA 5: FUENTES PRINCIPALES DE IDEAS UTILIZADAS EN LA GENERACIÓN DE LOS PROYECTOS DE LA VPIT.....	17
FIGURA 6: ESQUEMA DE LA METODOLOGÍA DE INGENIERÍA DE NEGOCIOS DE BARROS	22
FIGURA 7: ESQUEMA GRÁFICO DEL MODELO LDA	28
FIGURA 8: ESQUEMA DE LA METODOLOGÍA CRISP-DM	30
FIGURA 9: POSICIONAMIENTO ESTRATÉGICO DE MOLYMET SEGÚN EL MODELO DELTA DE HAX.....	31
FIGURA 10: MODELO DE NEGOCIOS DE MOLYMET SEGÚN EL MODELO DE OSTERWALDER	34
FIGURA 11: VISIÓN GENERAL DE MACROPROCESOS DE MOLYMET.....	35
FIGURA 12: MACROPROCESO DE DESARROLLO DE NUEVAS CAPACIDADES	37
FIGURA 13: PROCESO DE EVALUACIÓN DE LA NECESIDAD DE UNA NUEVA CAPACIDAD	38
FIGURA 14: BPMN DEL PROCESO DE INNOVACIÓN DE CAPACIDADES AS IS.....	39
FIGURA 15: RESULTADOS ESPERABLES DE LA IMPLEMENTACIÓN DE VIGILANCIA TECNOLÓGICA EN LA EMPRESA.....	42
FIGURA 16: PROCESO DE EVALUACIÓN DE LA NECESIDAD DE UNA NUEVA CAPACIDAD TO BE.....	46
FIGURA 17: BPMN DEL PROCESO DE INNOVACIÓN DE CAPACIDADES TO BE	48

FIGURA 18: RESULTADOS COMPARATIVOS ENTRE LOS DISTINTOS MODELOS DE CLUSTERING.....	52
FIGURA 19: RESULTADOS COMPARATIVOS DE CLUSTERING PARA 6 Y 7 TÓPICOS	54
FIGURA 20: DISTRIBUCIÓN DE TÓPICOS PARA LA BÚSQUEDA DEFINIDA..	55
FIGURA 21: DISTRIBUCIÓN DE SUBTÓPICOS PARA EL TÓPICO 2	58
FIGURA 22: ARQUITECTURA TECNOLÓGICA UTILIZADA	61
FIGURA 23: DIAGRAMA DE CASOS DE USO DE LA APLICACIÓN	62
FIGURA 24: DIAGRAMA DE SECUENCIA PARA EL CASO DE USO DE ANÁLISIS DE TÓPICOS	63
FIGURA 25: CAPTURA DE LA APLICACIÓN DE VISUALIZACIÓN DE RESULTADOS DESARROLLADA	64
FIGURA 26: REPORTE HTML GENERADO PARA LA BÚSQUEDA DEFINIDA EN EL PILOTO	66
FIGURA 27: REPORTE HTML GENERADO PARA LA BÚSQUEDA DEFINIDA EN EL PILOTO. SE INCLUYE EL ANÁLISIS DE SUBTÓPICOS	67

ÍNDICE DE TABLAS

TABLA 1: DEMANDA DE MOLIBDENO (EN MILES DE TONELADAS) PARA 2015-2016 Y EXPECTATIVAS DE DEMANDA PARA EL PERIODO 2017-2018	12
TABLA 2: INVERSIÓN EN I+D POR PARTE DE MOLYMET EN COMPARACIÓN CON EL PROMEDIO DE LA INDUSTRIA CHILENA	16
TABLA 3: RIESGOS POTENCIALES INVOLUCRADOS EN LA REALIZACIÓN DEL PROYECTO.....	21
TABLA 4: VARIABLES DE DISEÑO DE ESTRUCTURA DE EMPRESA Y MERCADO.....	44
TABLA 5: VARIABLES DE DISEÑO DE ANTICIPACIÓN	44
TABLA 6: VARIABLES DE DISEÑO DE COORDINACIÓN	44
TABLA 7: VARIABLES DE DISEÑO DE PRÁCTICAS DE TRABAJO	45
TABLA 8: VARIABLES DE DISEÑO DE INTEGRACIÓN DE PROCESOS CONEXOS.....	45
TABLA 9: VARIABLES DE DISEÑO DE MANTENCIÓN DE ESTADO	45
TABLA 10: CARACTERIZACIÓN DE LOS TÓPICOS OBTENIDOS.....	55
TABLA 11: EJEMPLO DE EXPORTACIÓN DE INFORMACIÓN PARA EL TÓPICO 3 DE LA TABLA 10	56
TABLA 12: CARACTERIZACIÓN DE LOS SUBTÓPICOS OBTENIDOS PARA EL TÓPICO 2	57
TABLA 13: RESULTADOS DEL ANÁLISIS DE INTERPRETACIÓN DE TÓPICOS	73
TABLA 14: RESULTADOS DEL ANÁLISIS DE INTERPRETACIÓN DE SUBTÓPICOS	73
TABLA 15: RESULTADOS OBTENIDOS EN LA VALIDACIÓN DE TÓPICOS	75
TABLA 16: COSTOS DE IMPLEMENTACIÓN DEL PROYECTO	77
TABLA 17: COSTOS OPERACIONALES DEL PROYECTO	77
TABLA 18: FLUJO DE CAJA DEL PROYECTO	79
TABLA 19: ANÁLISIS DE SENSIBILIDAD DEL VAN RESPECTO AL AUMENTO EN EL ÍNDICE DE PROYECTOS EXITOSOS	80

CAPÍTULO 1: INTRODUCCIÓN Y CONTEXTO

1.1 Antecedentes de la industria

El molibdeno es un metal de color gris plateado, de número atómico 42, que no existe en estado puro en la naturaleza, por lo que generalmente se encuentra asociado a otros elementos, como es el caso de los minerales sulfurados, de los cuales también se obtiene el cobre. Así, es común que el molibdeno sea considerado un subproducto de la operación de extracción de cobre.

Alrededor de las dos terceras partes de este metal se usan para la producción de acero inoxidable, incluyendo concentraciones variables de hasta un 6%. La aleación de acero soporta altas temperaturas y presiones por lo que se utiliza en construcción, piezas de aviones, piezas forjadas de automóviles, etc. El alambre de molibdeno se usa en tubos electrónicos y el metal sirve también como electrodo en los hornos de vidrio. Otro de los usos principales del molibdeno es la súper aleación que se obtiene en base a níquel, para obtener catalizadores que se usan en la remoción de azufre en la industria petrolera.

La producción de molibdeno proviene de dos fuentes:

- **Primaria:** Cuya producción deriva de yacimientos donde el molibdeno es el producto principal.
- **Secundaria:** La extracción como subproducto de la minería del cobre.

Chile ha sido históricamente uno de los países con las mayores reservas y producción mundial de molibdeno, junto con China y EE. UU., debido a la asociación con algunos pórfidos cupríferos masivos (Chuquicamata, Formación Río blanco, El Teniente, Pelambres, entre otros). La figura 1 muestra la distribución de las reservas de molibdeno al año 2016.

El precio del molibdeno ha sido fuertemente influenciado por el momento histórico, observándose valores muy altos que llegaron a superar los 30 US\$/lb

(boom de los *commodities*, 2005-2008) y valores por debajo de los 8 US\$/lb (crisis *subprime*, 2009-2015).

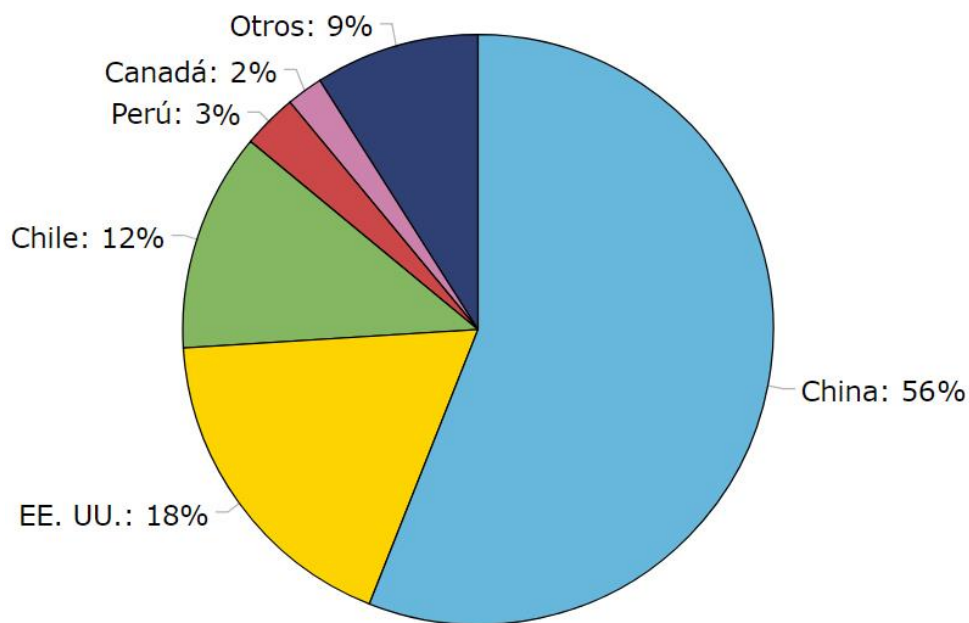


Figura 1. Distribución de las reservas de molibdeno al año 2016.¹

La figura 2 muestra la evolución de los precios del molibdeno desde 1975 a 2017, observándose un valor promedio de 13.1 US\$/lb.

La tabla 1 presenta la demanda total de molibdeno en 2015 y 2016 y las expectativas de demanda para los años 2017 a 2018.

¹ Fuente: Cochilco, Mercado internacional del molibdeno 2017.

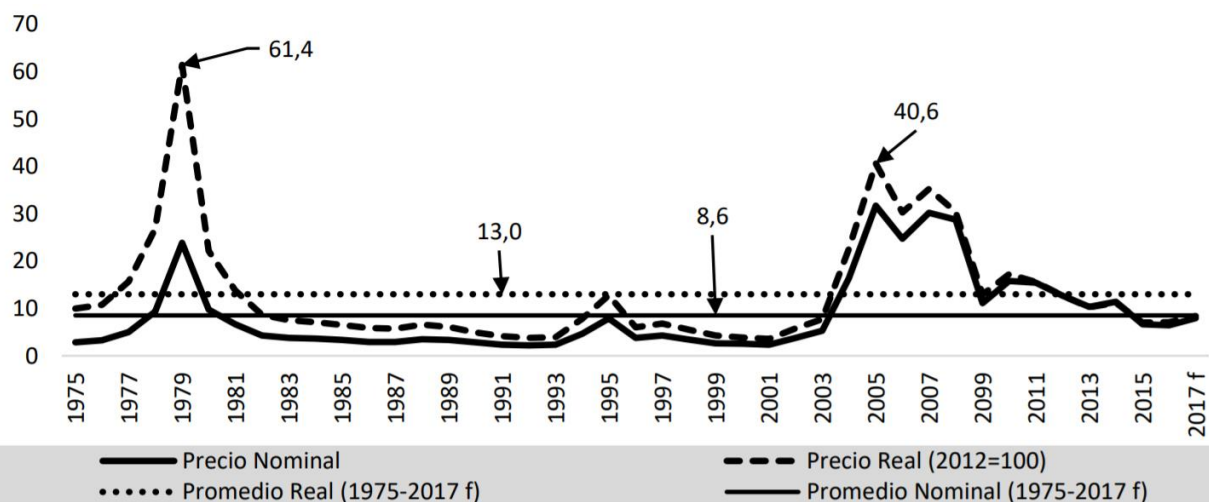


Figura 2. Evolución del precio del molibdeno en el periodo 1975-2017.¹

	2015	2016	2017 e	2018 e
Europa	70	71	74	77
EE. UU.	31	33	35	36
Japón	26	26	26	27
China	91	94	97	101
Otros	63	63	65	66
Mundo	281	286	296	307
Variación	-4.6%	1.8%	3.7%	3.6%

Tabla 1. Demanda de molibdeno (en miles de toneladas) para 2015-2016 y expectativas de demanda para el periodo 2017-2018.¹

1.2 Descripción general de la empresa

Molibdenos y metales S.A., mejor conocida por su nombre de fantasía Molymet, es la mayor procesadora mundial de molibdeno. Se formó en 1975 tras la división de Carburo y Metalurgia S.A., empresa dedicada principalmente a la producción de carburo de calcio y ferroaleaciones básicas. La empresa se fundó en la localidad de Nos en la comuna de San Bernardo, en la ciudad de Santiago de Chile, localidad que se ha mantenido hasta la fecha como sede

¹ Fuente: Cochilco, Mercado internacional del molibdeno 2017.

principal tanto a nivel de producción como a nivel corporativo. La figura 3 muestra una vista panorámica actual de las instalaciones de Molymet en Nos.



Figura 3. Panorámica de las instalaciones de Molymet en Nos al año 2017.

Los productos que comercializa la empresa tienen como origen principal la transformación del concentrado de molibdenita (MoS_2) en productos de valor agregado. Los productos de Molymet son comercializados, principalmente, en el mercado europeo, asiático y norteamericano, siendo utilizados por empresas de sofisticada tecnología de sectores tales como la industria aeroespacial, química, electrónica y siderúrgica.

Molymet lidera además la producción mundial de renio y tiene capacidad para procesar aproximadamente el 35% del molibdeno del mercado mundial.¹

La empresa posee plantas industriales en 5 países: Molymet en Santiago (Chile), Molynor en Mejillones (Chile), Molymex en Cumpas (México), Sadaci en Gante (Bélgica), Chemiemetall en Bitterfeld (Alemania) y Luoyang High-Tech Metals en Luoyang (China). Adicionalmente, posee oficinas comerciales en Londres, Beijing, Texas, Sao Paulo y Santiago de Chile.

¹ Fuente: Molymet, Memoria anual 2017.

Las principales líneas de negocio de Molybmet son las siguientes:

- Ventas propias: Molybmet compra concentrados de molibdenita y utiliza su propia tecnología para el tratamiento y procesamiento del concentrado, para así producir una amplia gama de productos de molibdeno (desde óxido de molibdeno hasta productos puros) y luego venderlos a distintos clientes en el mercado mundial.
- Maquila: Empresas de la industria minera le entregan molibdenita a Molybmet para que ésta sea procesada y luego devuelta a las empresas como un producto comercializable. Por este servicio se cobra un *fee*.
- Subproductos: Como resultado del proceso de tostación y oxidación al que se tiene que someter el molibdeno para su procesamiento, Molybmet recupera subproductos que luego comercializa en el mercado. Estos subproductos son principalmente renio, cobre y ácido sulfúrico.

La figura 4 presenta el organigrama de la empresa. El número de trabajadores al 31 de diciembre de 2017 es 1502.

La administración de la compañía destina importantes recursos a investigación, desarrollo e innovación, en una búsqueda constante para obtener nuevos o mejores productos y procesos alternativos asociados a la actividad de Molybmet y de sus filiales. Por tal razón se constituyó la Vicepresidencia de Investigación, Desarrollo e Innovación Tecnológica (VPIT) como una unidad autónoma, con dependencia solo del Presidente Ejecutivo, encargada del estudio de nuevos procesos que apunten a generar nuevos negocios para la compañía y/o mejora continua y optimización de los procesos actuales a través de la integración de tecnología de punta, con el fin de asegurar las mayores cotas de calidad para los productos y un bajo costo de producción.

La VPIT cuenta con 18 personas con dedicación exclusiva e instalaciones propias, incluyendo laboratorios equipados especialmente para el estudio de procesos y el desarrollo de actividades experimentales tanto a escala laboratorio como a escala piloto.

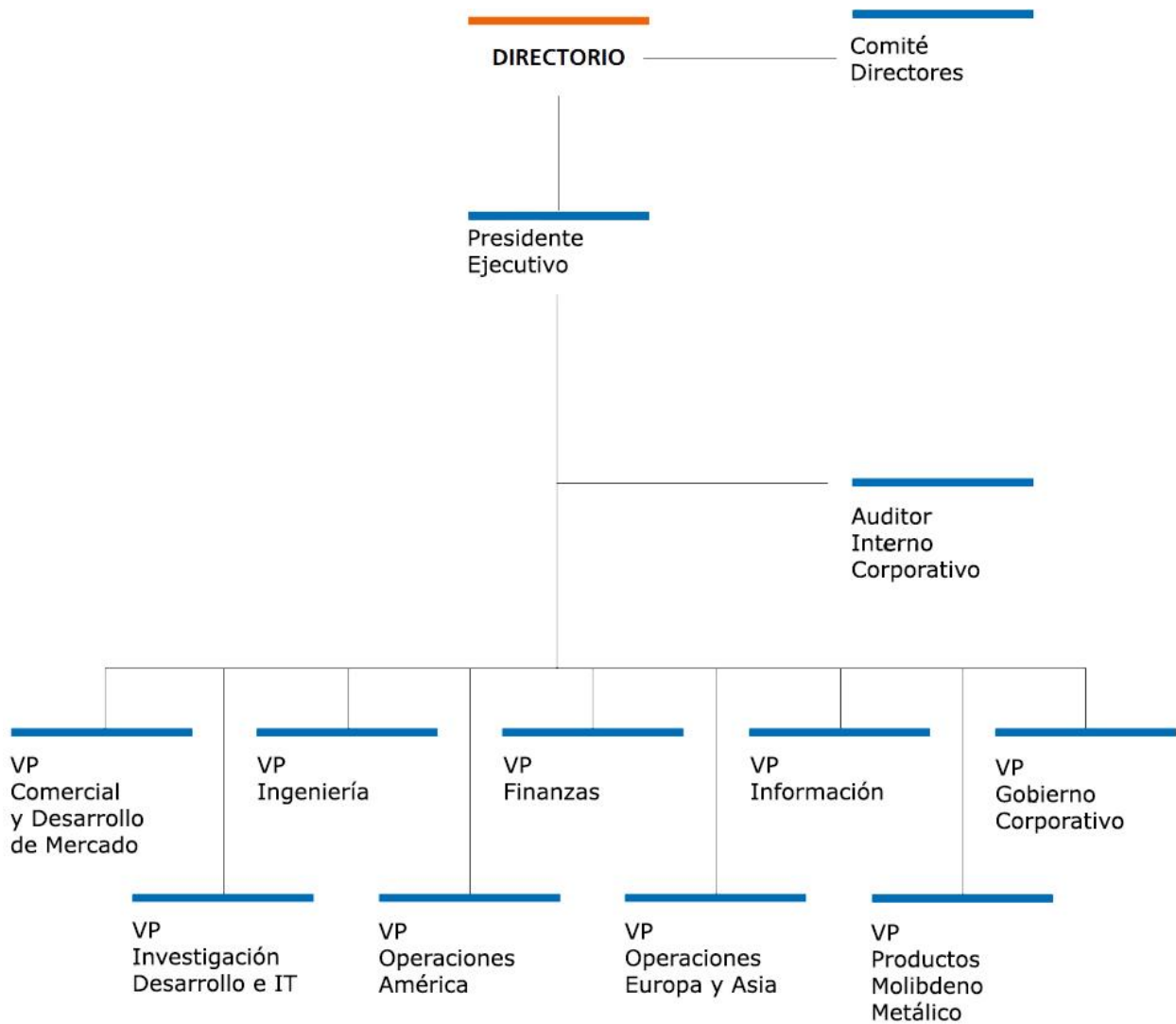


Figura 4. Organigrama de Molymet a diciembre de 2017.

Inicialmente la vicepresidencia era solo de investigación y desarrollo, pero a finales del 2012 se decidió añadir la innovación como eje central de los proyectos, de manera de potenciar el beneficio económico y al mismo tiempo impulsar la innovación al resto de la compañía.

La importancia que tiene Investigación y Desarrollo (I+D) para la empresa se justifica en la inversión que se realiza en ella. La tabla 2 muestra esta inversión en comparación con la inversión que realiza el promedio de la industria chilena.

	INVERSIÓN I+D (% VENTAS)	PROFESIONALES DEDICADOS A I+D (CADA 1000 TRABAJADORES)
MOLYMET	0.34%	7.4
INDUSTRIA CHILENA	0.25%	1.8

Tabla 2. Inversión en I+D por parte de Molymet en comparación con el promedio de la industria chilena.¹

1.3 Problema u oportunidad identificada

Los proyectos de investigación y desarrollo que realiza la VPIT comienzan con una idea o problemática que puede provenir de distintas fuentes, como se muestra en la figura 5. Al respecto, se puede comentar lo siguiente:

- Vicepresidencia de operaciones: Consultas técnicas por problemas operacionales internos que impiden el funcionamiento normal de las plantas y que no son capaces de solventar con sus propios recursos.
- Presidencia ejecutiva y Vicepresidencia comercial: Propuestas de estudio de factibilidad técnica para el desarrollo de nuevos productos o mejoras de los actuales y consultas técnicas con respecto a nuevas inversiones. Se basan principalmente en consultas directas de clientes.
- Vigilancia tecnológica: Con el advenimiento de la innovación se añadió la vigilancia tecnológica como fuente externa de nuevas ideas a través de la observación constante de todas las publicaciones que dicen relación con el modelo de negocios de Molymet.

El producto principal de la compañía es un *commodity*, es decir, un producto genérico y sin mayor diferenciación respecto de lo que puedan ofrecer los competidores. Debido a lo anterior, el crecimiento de la empresa está limitado al crecimiento de la demanda mundial de este producto. Un crecimiento mayor

¹ Fuente: OECD

solo puede lograrse a partir de la creación de un nuevo negocio, entendiendo esto como la generación de un nuevo producto o servicio, para la empresa.

La VPIT es la unidad básica generadora de nuevos negocios para la empresa a partir del estudio de proyectos que conduzcan a ello. Hoy en día se realizan proyectos tanto de mejora continua como conducentes a nuevos negocios, siendo el primero de éstos el enfoque tradicional y el que suele englobar el mayor número de proyectos y por lo tanto de recursos. La cantidad de proyectos conducentes a nuevos negocios está actualmente entre un 10-15%. Como objetivo estratégico y como forma de impulsar el crecimiento de la empresa, se hace necesario aumentar este porcentaje. Otro tema importante dice relación con el índice de proyectos exitosos. Actualmente este índice está entre 10-12%, porcentaje bajo para un centro de esta categoría y potencialmente mejorable. La idea global es potenciar la innovación y el crecimiento de la empresa desde la VPIT, al aumentar la generación de nuevos negocios a partir de proyectos exitosos que conduzcan a ello.



Figura 5. Fuentes principales de ideas utilizadas en la generación de los proyectos de la VPIT.

Las fuentes externas de ideas son aquellas que tienen la potencialidad de generar un nuevo negocio para la empresa. De ellas, el proceso de vigilancia tecnológica es la única fuente sustentable de ideas, pues permite analizar el entorno competitivo e incluso anteponerse a los movimientos futuros del mercado sin depender de consultas específicas.

Hoy en día no existe un proceso sistemático para realizar vigilancia tecnológica en Molyt y se entiende como tal, al análisis manual de algunos documentos escogidos según los intereses del analista o por alguna recomendación particular. Lo anterior es claramente ineficiente puesto que se hace inmanejable procesar la inmensa cantidad de información disponible de manera manual. Un proceso inteligente y bien definido de vigilancia tecnológica, con la inclusión de apoyo tecnológico, permitiría analizar grandes volúmenes de información en tiempos razonables, por lo que se vislumbra como una solución prometedora y perfectamente integrable a la operación actual. Esto ayudaría a aumentar el número de proyectos conducentes a nuevos negocios y al mismo tiempo permitiría aumentar el número de proyectos exitosos, puesto que la información permite disminuir el riesgo en la toma de decisiones acerca de que proyecto realizar a partir de una detección eficaz de amenazas y cambios del entorno.

1.4 Objetivos y resultados esperados del proyecto

1.4.1 Objetivo general

Mejorar la selección de proyectos de investigación y desarrollo con enfoque en la generación de nuevos negocios, mediante vigilancia tecnológica inteligente incluyendo modelos de minería de texto.

1.4.2 Objetivos específicos

- Desarrollar un modelo que permita la obtención de información masiva desde una base de datos online.
- Determinar el mejor modelo de minería de textos para el análisis en cuestión.

- Determinar las tendencias que hay en los documentos obtenidos y un sistema de recomendación de documentos en base a citas.
- Desarrollar una interfaz de consulta de tendencias para una búsqueda cualquiera.
- Comprobar la funcionalidad real y la utilidad de mejora de negocio que presenta la propuesta, a partir del análisis crítico de los resultados de un piloto de la aplicación, incluyendo juicio experto.

1.4.3 Resultados Esperados

- Crawler que permita la obtención de información desde una base de datos online. Plazo objetivo: octubre 2017.
- Determinación del modelo óptimo de minería de datos para el caso en estudio. Plazo objetivo: diciembre 2017.
- Modelo para la obtención de las tendencias de una búsqueda objetivo y recomendación de documentos en base a citas. Plazo objetivo: abril 2018.
- Aplicación de visualización de resultados con generación automática de reportes incluyendo la información obtenida. Plazo objetivo: junio 2018.
- Resultados del piloto de la aplicación y conclusión acerca de la utilidad de negocio de la propuesta. Plazo objetivo: agosto 2018.

1.5 Alcance

La mejora propuesta se realizará a partir del rediseño del proceso de I+D con énfasis en la selección de proyectos con apoyo tecnológico. La solución solo se aplicará en la Vicepresidencia de Investigación, Desarrollo e Innovación Tecnológica de MolyMet y no se aplicará en otras áreas de la empresa ni en áreas similares de otras empresas.

La mejora propuesta solo se enfoca en el proceso de selección de proyectos mediante vigilancia tecnológica. Las demás fuentes de ideas para proyectos de investigación de la figura 5 se dejarán tal como operan en la actualidad.

Las fuentes de información relevantes para los proyectos de investigación y desarrollo se encuentran en publicaciones científicas y en patentes de invención. Las publicaciones científicas representan lo que probablemente se hará en el futuro de una industria en cuestión mientras que las patentes representan algo probado y que ya se hace. En este sentido, la información contenida en las publicaciones científicas permitiría anteponerse y realizar algo que no se ha hecho en el mercado, por lo que representan innovaciones con mejores ingresos potenciales, aunque más riesgosas que aquellas seleccionadas a partir de la información contenida en patentes. El presente estudio se realizará con la base de datos de publicaciones científicas Scopus, que contiene información muy atinente al negocio de Molytmet y de los posibles mercados de interés.

1.6 Riesgos potenciales

La tabla 3 presenta un resumen de los riesgos potenciales detectados para el proyecto.

Riesgo	Tipo	Probabilidad	Impacto (1-10)	Mitigación
Extensión en el tiempo de desarrollo propuesto	Gestión de proyecto	70%	2	Realizar actividades de acuerdo con la carta Gantt comprobando siempre el estado de avance en las reuniones con tutor y sponsor
Extensión en los plazos propuestos por el MBE	Gestión de proyecto	70%	2	Solicitar extensión de tiempo para el desarrollo de la propuesta según sea necesario
Cambios en el alcance del proyecto	Gestión de proyecto	20%	6	Ajustar el proyecto al alcance actual y dejar nuevos requerimientos como trabajo futuro
Reestructuración de personal que incluya al sponsor	Gestión de proyecto	5%	10	Crear nueva estrategia para mostrar el proyecto al nuevo sponsor
Reestructuración de personal que incluya al tesista	Gestión de proyecto	5%	10	Concluir el proyecto de cualquier modo y comparar con información de años anteriores de la empresa
Fiabilidad de las fuentes de información	Información	5%	10	Utilización de bases de datos de primera categoría como Science Direct por ejemplo
Permisos de conexión a bases de datos externas	Información	50%	5	Tramitar permisos directamente en el área TI de la empresa con la justificación correspondiente
Modelo de minería de textos no apto para la propuesta	Modelamiento	5%	10	Utilización de varios modelos y comparación de resultados
Exceso de complejidad en el desarrollo de la propuesta	Modelamiento	20%	6	Utilización de un camino alternativo más sencillo como por ejemplo la utilización de modelos menos complejos
Incredulidad respecto a la efectividad de la propuesta	Gestión del cambio	40%	5	Presentación de resultados previos con muestras pequeñas de datos para mostrar efectividad
Rechazo a la utilización de la herramienta	Gestión del cambio	20%	8	Presentación de los beneficios que conlleva la utilización tanto en resultados como en disminución de tiempo
Restricción de publicación de tesis	Privacidad de información	90%	1	Limitar la publicación de la tesis por 5 años

Tabla 3. Riesgos potenciales involucrados en la realización del proyecto.

CAPÍTULO 2: MARCO TEÓRICO

2.1 Metodología de Ingeniería de Negocios

El proyecto planteado se desarrolló siguiendo la metodología de Ingeniería de Negocios de Oscar Barros (Barros, 2015). La figura 6 muestra un esquema de esta metodología.

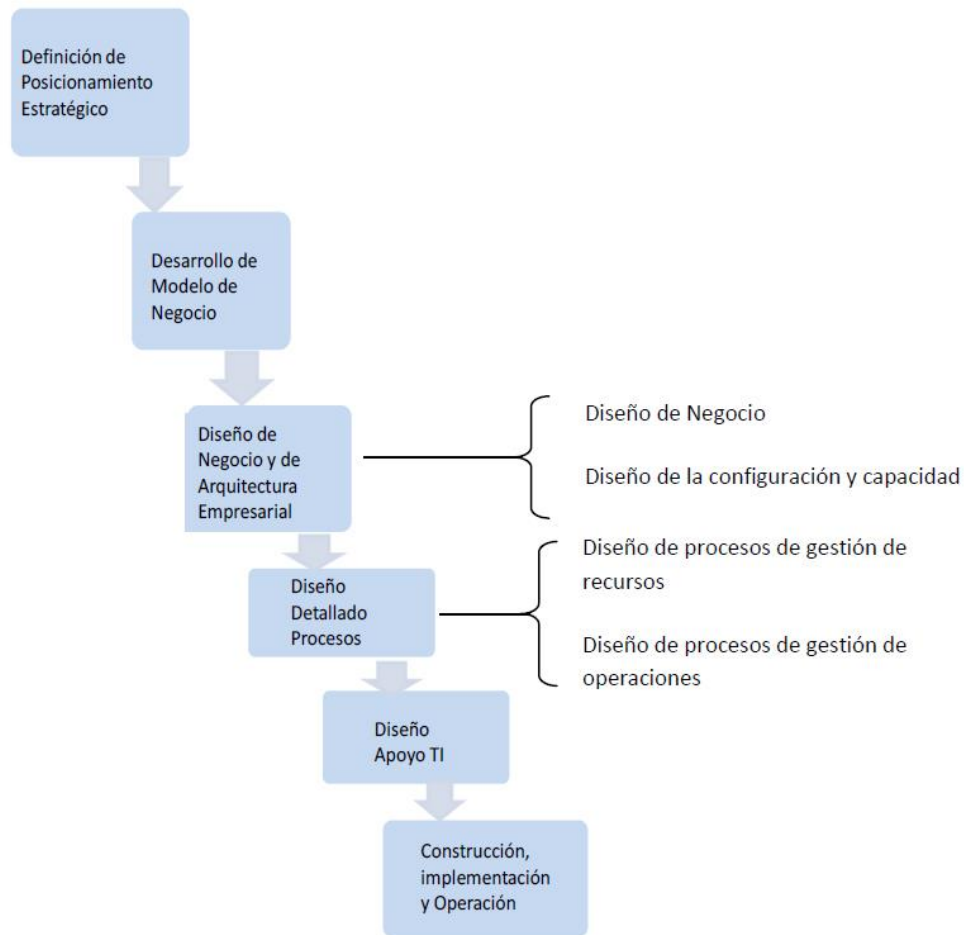


Figura 6. Esquema de la metodología de Ingeniería de Negocios de Barros (Barros, 2015).

Las etapas de esta metodología se describen a continuación:

- **Definición de posicionamiento estratégico:** Determina el curso a seguir de la empresa en forma de estrategias, que finalmente se convierten en planes que gobiernan a los demás procesos. Para su desarrollo se utilizan

herramientas como las 5 fuerzas de Porter (Porter, 1996) y el modelo Delta de Hax (Hax y Wilde, 2001).

- Desarrollo de modelo de negocio: Se define el valor que se entregará a los clientes enfocado en las acciones que se necesitan para alcanzar los objetivos estratégicos planteados anteriormente. Existen diversas herramientas que ayudan a definir el modelo de negocios, siendo una de las más utilizadas el modelo Canvas de Osterwalder (Osterwalder, 2004), por su simplicidad y su facilidad de entendimiento.
- Diseño de negocio y de arquitectura empresarial: Se define el diseño de negocio en base a patrones y la arquitectura empresarial necesaria para sustentar su diseño. Esta arquitectura se construye en base a macroprocesos, que representan todas las actividades de la empresa en forma de agrupaciones grandes de procesos (Barros, 2015).
- Diseño detallado de procesos: Considera el diseño detallado de todos aquellos procesos constituyentes de los macroprocesos. Este diseño considera varios niveles de profundidad, llegando incluso a la operación misma, y considera la utilización de la notación estándar de procesos de negocios (BPMN).
- Diseño apoyo TI: Considera el diseño del apoyo tecnológico que se requerirá y puede ir desde la definición de una arquitectura tecnológica empresarial hasta la definición de un apoyo tecnológico para un proceso en concreto.
- Construcción, implementación y operación: Considera la puesta en práctica tanto de los procesos diseñados como del apoyo tecnológico, sujeto a la factibilidad técnica y económica del negocio. Debe considerar además una implementación adecuada considerando la influencia de las personas (gestión del cambio).

2.2 Minería de datos

2.2.1 Introducción a la minería de datos

La minería de datos es el estudio que considera la recolección, limpieza, procesamiento, análisis y obtención de información útil desde fuentes de datos. Existe una amplia variación en cuanto a las posibles aplicaciones, así como también a los tipos de datos con los que se puede trabajar. Algunos ejemplos de diferentes tipos de datos son los siguientes:

- **Internet:** El orden de magnitud de los documentos pertenecientes a la web está en los miles de millones, esto sin considerar la información no indexada que es mucho más grande. Los usuarios acceden a los documentos, se crean registros de acceso a la web en los servidores y perfiles de comportamiento de clientes en sitios comerciales.
- **Interacciones financieras:** Transacciones comunes como el uso de un cajero automático o una tarjeta de crédito pueden crear datos de manera automática.
- **Interacciones de usuarios:** Muchas formas de interacciones entre usuarios pueden crear altos volúmenes de datos. Como por ejemplo las interacciones en redes sociales.
- **Sensores:** Mediciones en tiempo real de diversas condiciones tales como temperatura, presión, humedad, etc. Lo anterior se ha reforzado con el desarrollo de sensores de bajo costo acompañado de dispositivos móviles como los smartphones, que pueden comunicarse entre ellos.

Las técnicas de minería de datos permiten encontrar patrones no triviales en los datos y transformarlos en una estructura comprensible para su uso posterior. Los modelos pueden ser de dos tipos:

- **Modelos supervisados:** Predicen un dato o conjunto de datos desconocido, a partir de otros datos conocidos. Los modelos más comunes son los de clasificación y predicción.

- Modelos no supervisados: Descubren patrones y tendencias intrínsecas de los datos. Los modelos más comunes son los de segmentación y asociación.

Las técnicas más utilizadas en minería de datos son las siguientes:

- Árboles de decisión: Construcciones lógicas basadas en reglas de clasificación automática que sirven para representar y categorizar una serie de condiciones conducentes a la obtención de una resolución. En muchos casos estas reglas pueden ajustarse convenientemente para formar una estructura tipo árbol.
- Redes neuronales: Paradigma de aprendizaje inspirado en el funcionamiento del sistema nervioso humano. Se trata de un conjunto interconectado de nodos que colabora para producir un estímulo de salida, de manera similar a como lo hacen las neuronas.
- Regresiones: Construcciones matemáticas cuyo objetivo es generalizar un conjunto numérico de datos creando una ecuación que relaciona una o más variables de entrada con una variable de salida. La más típica es la regresión lineal.
- Agrupamiento: Procedimiento de agrupación según diversos criterios. Los más comunes incluyen agrupación por distancia y agrupación por densidad.

2.2.2 El proceso de la minería de datos

El proceso de la minería de datos contiene las siguientes fases:

- Recolección de datos: Es la etapa inicial del proceso, en donde se recopilan los datos que se utilizarán en las etapas posteriores. Dependiendo del tipo de dato, puede requerirse el uso de hardware especializado como por ejemplo sensores, o labores manuales como por ejemplo realización de encuestas.

- Limpieza de datos: Los datos recolectados rara vez se encuentran en un formato adecuado para trabajar con ellos, por lo que se hace necesario realizarles una transformación previa. El tipo de transformación necesaria depende del tipo de datos disponibles y del tipo de procesamiento que se pretenda hacer. Además, puede ser necesario remover data errónea o estimar datos faltantes según sea necesario. Finalmente, es necesario integrar los datos en un solo formato coherente.
- Procesamiento de datos: Es la etapa final del proceso y corresponde al procesamiento de la data con modelos analíticos. La elección de los modelos analíticos a utilizar depende del problema que se pretende resolver y del tipo de datos que se está utilizando. Lo ideal es procesar los datos con distintos modelos analíticos de manera de obtener el mejor resultado dentro de lo posible.

2.2.3 Minería de textos

La minería de textos es el proceso de extracción de patrones de interés desde una gran base de datos de texto, con el fin de descubrir información antes desconocida. La minería de texto aplica los mismos modelos analíticos que la minería de datos, pero utiliza además funciones analíticas del lenguaje natural y técnicas de recuperación de información. Debido a la naturaleza no estructurada de los documentos de textos, estos requieren un preprocesamiento especial antes de poder aplicar modelos analíticos.

Los principales usos son los siguientes:

- Extracción de información relevante desde una colección de documentos a través de algoritmos de asociación o ajuste de patrones.
- Búsqueda de tendencias o relaciones entre personas, lugares u organizaciones a través de comparaciones entre documentos agregados.
- Clasificar y organizar documentos de acuerdo con su contenido.
- Agrupar textos de acuerdo con su contenido.

- Obtener información respecto a la actitud de un ente de interés respecto de un tema particular, en función de la connotación del lenguaje utilizado (análisis de sentimientos).

La clasificación (en inglés *classification*) es un modelo de tipo supervisado en donde se busca organizar una serie de documentos en ciertos grupos definidos previamente, ya sea porque se cuenta con información de la constitución de aquellos grupos o porque se definen externamente los tópicos de cada grupo mediante conocimiento experto. El agrupamiento (en inglés *clustering*) es un modelo de tipo no supervisado en donde se busca agrupar una serie de documentos a través de patrones internos encontrados en el mismo set de documentos. El agrupamiento permite separar un set de documentos en distintos grupos a través de algoritmos de similitud. Así mismo, pueden utilizarse para obtener los distintos tópicos y tendencias que pueden haber dentro de un grupo de documentos, sin tener que leer forzosamente cada documento.

Los principales métodos utilizados en agrupamiento son los siguientes:

- K-Means: Método de agrupamiento basado en similitud entre vectores de palabras (distancia), siendo las más comunes la distancia euclidiana y la similitud coseno. El sistema escoge K centroides (Donde K es un parámetro del modelo), cada documento se asocia a su centroide más cercano, se calcula la suma total de las distancias de cada documento a su centroide y se itera respecto a la elección de centroides. El algoritmo termina cuando dicha suma de distancias es la mínima.
- K-Medoids: Método de agrupamiento similar a K-Means que utiliza la mediana en vez del promedio para el cálculo de los centroides y que siempre escoge alguno de los documentos como centroide en vez de un punto arbitrario en el espacio.
- Análisis de semántica latente (LSA): Modelo de agrupamiento basado en proximidad que utiliza la descomposición matricial (descomposición en valores singulares) de la matriz que contiene las palabras por cada documento. Los vectores se comparan finalmente utilizando similitud coseno. El efecto general es la disminución de la dimensionalidad de los

datos, mejorando los resultados ya que el exceso de dimensionalidad provoca ruidos en las aplicaciones basadas en similitud. La disminución de dimensionalidad también ayuda a magnificar los efectos semánticos en los datos subyacentes.

- Latent Dirichlet allocation (LDA, Blei, Ng y Jordan, 2003): Modelo probabilístico que asume que, en una colección de documentos, cada documento es una distribución respecto de los tópicos y por su parte cada tópico es una distribución respecto de las palabras contenidas en el set de documentos. La idea detrás del modelo es que los documentos pueden representarse como mezclas aleatorias entre tópicos escondidos (latentes) pertenecientes al mismo set de documentos. Permite simplificar los algoritmos de semántica latente reduciendo considerablemente el número de parámetros y proporcionando una probabilidad bien definida para documentos arbitrarios. La figura 7 presenta un esquema gráfico del modelo LDA.

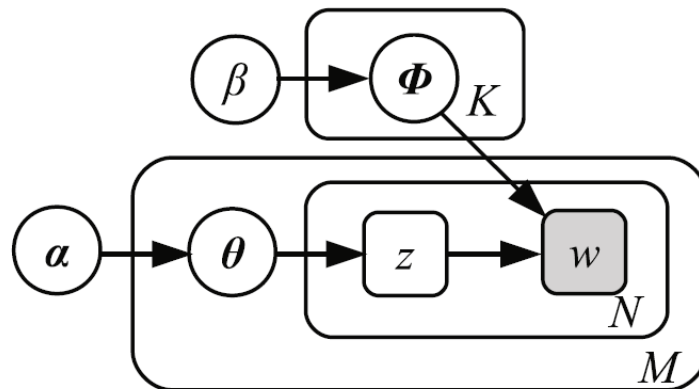


Figura 7. Esquema gráfico del modelo LDA.

En la figura 7, la distribución Φ de los términos en cada tópico es una distribución de Dirichlet de parámetro β , la distribución θ de los tópicos en cada documento es una distribución de Dirichlet de parámetro α , z es el tópico para cada palabra en cada documento y se escoge desde la distribución multinomial de tópicos del documento (multinomial de parámetro θ) y w es la palabra específica y se escoge desde la distribución multinomial asociada con la selección de tópicos (multinomial de parámetro Φ).

- Correlated topic model (CTM, Blei y Lafferty, 2007): Extensión del modelo LDA en donde se permiten correlaciones entre los tópicos. El modelo LDA no permite correlaciones entre tópicos debido al supuesto de independencia implícito en la distribución de Dirichlet utilizada. El modelo CTM reemplaza la distribución de Dirichlet por una distribución logística normal que es más flexible, incorporando una estructura de covarianza entre los componentes. Esto permite que la presencia de un tópico pueda ser correlacionado con la existencia de otro.

Varios autores han utilizado estas técnicas para agrupar documentos de algún tema en particular o encontrar tendencias dentro de ellos. Dentro de la investigación reciente, se han estudiado las tendencias en investigación académica respecto del tema de la privacidad de la información personal (Chen, Zhang, Zhu y Lu, 2017), tendencias de las presentaciones en conferencias (Moro, Alturas, Esmeraldo y Costa, 2017; Mathew, Agrawal y Menzies, 2017) y tendencias y tópicos en patentes (Choi, Lee y Sohn, 2017; Ma, Luo, Xuan, Xue y Guo, 2017), entre otros temas. Otros investigadores han aplicado modificaciones a las técnicas de manera de mejorar los resultados obtenidos. Entre las modificaciones puede mencionarse la utilización de conocimiento previo para mejorar la selección de tópicos (Shams y Baraani-Dastjerdi, 2017; Yao et al., 2017), la utilización de sistemas semi-supervisados para mejorar la interpretación de los tópicos (Ramage, Manning y Dumais, 2011) y la combinación con redes de colaboración de documentos (Wang y Blei, 2011; Gao, Li, Yu, Qin y Zhang, 2017).

2.2.4 Procesos para la implementación de minería de datos

Debido a la consolidación del campo de la minería de datos, se han desarrollado varios estándares cuyo fin es facilitar la implementación de sus aplicaciones. Los más populares son KDD (Knowledge Discovery in Databases, Fayyad, Piatetsky-Shapiro y Smyth, 1996), CRISP-DM (Cross Industry Standard Process for Data Mining, Shearer, 2000) y SEMMA (SAS Institute). De los tres, CRISP-DM es el más completo a la hora de implementar una solución de minería de datos dentro de un negocio debido a lo siguiente (Azevedo y Santos, 2008; Shafique y Qaiser, 2014):

- CRISP-DM extiende a KDD incorporando además dos etapas extras: Una etapa previa concerniente al entendimiento del negocio, que dice relación con el entendimiento del dominio donde se utilizará la aplicación, el conocimiento previo pertinente y los objetivos del usuario final; y una etapa posterior de implementación, que dice relación con la consolidación e incorporación del conocimiento al sistema. La figura 8 muestra un esquema del proceso CRISP-DM.
- SEMMA puede ser visto como una implementación práctica de KDD y es muy similar a él. No incluye explícitamente los procesos de entendimiento del negocio y despliegue y se centra principalmente en el modelamiento de los datos.

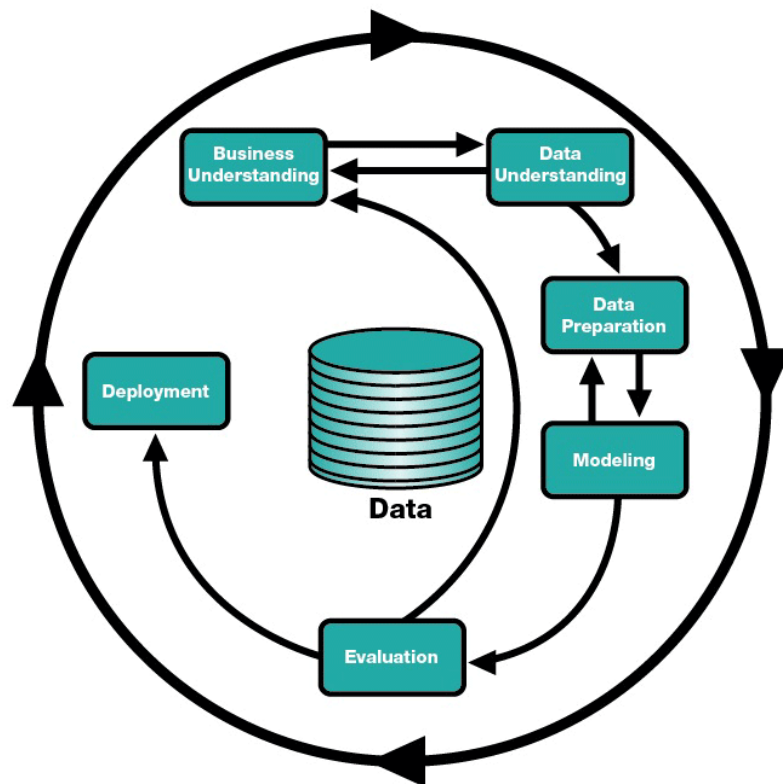


Figura 8. Esquema de la metodología CRISP-DM.

CAPÍTULO 3: PLANTEAMIENTO ESTRATÉGICO Y MODELO DE NEGOCIOS

3.1 Posicionamiento estratégico

Para el análisis del posicionamiento estratégico de la empresa se utilizará el modelo Delta de Hax (Hax y Wilde, 2001). Al respecto la empresa se inclinaría a mejor producto según se muestra en la figura 9.

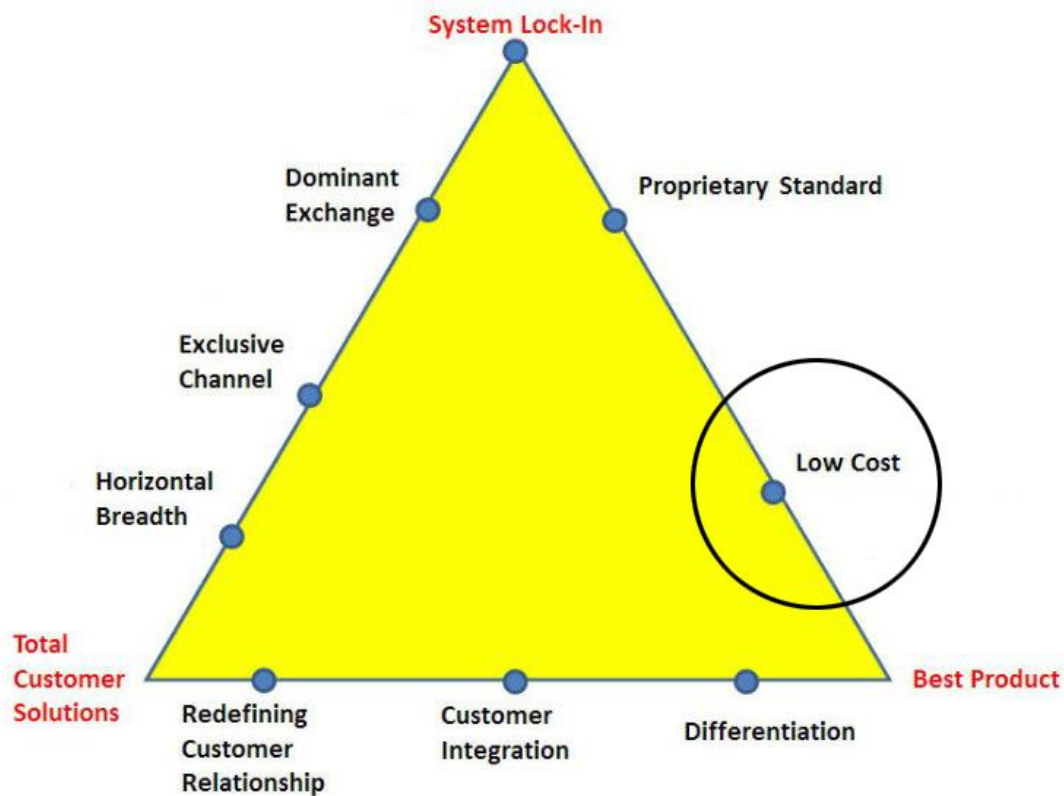


Figura 9. Posicionamiento estratégico de Moly met según el modelo delta de Hax (Hax y Wilde, 2001).

El enfoque estratégico es directo ya que se comercializa un producto relativamente estandarizado (*commodity*) y los esfuerzos de la empresa siempre han apuntado a disminuir los costos de producción. Lo anterior se expresa también en la misión de la empresa, en donde se establece el enfoque en la capacidad de procesamiento y optimización e innovación de procesos, puntos clave a la hora de disminuir los costos de producción. Por otro lado,

con el nivel de demanda que se supe hoy en día sería imposible alcanzar las cotas de lock-in sistémico.

3.2 Lineamientos estratégicos

La misión de MolyMet establece lo siguiente:

“MolyMet fortalecerá su posición líder a nivel mundial en el mercado del molibdeno y renio, mediante una capacidad de procesamiento eficiente y flexible, la continua optimización de procesos y la innovación, además de la búsqueda de nuevas oportunidades de negocio en metales estratégicos.

En especial, procuraremos aumentar el valor para nuestros grupos de interés más relevantes, entre los que cuentan:

- *Nuestro equipo humano, pilar fundamental en todo lo que hacemos y de quienes nos preocuparemos incansablemente.*
- *Nuestros clientes, socios estratégicos de largo plazo, a quienes deseamos servir de manera cercana para satisfacer sus necesidades.*
- *La comunidad en la cual estamos insertos y de la cual debemos ser parte activa y comprometida.*
- *Nuestros proveedores de insumos y servicios, quienes nos apoyan con la adecuada operación de las plantas industriales y unidades corporativas.*
- *Las autoridades, cuyas regulaciones nos proponemos superar permanente y estrictamente, generando así un ambiente de confianza.*

En este contexto, MolyMet espera obtener resultados superiores y crear valor económico sustentable para sus accionistas e inversionistas.

Incentivaremos una cultura empresarial que actúe consistentemente de manera responsable, ética, moral y transparente.”

Por otro lado, la empresa define su estrategia en función de cuatro ejes estratégicos principales:

- Industria del molibdeno y renio: Mantener la dominancia en el procesamiento de molibdeno y renio.
- Industria del molibdeno metálico: Fortalecer y rentabilizar la integración vertical de la empresa hacia la producción de molibdeno metálico en las empresas Chemiometall y Luoyang High-Tech Metals.
- Industria de metales afines: Fomentar el crecimiento de la empresa a partir de la creación de nuevos negocios dentro de la industria de metales afines. Estos metales son todos aquellos pertenecientes al grupo de los refractarios: Molibdeno, renio, tungsteno, niobio y tantalio.
- Sustentabilidad de la empresa: Asegurar la permanencia de MolyMet y en especial la permanencia de la casa matriz en la localidad de Nos.

La relación con el posicionamiento estratégico se evidencia en los 4 ejes estratégicos:

- Industria del molibdeno y renio: Se pretende aumentar la productividad y reducir los costos, por lo que claramente se apunta a una estrategia de mejor producto.
- Industria del molibdeno metálico: Se pretende rentabilizar el negocio del molibdeno metálico, aumentando la producción de las plantas y mejorando o creando nuevos productos. También se apunta a la estrategia de mejor producto.
- Industria de metales afines: Se pretende aumentar el crecimiento de la empresa a través de nuevos negocios en mercados de metales afines. Debido a que los productos seguirán siendo *commodities*, se mantiene el posicionamiento estratégico del mejor producto.

- **Sustentabilidad:** Apunta a asegurar la permanencia del negocio y por lo tanto es sustancial para cualquier posicionamiento estratégico.

3.3 Modelo de negocios

Para el modelo de negocio se utilizará el modelo Canvas de Osterwalder (Osterwalder, 2004). El resultado se muestra en la figura 10.










ACTIVIDADES CLAVE	ASOCIACIONES CLAVE	PROPUESTAS DE VALOR	RELACIONES CON CLIENTES	SEGMENTOS DE MERCADO		
 Negociación con proveedores y clientes. Optimización continua de la operación con énfasis en la disminución de costos manteniendo la calidad de los productos. Integración de nuevas tecnologías de proceso.	 Empresas de la minería del cobre. Empresas de transporte.	 Procesamiento de concentrados de molibdenita para producir diversos productos de molibdeno y subproductos como renio, cobre y ácido sulfúrico.	 Comunicación directa a través de la Vicepresidencia comercial y las oficinas comerciales a lo largo del mundo.	 Empresas siderúrgicas y hierro fundido. Empresas de catalizadores para industria del petróleo. Empresas de aceros especiales de alta velocidad y superaleaciones. Minería del cobre (maquila).		
	RECURSOS CLAVE		CANALES			
	 Plantas productivas. Mano de obra especializada. Oficinas comerciales.		 Venta directa a los clientes y despacho por aire, mar o tierra.			
 ESTRUCTURA DE COSTES		FUENTES DE INGRESOS 				
Costos de materia prima (concentrado de molibdenita). Costos operacionales. Costos de transporte.		Venta directa de productos y subproductos. Cobro de fee por procesamiento de concentrado de molibdenita externo (maquila).				

Figura 10. Modelo de negocios de Molymet según el modelo de Osterwalder.

CAPÍTULO 4: ANÁLISIS SITUACIÓN ACTUAL

4.1 Arquitectura de procesos

La figura 11 presenta la visión general de macroprocesos de Molymet. Estos macroprocesos representan todas las actividades de la empresa en forma de agrupaciones grandes de procesos.

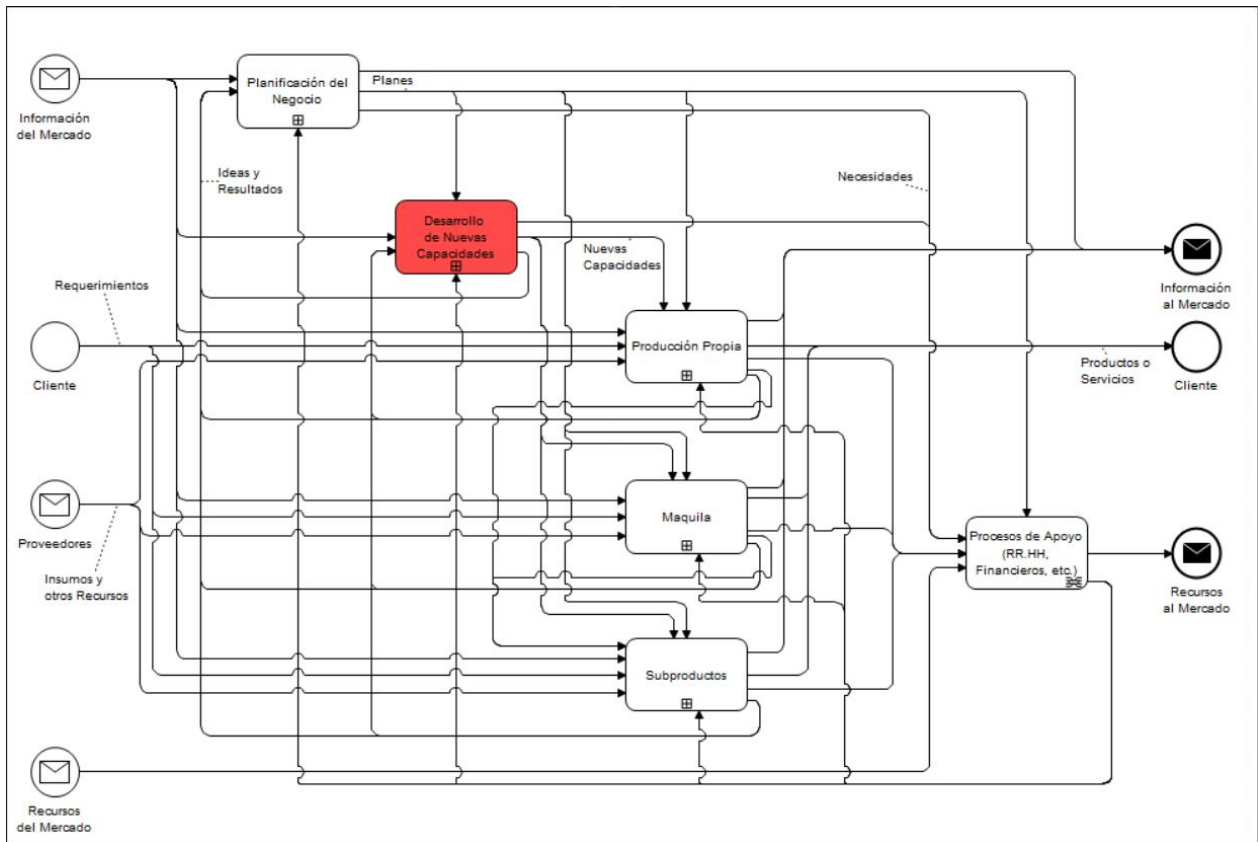


Figura 11. Visión general de macroprocesos de Molymet. Se indica en rojo el macroproceso en el que se encuentra el proyecto MBE.

En primer lugar, está la planificación del negocio, en donde se define el curso a seguir de la empresa en forma de estrategias, que finalmente se convierten en planes que gobiernan a los demás procesos. En estos procesos se encuentra la Vicepresidencia de Información y el Presidente Ejecutivo.

En segundo lugar, está el desarrollo de nuevas capacidades, que agrupa una serie de procesos que se encargan de proponer mejoras a los procesos

actuales de la empresa (para lo que eventualmente pueden monitorear su desempeño actual) e innovación, permitiendo a la empresa la creación de capacidades que antes no existían. En estos procesos se encuentra la VPIT y la Vicepresidencia de Ingeniería. Además, aquí se ubican algunas de las actividades clave del modelo de negocio.

La generación de un nuevo negocio, entendiéndolo como la generación de una capacidad nueva para la empresa o innovación, dice directa relación con este macroproceso por lo que el proyecto desarrollado se ubicará aquí. El detalle del proceso se incluirá en la sección siguiente.

En tercer lugar, está la cadena de valor, que para el caso de Molymet son 3: producción propia, maquilas y subproductos, es decir, aquí está el desarrollo principal de la propuesta de valor. En estos procesos se encuentran la Vicepresidencia de Productos de Molibdeno Metálico, Operaciones y Comercial.

Finalmente están los procesos de apoyo, que son el conjunto de procesos que tienen que ver con el manejo de los recursos necesarios para el desarrollo de los demás procesos, como lo son RR. HH., finanzas, etc. En estos procesos se encuentran las Vicepresidencias de Finanzas, Gobierno Corporativo e Información.

4.2 Modelamiento detallado de procesos

La figura 12 muestra el diagrama correspondiente al macroproceso de desarrollo de nuevas capacidades. Este macroproceso abarca todas las tareas relacionadas con la evaluación de la necesidad de una nueva capacidad, la gestión del diseño y construcción de la capacidad y finalmente el diseño y construcción de la nueva capacidad.

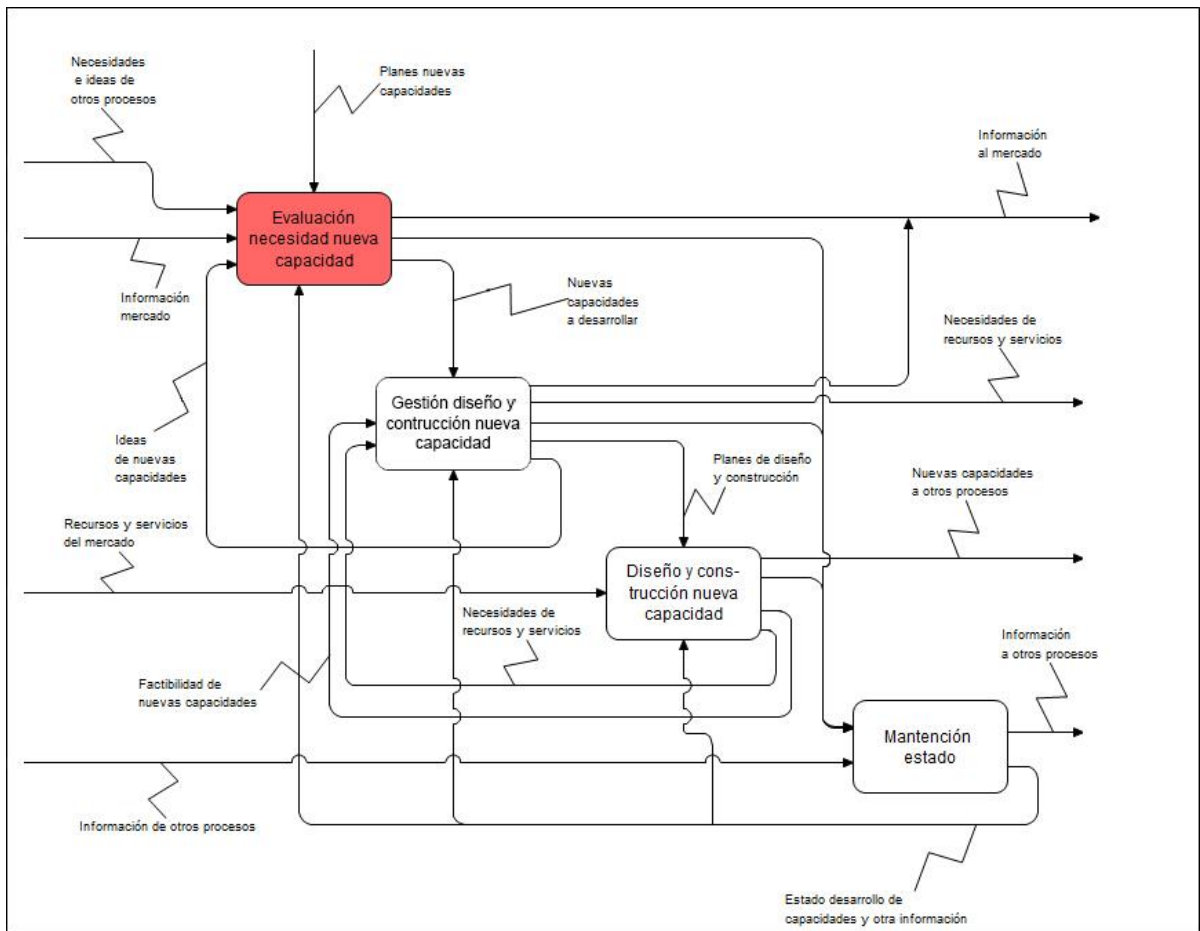


Figura 12. Macroproceso de desarrollo de nuevas capacidades. Se incluye en rojo el proceso relacionado con el proyecto MBE.

La evaluación de la necesidad de una nueva capacidad consiste en el análisis de la información que motiva y valida el proyecto y la evaluación económica formal para decidir su implementación. Lo anterior dice relación directa con la generación de proyectos conducentes a nuevos negocios por lo que el presente proyecto se ubicará en este proceso. Se enfatiza la necesidad de contar con información de mercado para poder evaluar de manera eficiente la necesidad de una nueva capacidad.

Gestión del diseño y construcción de la nueva capacidad ejecuta los subprocesos y actividades necesarias para determinar los recursos necesarios para el proyecto, obtiene y asigna tales recursos y genera un plan para el proyecto. Nuevamente se enfatiza la necesidad de contar con información estratégica para establecer los mecanismos adecuados de gestión para los proyectos.

Finalmente, Diseño y construcción de la nueva capacidad ejecuta las asignaciones y los planes originados en los procesos de gestión, llevando a cabo las actividades necesarias.

La figura 13 muestra el proceso de evaluación de la necesidad de una nueva capacidad. Este comprende los subprocesos de análisis de mejora de capacidades e innovación de capacidades.

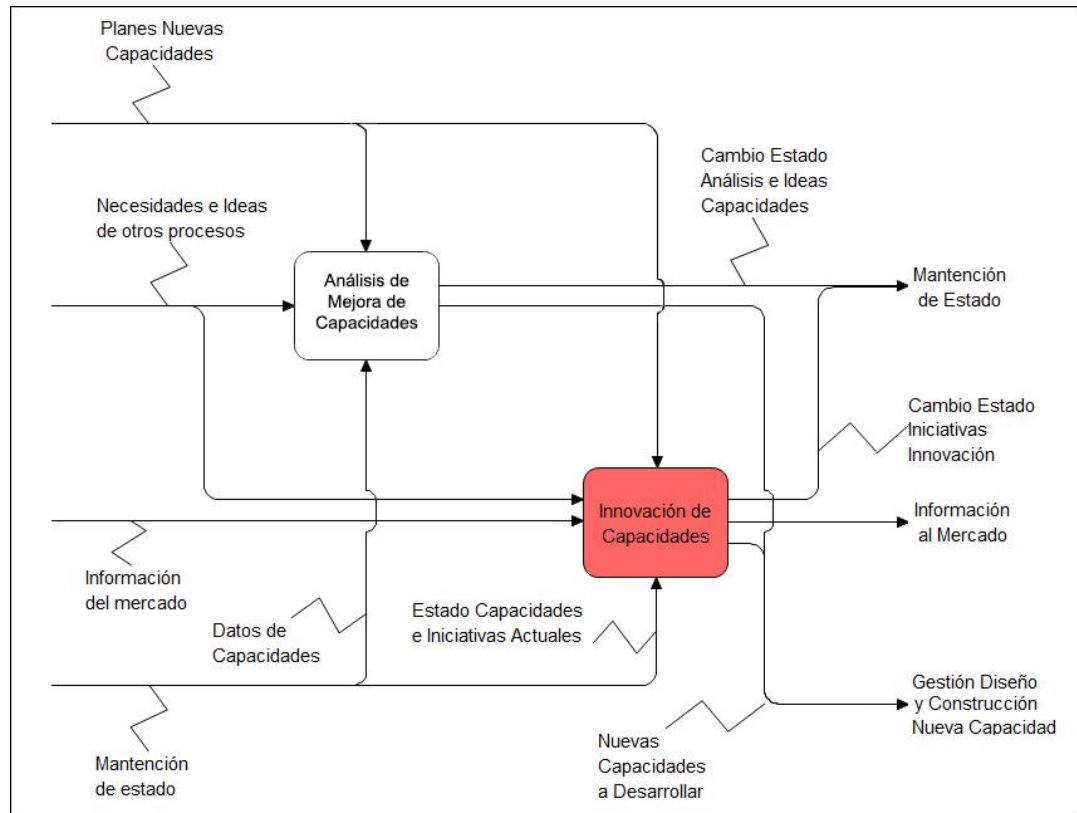


Figura 13. Proceso de evaluación de la necesidad de una nueva capacidad. Se incluye en rojo el subproceso objetivo del proyecto MBE.

El proceso de análisis de mejora de capacidades se centra en sacarle partido a toda la información operativa que se genera en los macroprocesos de cadena de valor y procesos de apoyo y a las sugerencias que provienen de éstos, con el fin de proponer mejoras a los procesos actuales. El objetivo principal es la mejora continua.

La innovación de capacidades tiene que ver con todos aquellos procesos formales, que pretenden generar en forma sistemática nuevas capacidades en

todos los ámbitos: nuevos productos y servicios, nueva tecnología en los procesos productivos, etc., es decir, son los procesos que apuntan a generar capacidades que antes no existían en la empresa. Notar que se enfatiza la necesidad de contar con información de mercado para llevar a cabo el proceso.

Haciendo un paralelo con la situación actual de la VPIT:

- Mejora de capacidades: Propuestas provenientes de Vicepresidencia de Operaciones, Presidencia Ejecutiva y Vicepresidencia Comercial.
- Innovación de capacidades: Propuestas provenientes de Presidencia Ejecutiva, Vicepresidencia Comercial y Vigilancia tecnológica.

La figura 14 muestra el BPMN del proceso de innovación de capacidades. El proceso de vigilancia tecnológica es un proceso manual de análisis de un pequeño subconjunto de la información disponible en función de alguna recomendación particular o de los intereses del analista. En este proceso se enfoca el presente proyecto.

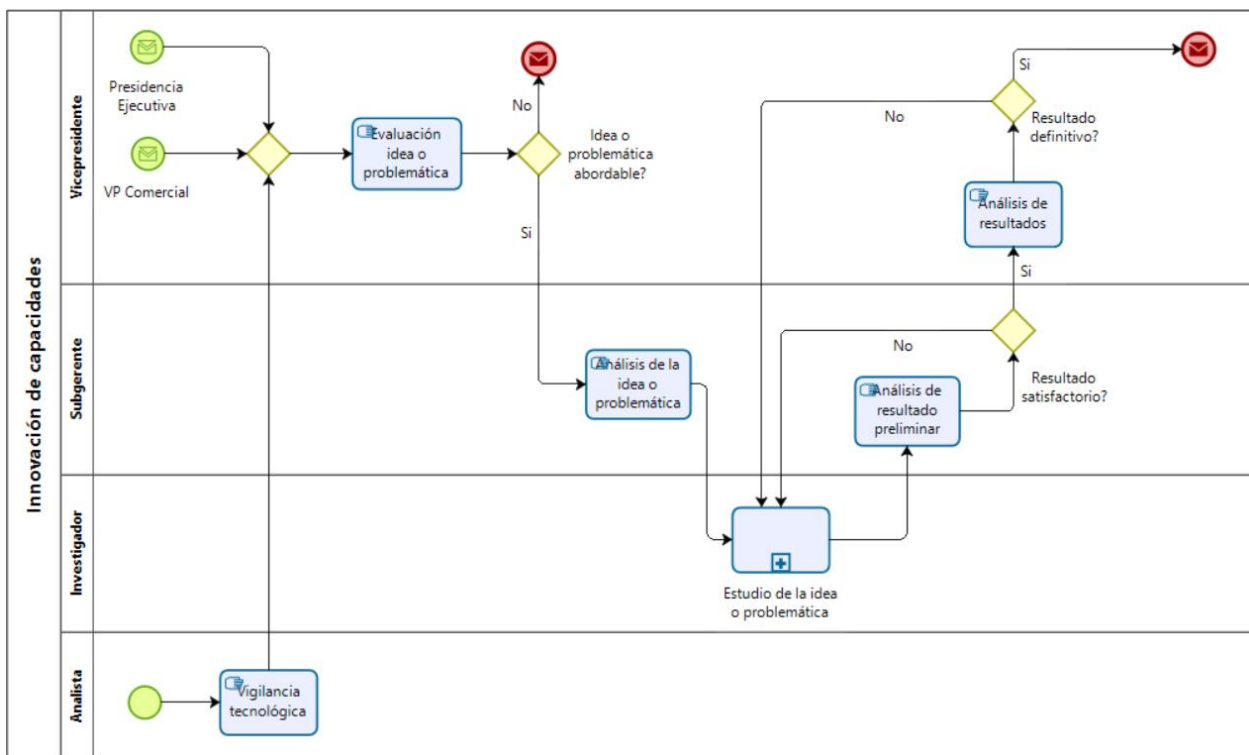


Figura 14. BPMN del proceso de innovación de capacidades as is.

4.3 Diagnóstico de la situación actual

Las empresas que comercializan *commodities* ven limitada su capacidad de crecimiento al crecimiento mundial de la demanda del producto, puesto que se trata de materiales demasiado estandarizados como para generar diferenciación. La innovación es una fuente importante de crecimiento en los negocios, generando diferenciación en los productos y/o creando nuevos negocios para la empresa. Para el caso de Molymet en concreto, el crecimiento está limitado por el crecimiento de la demanda mundial de molibdeno, que suele rondar el 2% anual en promedio en el largo plazo. Para fomentar el crecimiento, se ha establecido el uso estratégico de la innovación como fuente generadora de nuevos negocios en metales afines, entendiendo a estos como aquellos pertenecientes a los del grupo de los refractarios: molibdeno, renio, tungsteno, niobio y tantalio. Estos metales poseen un comportamiento químico similar, por lo que podría aprovecharse el know-how actual de Molymet en cuanto al procesamiento de molibdeno y renio y extenderlo hacia los demás elementos.

De las fuentes de ideas para innovación de capacidades, la única sustentable es la vigilancia tecnológica porque permite generar nuevas ideas a partir del análisis del entorno competitivo sin depender de consultas específicas o eventualidades. Sin embargo, la vigilancia tecnológica es hoy en día en Molymet más un concepto que una herramienta útil dentro de la generación de ideas para proyectos, puesto que actualmente se limita a ser un proceso de análisis de un subconjunto muy pequeño de toda la información relevante disponible; subconjunto que se escoge por alguna recomendación particular o por los intereses del analista, por lo que está lejos de representar realmente toda la información contenida. Dada la gran cantidad de información disponible, tanto en publicaciones científicas como en patentes de invención, es imposible realizar un análisis manual a tiempos y costos razonables. De aquí la necesidad de establecer un proceso inteligente y bien definido de vigilancia tecnológica que incluya un apoyo tecnológico, para poder analizar los grandes volúmenes de información de manera global.

Por otro lado, no basta con generar más proyectos conducentes a nuevos negocios, pues el verdadero valor se encuentra en los proyectos exitosos. A pesar de que el éxito de un proyecto de investigación nunca está asegurado, hay que disminuir la incerteza al mínimo valor posible. Para ello se debe contar

con las mejores fuentes de información y seleccionar adecuadamente que contenido va a utilizarse finalmente para la generación de una idea conducente a un nuevo proyecto.

4.4 Cuantificación del problema u oportunidad

Existen dos problemas fundamentales derivados de una selección de proyectos ineficiente:

- **Proyectos conducentes a nuevos negocios:** Actualmente se encuentran en torno a un 10-15%. Como objetivo estratégico de la compañía se espera aumentar este porcentaje al menos a un 30%.
- **Proyectos exitosos:** Actualmente se encuentran en torno a un 10-12%. El porcentaje de éxito típico en centros de investigación similares (de industria minera) debería estar en torno a 20% (según juicio experto y benchmarking con Soquimich).

Un proceso inteligente de vigilancia tecnológica, que incluye apoyo informático, permite apuntar a ambos objetivos a la vez. Por una parte, permite conocer la información externa de manera global y las tendencias contenidas en los grandes volúmenes de documentos. Esta información externa es la que permite generar nuevos negocios para la compañía, ya sea anteponiéndose a los movimientos del mercado con negocios riesgosos pero con una mayor utilidad potencial (desde publicaciones científicas que representan los movimientos futuros del mercado) o imitando las tendencias actuales del mercado con negocios con funcionalidad probada, por lo que conllevan menor riesgo, pero con menor utilidad potencial pues se estaría replicando lo que ya se hace (desde patentes de invención que representan los movimientos actuales y pasados del mercado). Por otra parte, la información obtenida permite disminuir la incerteza a la hora de decidir respecto de que proyecto se debería realizar, por lo que se espera un aumento en el índice de proyectos exitosos. La figura 15 muestra un resumen de los beneficios esperados de la implementación de un proceso bien definido de vigilancia tecnológica en la compañía.

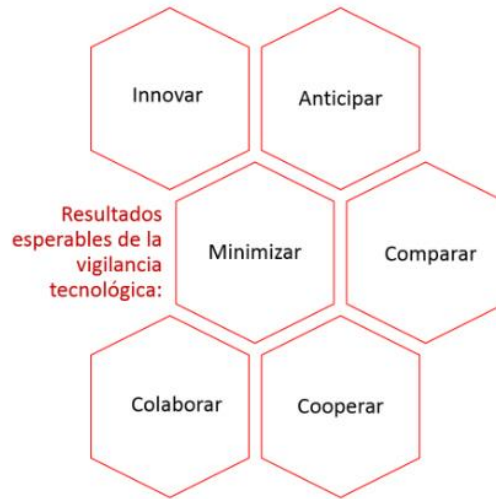


Figura 15. Resultados esperables de la implementación de vigilancia tecnológica en la empresa.¹

Al respecto de la figura 15 se puede comentar que un proceso de vigilancia tecnológica bien implementado permitiría:

- Anticipar los cambios del entorno al detectar tempranamente las tendencias e información de carácter estratégico.
- Minimizar los riesgos asociados a la innovación a partir de una detección de posibles amenazas y cambios relevantes en el entorno.
- Comparar el rendimiento actual del negocio con lo que ya existe en el entorno competitivo.
- Innovar al detectar oportunidades de mejora y fuentes de ideas generadoras de proyectos de investigación.
- Colaborar y cooperar con socios estratégicos a partir de la detección de oportunidades de colaboración y cooperación.

En particular para la VPIT, se espera aplicar vigilancia tecnológica para seleccionar un mayor número de proyectos conducentes a nuevos negocios y al mismo tiempo con mayor probabilidad de éxito, pues se está tomando la

¹ Fuente: MoocVT

decisión de acerca de que proyecto realizar contando con la mejor información disponible.

El gasto anual promedio de la VPIT para el periodo 2010-2017 fue de 1.5 MMUSD/año. El índice de proyectos exitosos actuales de la Vicepresidencia debe permitir como mínimo justificar su existencia (al menos en el largo plazo) puesto que de otro modo ya se hubiera considerado su eliminación. Dicho de otro modo, un índice de proyectos exitosos de un 11% en promedio permitiría generar un ingreso de 1.5 MMUSD/año para la compañía. Si se asume una proporcionalidad directa de los ingresos respecto del porcentaje de éxito de los proyectos y una mejora de un 3% en este mismo índice debido a la aplicación de la mejora propuesta (valor relativamente bajo y conservador pues se estaría aumentando el índice a 14% aún muy por debajo del 20%), se tendría un aumento en los ingresos de 0.4 MMUSD/año como mínimo.

CAPÍTULO 5: PROPUESTA DE DISEÑO DE PROCESOS

5.1 Direcciones de cambio y alcance

A continuación, se detallan las variables de diseño asociadas al proyecto con el fin de contrastar la situación actual con la situación propuesta.

Variable	Actual	Propuesto
Servicio integral al cliente	No	No
Lock-in sistémico	No	No
Integración con proveedores	No	No
Estructura interna centralizada o descentralizada	Descentralizada	Descentralizada
Toma de decisiones centralizada o descentralizada	Descentralizada	Descentralizada

Tabla 4. Variables de diseño de estructura de empresa y mercado.

Variable	Actual	Propuesto
Modelo de agrupamiento de publicaciones de interés	No	Permite mejorar la selección de proyectos con base en la información del comportamiento del mercado
Modelo de recomendación de publicaciones de interés	No	Permite contar con la mejor información para seleccionar proyectos con menos riesgo

Tabla 5. Variables de diseño de anticipación.

Variable	Actual	Propuesto
Reglas	No	Reglas formales con apoyo computacional para optimizar la selección de proyectos
Jerarquía	No	No
Colaboración	No	No
Partición	No	No

Tabla 6. Variables de diseño de coordinación.

Variable	Actual	Propuesto
Lógica de negocio automatizada o semiautomatizada	No	Lógica de negocio semiautomatizada
Lógica de apoyo a actividades tácitas	No	No
Procedimientos de comunicación e integración	No	Definición de procesos y manuales de uso
Lógica y procedimientos de desempeño y control	No	Se agrega criterio de éxito de proyectos y de proyectos conducentes a nuevos negocios

Tabla 7. Variables de diseño de prácticas de trabajo.

Variable	Actual	Propuesto
Proceso aislado	Si	Si
Todos o la mayor parte de los procesos de un macroproceso	No	No
Dos o más macros que interactúan	No	No

Tabla 8. Variables de diseño de integración de procesos conexos.

Variable	Actual	Propuesto
Datos propios	No	No
Integración con otros sistemas de la empresa	No	No
Integración con datos de sistema de otras empresas	No	No

Tabla 9. Variables de diseño de mantención de estado.

La mejora propuesta no pretende cambiar la estructura de la empresa y se centra en un proceso particular dentro de un área concreta. Los cambios se centran en las variables de anticipación y en particular en la mejora de la selección de proyectos con información de los movimientos futuros probables del mercado, y al mismo tiempo permite anteponerse a futuros proyectos fallidos al contar con la mejor información, que permite tomar decisiones

respaldadas y justificadas y por lo tanto con menor riesgo. Se añaden además reglas computacionales de apoyo a la toma de decisiones y la lógica de negocio semiautomatizada de vigilancia tecnológica inteligente.

5.2 Diseño detallado de procesos to be

El diseño de procesos para la situación con proyecto se centra en el proceso de evaluación de una nueva capacidad. El detalle del rediseño se muestra en el diagrama de procesos de la figura 16 y en el BPMN de la figura 17.

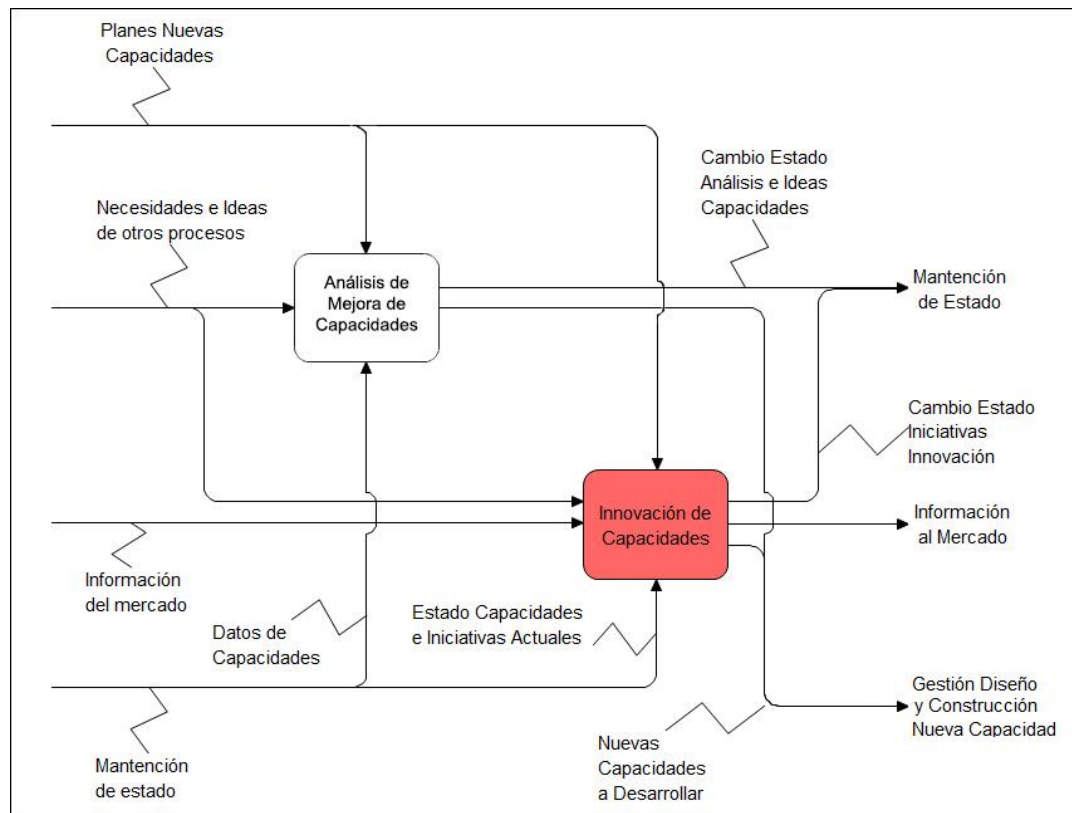


Figura 16. Proceso de evaluación de la necesidad de una nueva capacidad to be. Se incluye en rojo el subproceso a rediseñar en el proyecto MBE.

En esta nueva situación el proceso de vigilancia tecnológica se ha rediseñado, tal como se muestra en la figura 17. El proceso comienza con el ingreso de algún criterio de búsqueda a la herramienta de vigilancia tecnológica por parte del analista. La aplicación consulta la base de datos externa, procesa los datos obtenidos, y permite visualizar directamente los resultados. Si el analista

considera que los resultados obtenidos son coherentes puede generar informes de resultados que incluyen el agrupamiento de los documentos y una recomendación de los mejores documentos por tópico, en caso contrario se volverá a solicitar el procesamiento de datos (cambiando por ejemplo el número de *clusters*). Una vez generados los informes, se procederá a analizar la información para obtener alguna idea interesante conducente a la realización de un proyecto. Si la información obtenida no permite generar alguna idea interesante, entonces se volverá al comienzo del proceso para ingresar un nuevo criterio de búsqueda.

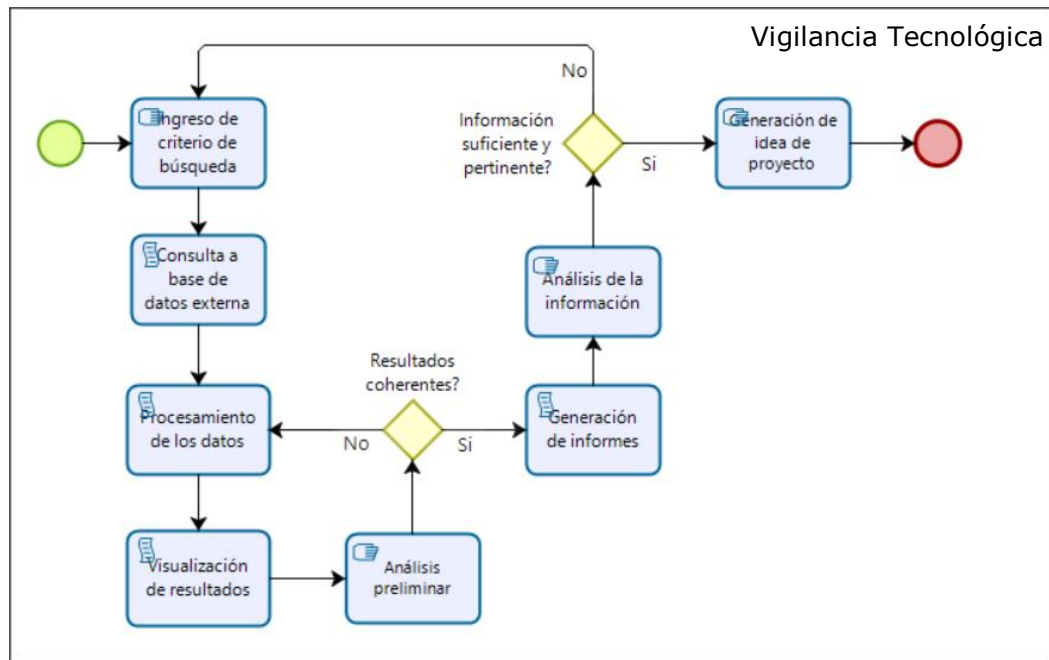
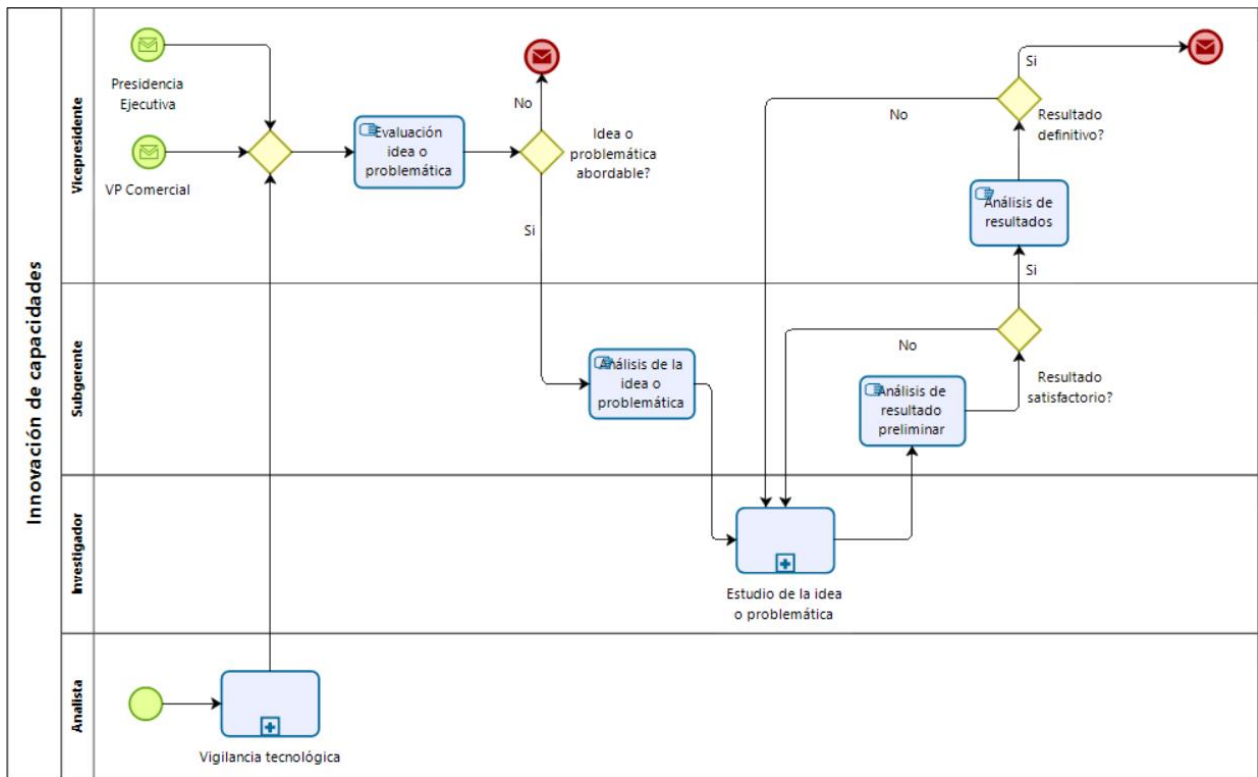


Figura 17. BPMN del proceso de innovación de capacidades to be.

5.3 Diseño de lógica de negocios

5.3.1 *Diseño del piloto*

El piloto se diseñará considerando los siguientes parámetros:

- Información: Base de datos de publicaciones científicas Scopus. Se trata de una base de datos muy completa y con información atinente al negocio de la empresa.
- Ecuación de búsqueda: "molybdenum". Básicamente porque se trata del producto principal de la empresa y cuya información es la más atinente al negocio actual.
- Año de búsqueda: 2016. El proceso está pensado para ejecutarse una vez al año con toda la información publicada el año inmediatamente anterior (sin perjuicio de que se podrá analizar otros años anteriores si ese es el objetivo), de manera de obtener varias ideas que puedan convertirse en los proyectos que se desarrollarán durante el año.
- Información por exportar: Informaciones de títulos, resúmenes y citas. Se utilizará la información de los títulos para la realización del *clustering* de los documentos y la información de citas para el sistema de recomendación de documentos.
- Resultados por obtener: Informes que incluirán los diferentes *clusters* y los documentos dentro de cada *cluster* ordenados por citas. Se incluirá el título y el resumen de cada documento.

Las etapas del proceso son las siguientes:

- Selección de los datos: Se hará a través de un *crawler* que se conecta a Scopus a través de su API. Se obtuvieron 5665 documentos.
- Preprocesamiento de datos: Se incluye remoción de *missing values*, conversión a minúscula, remoción de signos de puntuación, números, espacios en blanco, signos varios, *stopwords*, eliminación de la ecuación de búsqueda y aplicación de stemming. Finalmente, se crea la matriz

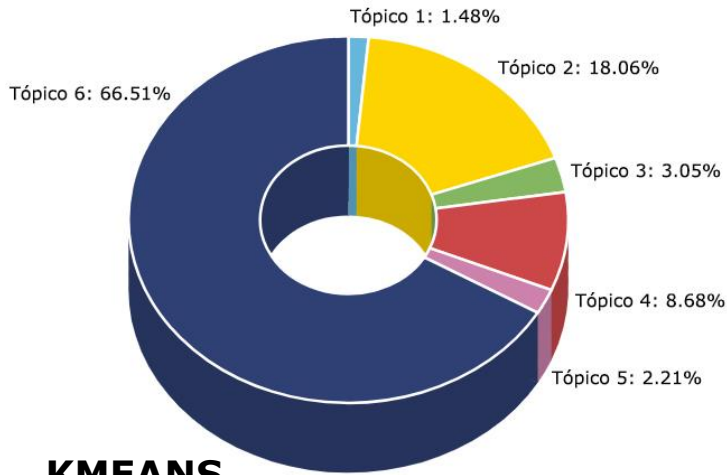
término-documento, que es la matriz base para la aplicación de los modelos de *clustering*.

- Modelamiento: Se aplicará un modelo de *clustering* y se seleccionará adecuadamente el número de tópicos coherente.
- Interpretación y evaluación: Se determinarán los tópicos de manera explícita y se exportarán los documentos por cada tópico ordenados por citas.

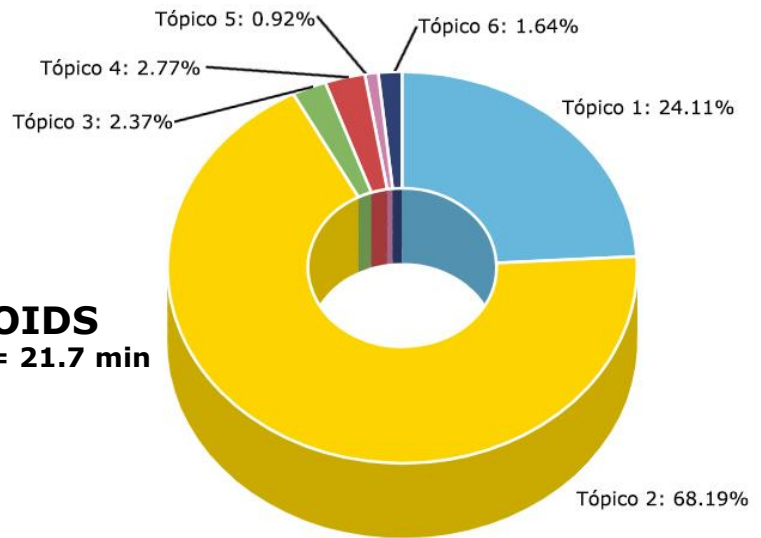
5.3.2 Selección del modelo de clustering

Se utilizaron 5 modelos de *clustering* diferentes: Kmeans con distancia euclidiana, Skmeans (Kmeans con similitud coseno), Kmedoids, LDA y CTM. El resultado comparativo se muestra en la figura 18. Se realizó un análisis preliminar del *cluster* de documentos utilizando LDA, la herramienta de visualización de tópicos de la sección 5.3.3 y juicio experto y se obtuvieron 6 tópicos como lo óptimo. Por lo tanto, se utilizó este mismo número para el análisis comparativo.

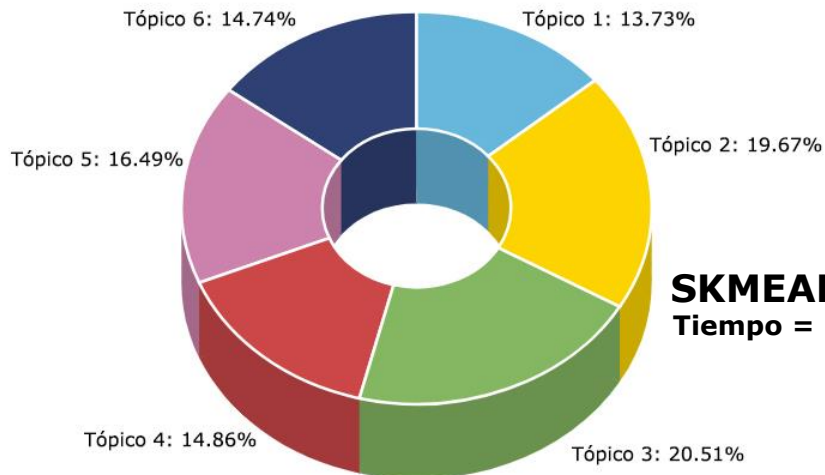
Los resultados de la figura 18 muestran que Kmeans y Kmedoids tienden a agrupar la mayor cantidad de documentos en un solo gran *cluster* para dividir el resto en porciones muy pequeñas. Debido a la naturaleza de la información que se está analizando, no es coherente que sobre el 50% de los documentos pertenezcan a un solo tópico, por lo que se infiere que el *clustering* no fue efectivo. Skmeans, CTM y LDA presentan *clusters* aceptables para este tipo de información. Sin embargo, CTM presenta un tiempo de procesamiento mucho más alto que LDA y Skmeans y además los tópicos tanto de Skmeans como de CTM no son fáciles de interpretar y son difícilmente asociables a un tópico real, por lo que finalmente se escogió el LDA como el modelo idóneo para el problema en cuestión.



KMEANS
Tiempo = 10.6 seg



KMEDOIDS
Tiempo = 21.7 min



SKMEANS
Tiempo = 6.1 seg

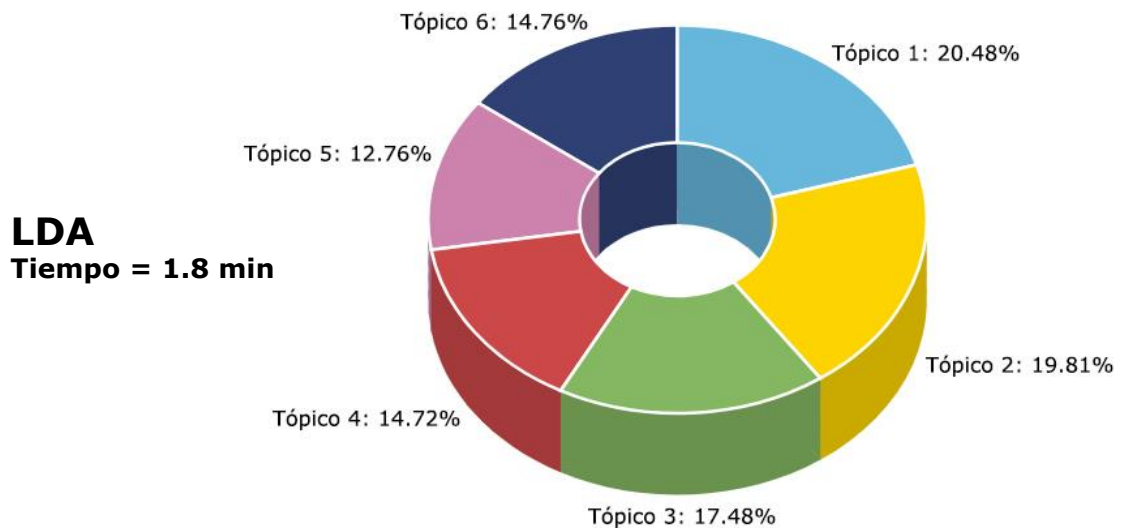
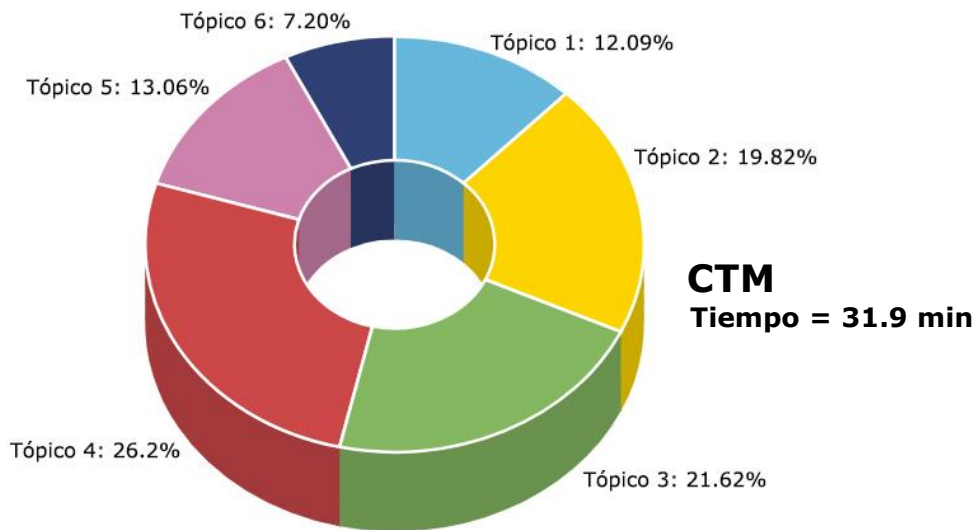


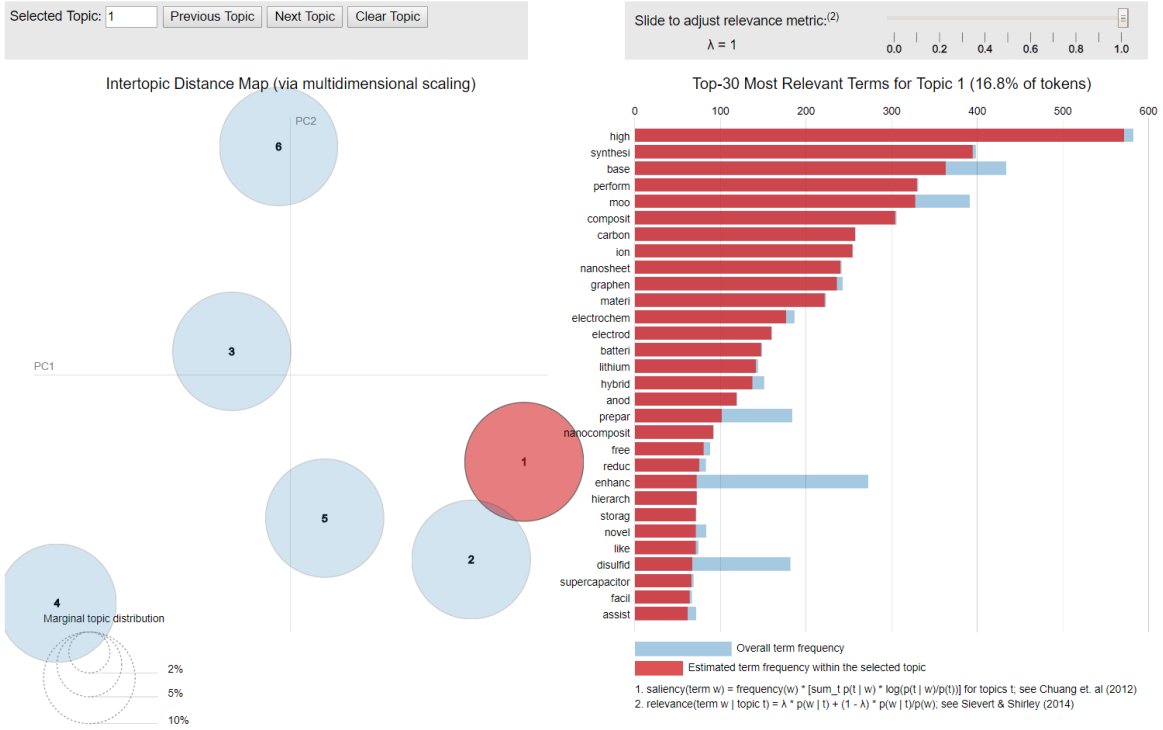
Figura 18. Resultados comparativos entre los distintos modelos de *clustering* incluyendo el tiempo de procesamiento.

5.3.3 Selección del número de tópicos

La selección del número óptimo de tópicos no es una actividad automática y depende necesariamente de la intervención del analista. Se utilizó una herramienta que facilita la selección del número de tópicos al visualizar la

distancia intertópica. La idea es seleccionar el mayor número de tópicos posible evitando que se traslapen, pues esto significaría en la práctica que dos o más tópicos están representando a un mismo grupo. Después de obtenido este valor se debe realizar el análisis de interpretación de tópicos y comprobar la coherencia obtenida. La figura 19 muestra una imagen comparativa de los resultados obtenidos con 6 y 7 grupos.

K = 6



K = 7

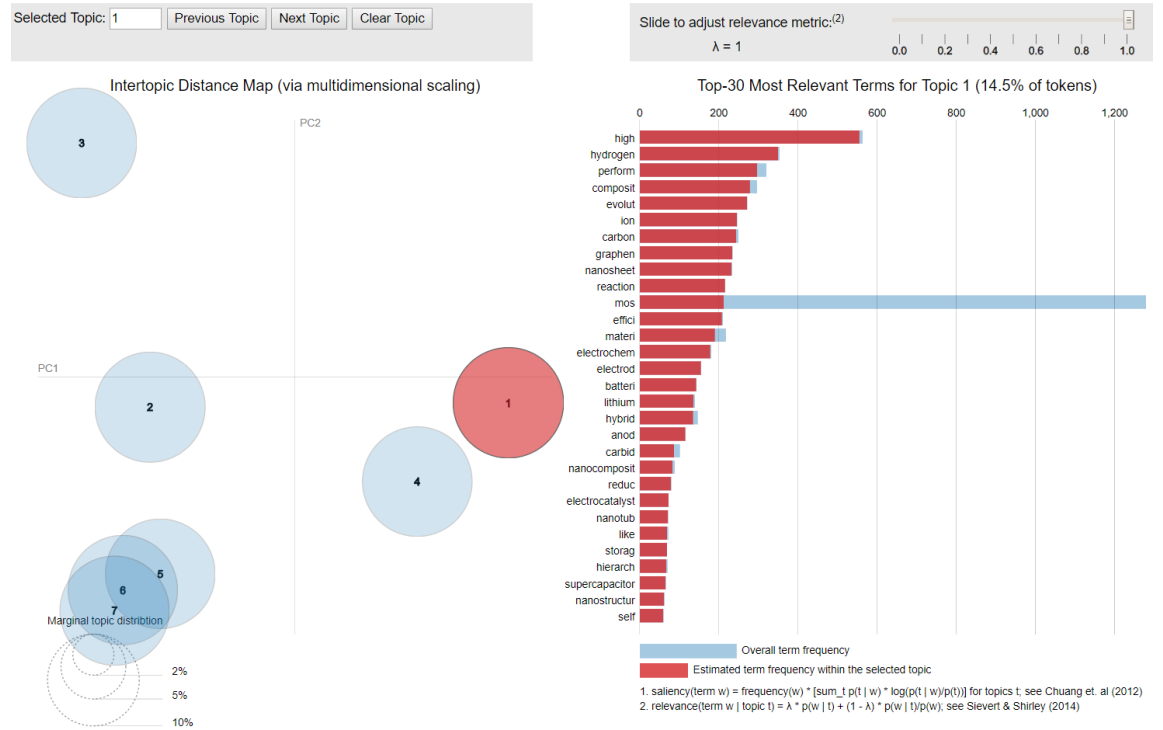


Figura 19. Resultados comparativos de clustering para 6 y 7 tópicos.

5.3.4 Caracterización de tópicos

La caracterización de cada tópico se realiza a partir de las 10 palabras que aparecen con mayor frecuencia. La interpretación de esta información depende del analista, quien debe dar un sentido coherente a cada tópico. La caracterización obtenida en el piloto se muestra en la tabla 10. La distribución de los tópicos se muestra en el gráfico de la figura 20.

Tópico 1	Tópico 2	Tópico 3	Tópico 4	Tópico 5	Tópico 6
metal	effect	oxid	use	high	mos
sup	alloy	hydrogen	film	synthesi	layer
studi	properti	catalyst	deposit	base	monolay
structur	mechan	activ	cell	perform	electron
phase	coat	evolut	thin	moo	dimension
complex	microstructur	dope	solar	composit	two
transit	temperatur	reaction	process	carbon	optic
character	steel	effici	organ	ion	growth
system	behavior	enhanc	applic	nanosheet	disulfid
model	surfac	catalyt	element	graphen	chemic

Tabla 10. Caracterización de los tópicos obtenidos.

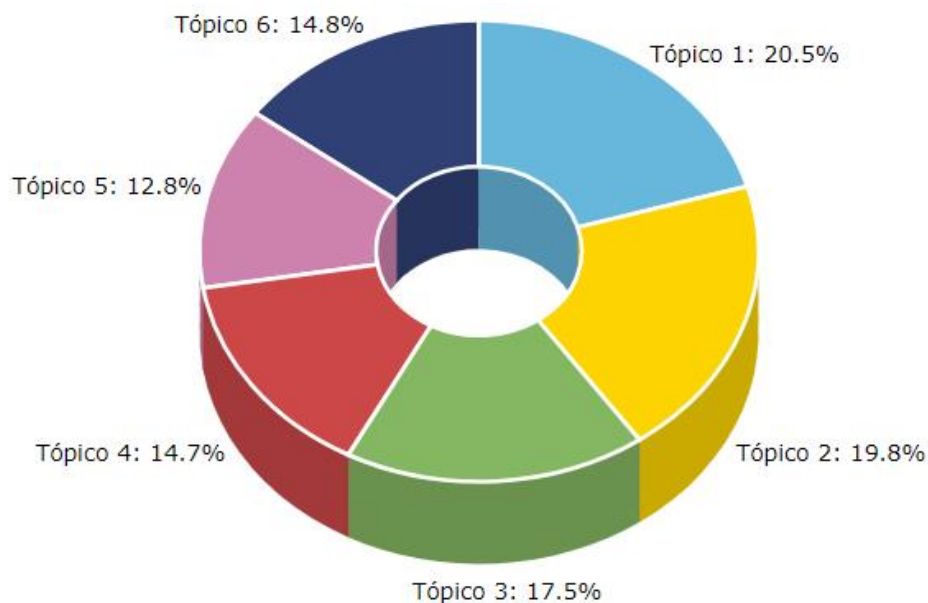


Figura 20. Distribución de tópicos para la búsqueda definida.

5.3.5 Sistema de recomendación

Una vez establecido cada t3pico, se exportar3 la informaci3n de los documentos contenidos en cada uno de ellos ordenados por citasiones, de manera de contar con la mejor informaci3n conducente a la toma de decisiones respecto de que proyecto realizar. La tabla 11 muestra un ejemplo de exportaci3n de informaci3n para el t3pico 3 de la tabla 10, incluyendo los primeros 5 documentos.

Documento	Citaciones	T3tulos	Resumen
1	103	Design of active and stable Co-Mo-Sx chalcogels as pH-universal catalysts for the hydrogen evolution reaction	© 2016 Macmillan Publishers Limited. All rights reserved. Three of the fundamental catalytic limitations that have plagued...
2	102	Coupled molybdenum carbide and reduced graphene oxide electrocatalysts for efficient hydrogen evolution	Electrochemical water splitting is one of the most economical and sustainable methods for large-scale hydrogen production. However, the development of low-cost and earth-abundant...
3	93	General Formation of M-MoS ₃ (M = Co, Ni) Hollow Structures with Enhanced Electrocatalytic Activity for Hydrogen Evolution	© 2015 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim. Complex molybdenum-based ternary or multinary sulfides hollow structures are developed via a fast precipitation process...
4	81	Porous MoO ₂ Nanosheets as Non-noble Bifunctional Electrocatalysts for Overall Water Splitting	The porous MoO ₂ nanosheets were directly grown on commercial nickel foam without binder by a simple wet-chemical route first and then with an annealing treatment...
5	74	Contributions of Phase, Sulfur Vacancies, and Edges to the Hydrogen Evolution Reaction Catalytic Activity of Porous Molybdenum Disulfide Nanosheets	© 2016 American Chemical Society. Molybdenum disulfide (MoS ₂) is a promising nonprecious catalyst for the hydrogen evolution reaction (HER) that has been extensively studied due to its excellent...

Tabla 11. Ejemplo de exportaci3n de informaci3n para el t3pico 3 de la tabla 10.

5.3.6 Subtópicos

Dependiendo de la cantidad de documentos resultantes de la ecuación de búsqueda, así como el número de tópicos determinado como óptimo, el proceso de *clustering* puede realizarse nuevamente, pero aplicado a los documentos contenidos en cada tópico particular. La idea de este proceso es poder especificar aún más la información contenida en el *cluster* de documentos y obtener subtópicos, es decir, tópicos específicos dentro del tópico más general. Básicamente consiste en la realización del proceso ya descrito (Selección de tópicos, interpretación de tópicos y exportación de información) pero utilizando un subconjunto particular de documentos pertenecientes a alguno de los tópicos cuya información se requiera especificar en vez del *cluster* completo de documentos. Para que la información obtenida sea práctica y efectiva, el subconjunto de documentos utilizados no debería disminuir de 500 documentos. A modo de ejemplo se realizó el análisis de subtópicos para el tópico 2 de la tabla 10. La caracterización de los subtópicos se muestra en la tabla 12. La distribución de los subtópicos se muestra en la figura 21.

Subtópico 1	Subtópico 2	Subtópico 3	Subtópico 4
high	coat	alloy	effect
temperatur	mos	properti	steel
oxid	base	mechan	behavior
low	composit	microstructur	corros
laser	surfac	metal	microstructur
structur	luenc	titanium	resist
addit	properti	applic	weld
investig	tribolog	film	differ
use	lubric	phase	wear
properti	studi	carbid	heat

Tabla 12. Caracterización de los subtópicos obtenidos para el tópico 2.

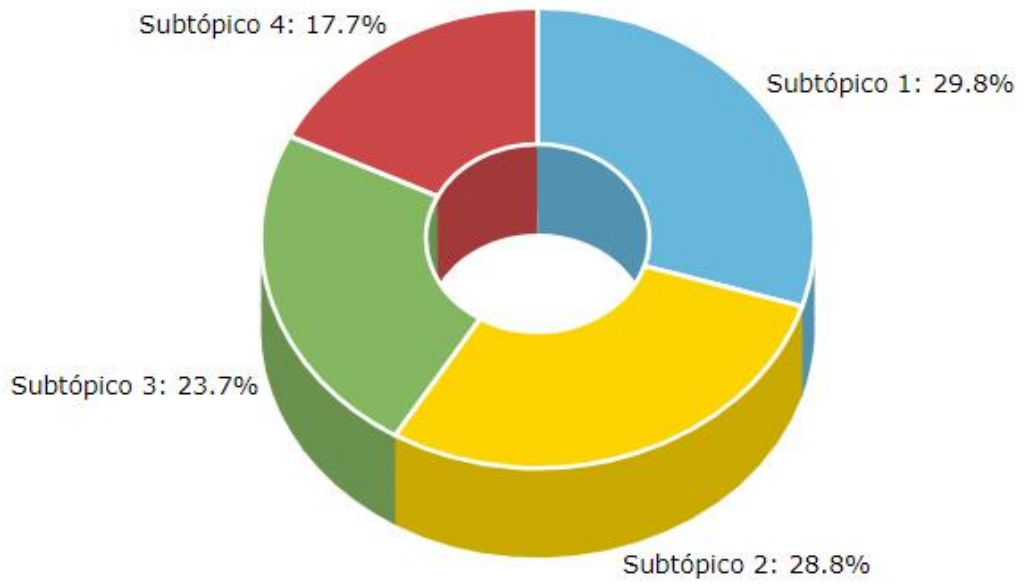


Figura 21. Distribución de subtópicos para el tópico 2.

CAPÍTULO 6: PROPUESTA DE APOYO TECNOLÓGICO

6.1 Especificación de requerimientos

De acuerdo con lo especificado en las secciones anteriores, el sistema debe permitir obtener información desde fuente de interés para el objetivo de negocio (publicaciones científicas o patentes), para un criterio de búsqueda definido por el usuario. Al respecto, los requerimientos funcionales y no funcionales se detallan a continuación.

6.1.1 Requerimientos funcionales

- La importación de información se realizará desde una base de datos competente y atinente y que permita la importación masiva de información desde una aplicación (es decir, que posea una API). La mayor parte de este tipo de bases de datos son de pago. Actualmente la empresa posee dos membresías: Una en Scopus (base de datos de publicaciones científicas) y otra en Acclaim IP (base de datos de patentes).
- La información se importará desde la aplicación desarrollada en R, para trabajar directamente con ella sin necesidad de procesos intermedios.
- El usuario debe indicar el criterio de búsqueda, incluyendo el año de interés.
- La información importada incluye: Título, resumen y citas de cada publicación.
- El modelo no selecciona automáticamente el número óptimo de tópicos K, por lo que este es necesariamente un parámetro que debe ingresar el usuario. La aplicación ofrece un asistente de ayuda para facilitar la selección del K.
- El asistente para ayudar a seleccionar el K muestra el mapa con la distancia intertópica y las palabras contenidas en cada tópico.

- La herramienta de visualización permite exportar un documento que muestra cada tópico obtenido del análisis de los datos. Dentro de cada tópico, se muestran los documentos ordenados por citas, incluyendo la información de título y resumen.

6.1.2 Requerimientos no funcionales

- En cuanto al rendimiento, no existe un requisito específico respecto del tiempo de operación, aunque para no extender excesivamente el análisis, se espera una respuesta en un tiempo inferior a las 2 horas.
- En cuanto a la disponibilidad, la herramienta debe estar siempre disponible.
- En cuanto a la mantención, no se considera un plan de mantención específico y ésta solo se realizará si se solicita.
- En cuanto al formato, solo se considera aplicación de escritorio y no aplicaciones portátiles.
- En cuanto a la seguridad y acceso, solo se considera la instalación en un computador dentro de la VPIT y no instalaciones en otros computadores. Se dará acceso solo al analista, mientras que las jefaturas podrán consultar los reportes de búsqueda.

6.2 Arquitectura tecnológica

El presente proyecto considera la utilización de modelos de minería de textos para poder analizar grandes volúmenes de información atinentes al negocio de la compañía y apoyar así la toma de decisiones respecto de que proyectos de investigación convendría realizar. Al respecto, toda la aplicación se ha desarrollado utilizando una sola herramienta tecnológica que es el lenguaje de programación R.

El desarrollo se ha realizado de modo de que pueda integrarse de manera sencilla pero efectiva con el proceso actual de I+D. La figura 22 muestra un esquema de la arquitectura tecnológica utilizada.



Figura 22. Arquitectura tecnológica utilizada.

Con respecto a la arquitectura tecnológica se puede comentar lo siguiente:

- Ingreso del criterio de búsqueda: El usuario ingresa el criterio de búsqueda a través de la interfaz de usuario.
- Procesamiento de consulta a base de datos: Se procesa la consulta a la base de datos utilizando la información del criterio de búsqueda. Para ello se utiliza un *crawler*.
- Consulta a base de datos: Se consulta a la base de datos correspondiente y se importa la información de títulos, resúmenes y citas.
- Procesamiento de los datos: Con la información importada, se procede a la limpieza de la información y al procesamiento utilizando modelos de *clustering* de minería de textos.
- Visualización de la información: Se visualiza el resultado del *clustering* obtenido, incluyendo recomendación de documentos, así como el análisis de los tópicos de ser necesario (Determinación del número óptimo de tópicos).
- Generación de reportes: Se exportan reportes que resumen la información del procesamiento realizado incluyendo los tópicos y la recomendación de documentos.

6.3 Diseño de la aplicación

El diseño de la aplicación, incluyendo los diagramas correspondientes, se detalla a continuación.

6.3.1 Diagrama de casos de uso

El diagrama de casos de uso de la aplicación se muestra en la figura 23.

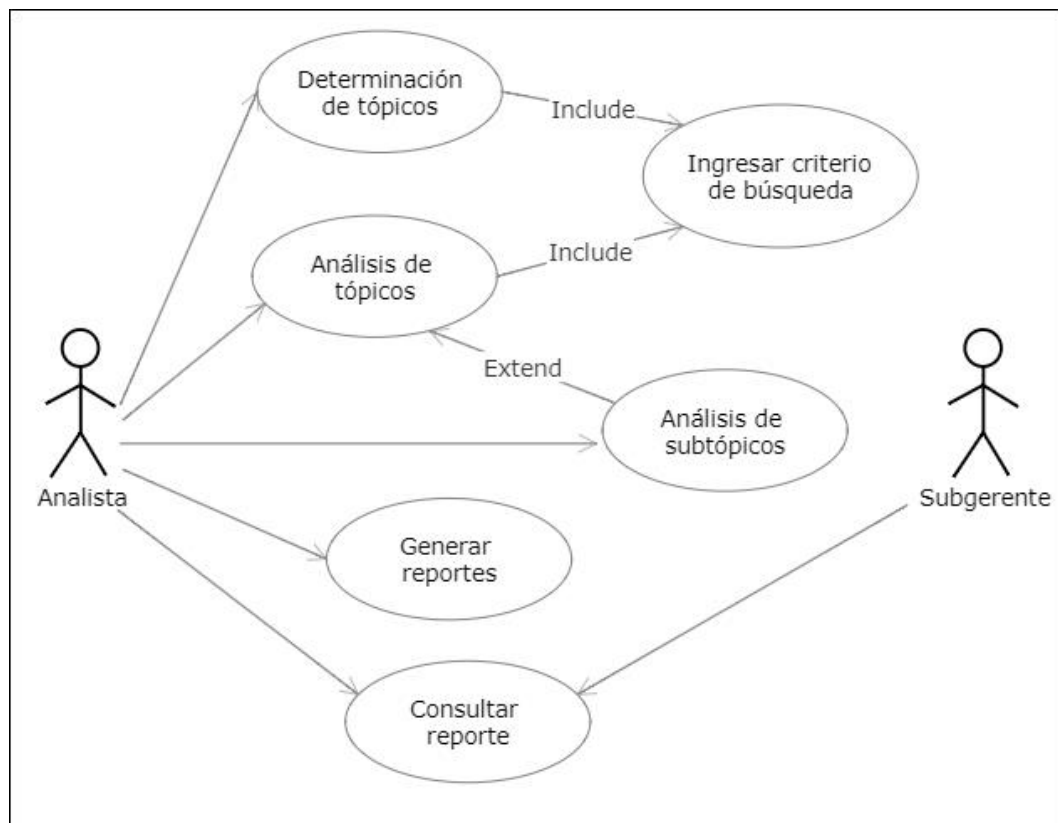


Figura 23. Diagrama de casos de uso de la aplicación.

Al respecto se puede comentar lo siguiente:

- Determinación de tópicos: Analista consulta la herramienta de ayuda para la selección de tópicos para lo cual debe ingresar un criterio de búsqueda.

- Análisis de tópicos: Analista consulta las tendencias dentro del *cluster* de documentos. Debe ingresar un criterio de búsqueda.
- Análisis de subtópico: Analista consulta las tendencias más específicas dentro de alguno de los tópicos.
- Generar reporte: Analista genera reporte con la información resumida del resultado de la búsqueda.
- Consulta reporte: Analista o subgerente pueden consultar los reportes de búsqueda generados anteriormente.

6.3.2 Diagrama de secuencia

La figura 24 muestra el diagrama de secuencia del caso de uso principal que es el análisis de tópicos.

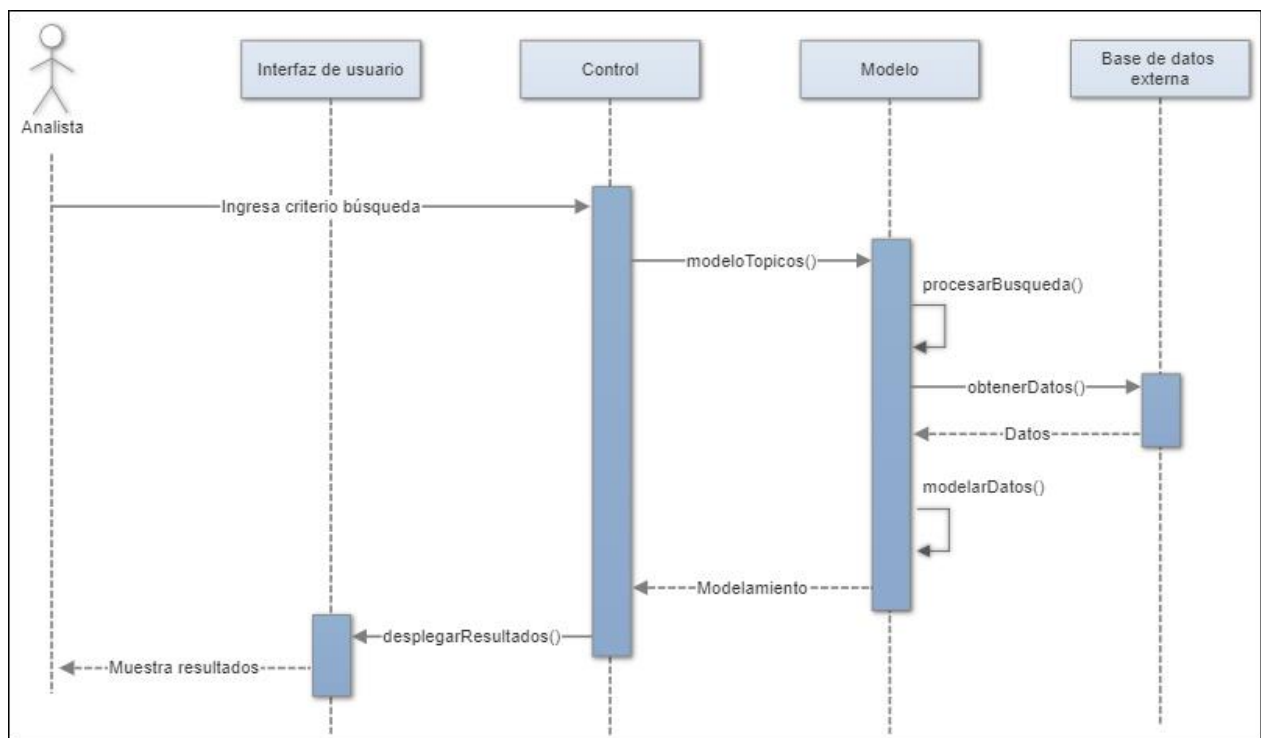


Figura 24. Diagrama de secuencia para el caso de uso de análisis de tópicos.

6.4 Prototipo funcional desarrollado

La figura 25 muestra una captura de la aplicación de visualización realizada.

Aplicación para análisis de publicaciones

The screenshot displays a web application interface for publication analysis, divided into two main sections:

- Búsqueda (Search):** This section includes a header, a search equation input field, a search year input field (containing '0'), and a 'Buscar y guardar' button.
- Análisis de la información (Information Analysis):** This section is further divided into two sub-sections:
 - Análisis de tópicos (Topic Analysis):** Includes a search equation input field, a search year input field (containing '0'), a note stating '*La búsqueda debe pertenecer a la base de datos', a number of topics input field (containing '0'), and buttons for 'Analizar' and 'Guardar reporte'.
 - Análisis de subtemas (Subtopic Analysis):** Includes a note stating '*Debe guardarse el reporte de tópicos primero', a topic to analyze input field (containing '0'), a number of topics within the subtopic input field (containing '0'), and buttons for 'Analizar' and 'Guardar reporte'.

Figura 25. Captura de la aplicación de visualización de resultados desarrollada.

La aplicación posee dos funcionalidades definidas y separadas en bloques:

- **Búsqueda:** Permite realizar una búsqueda en Scopus, para lo cual se solicita una ecuación de búsqueda y un año de búsqueda. La información obtenida se guarda y puede utilizarse en el otro bloque. La idea de separar búsqueda y análisis reside en que la búsqueda debe realizarse

necesariamente en un computador con acceso a la base de datos (lo que implica una cuenta de pago para el caso de Scopus).

- **Análisis de la información:** Permite realizar el análisis de alguna búsqueda anteriormente guardada. El botón Analizar permite determinar el número de tópicos utilizando la herramienta de selección de tópicos de la sección 5.3.3 y el botón Guardar reporte permite exportar la información en forma de un archivo html. Se considera además tanto el análisis de tópicos de una búsqueda como de los subtópicos dentro de cada tópico. La figura 26 muestra un ejemplo de exportación de un reporte incluyendo el análisis de tópicos. La figura 27 muestra un ejemplo de exportación de un reporte incluyendo hasta el análisis de subtópicos de un tópico particular.

Reporte de vigilancia tecnológica - tópicos

Ecuación de búsqueda: molybdenum

Año de búsqueda: 2016

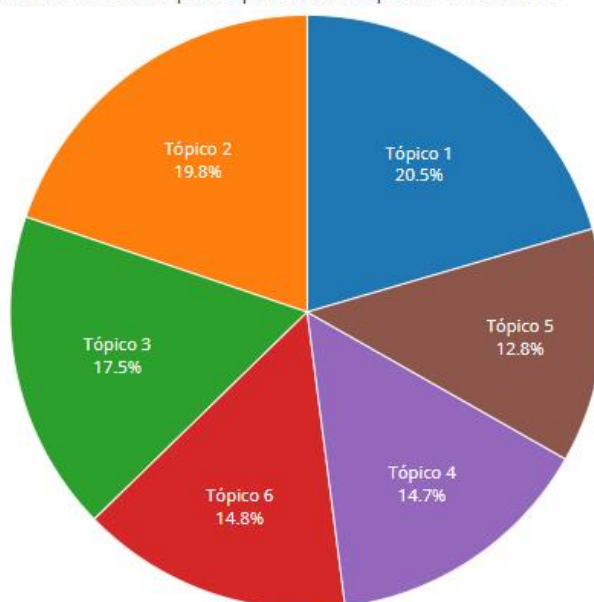
Análisis de la información

Cantidad de documentos: 5665

Número de tópicos: 6

La distribución de tópicos se muestra en la siguiente figura:

Distribución de tópicos para la búsqueda de interés



Tópico 1	Tópico 2	Tópico 3	Tópico 4	Tópico 5	Tópico 6
metal	effect	oxid	use	high	mos
sup	alloy	hydrogen	film	synthesi	layer
studi	properti	catalyst	deposit	base	monolay
structur	mechan	activ	cell	perform	electron
phase	coat	evolut	thin	moo	dimension
complex	microstructur	dope	solar	composit	two
transit	temperatur	reaction	process	carbon	optic
character	steel	effici	organ	ion	growth
system	behavior	enhanc	applic	nanosheet	disulfid
model	surfac	catalyt	element	graphen	chemic

Documentos por tópico:

[Tópico 1](#) - [Tópico 2](#) - [Tópico 3](#) - [Tópico 4](#) - [Tópico 5](#) - [Tópico 6](#)

Figura 26. Reporte HTML generado para la búsqueda definida en el piloto.

Reporte de vigilancia tecnológica - subtópicos

Ecuación de búsqueda: molybdenum

Año de búsqueda: 2016

Se obtuvieron los siguientes tópicos:

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6
metal	effect	oxid	use	high	mos
sup	alloy	hydrogen	film	synthesi	layer
studi	properti	catalyst	deposit	base	monolay
structur	mechan	activ	cell	perform	electron
phase	coat	evolut	thin	moo	dimension
complex	microstructur	dope	solar	composit	two
transit	temperatur	reaction	process	carbon	optic
character	steel	effici	organ	ion	growth
system	behavior	enhanc	applic	nanosheet	disulfid
model	surfac	catalyt	element	graphen	chemic

Tópico de interés: 2

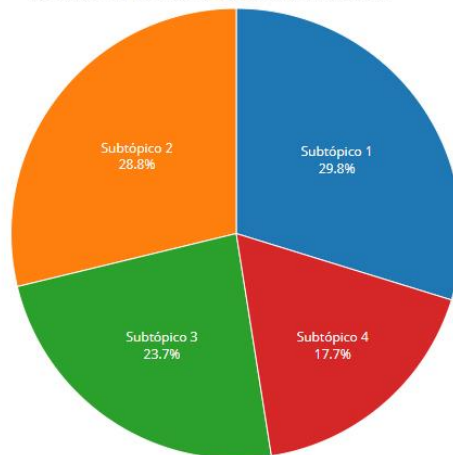
Análisis de la información

Cantidad de documentos en tópico 2: 1122

Número de subtópicos: 4

La distribución de subtópicos se muestra en la siguiente figura:

Distribución de subtópicos para el tópico 2



Subtópico 1	Subtópico 2	Subtópico 3	Subtópico 4
high	coat	alloy	effect
temperatur	mos	properti	steel
oxid	base	mechan	behavior
low	composit	microstructur	corros
laser	surfac	metal	microstructur
structur	luenc	titanium	resist
addit	properti	applic	weld
investig	tribolog	film	differ
use	lubric	phase	wear
properti	studi	carbid	heat

Documentos por subtópico:

Subtópico 1 - Subtópico 2 - Subtópico 3 - Subtópico 4

Figura 27. Reporte HTML generado para la búsqueda definida en el piloto. Se incluye el análisis de subtópicos para el tópico 2.

CAPÍTULO 7: GESTIÓN DEL CAMBIO

7.1 Contexto del cambio

Como ya se ha comentado en secciones anteriores, el cambio se implementará en el proceso de innovación de capacidades de Molytmet, que es llevado a cabo por la Vicepresidencia de Investigación, Desarrollo e Innovación Tecnológica (VPIT).

La VPIT cuenta con 18 personas con dedicación exclusiva al estudio de proyectos de investigación, pudiendo ser estos de mejora de capacidades o de innovación en capacidades. Inicialmente era solo investigación y desarrollo, pero a finales de 2012 se decidió añadir a la innovación como eje central impulsora de nuevos negocios para la compañía, es decir, como impulsora del proceso de innovación de capacidades puesto que la mayoría de los proyectos solían apuntar solo hacia mejora continua. Todo esto obedece al fin último de potenciar el crecimiento de la empresa utilizando a la innovación como motor.

La selección de los proyectos de investigación es un tema crucial puesto que permite enfocar los esfuerzos ya sea en mejora o en innovación. De ahí la idea de mejorar el proceso de vigilancia tecnológica, entendiéndolo como un medio sustentable para generación de ideas conducentes a nuevos negocios.

Actualmente el proceso es llevado a cabo de manera manual por la Subgerencia de Innovación Tecnológica.

7.2 Observación de la implementación a realizar

La implementación por realizar propone que el proceso de vigilancia tecnológica se lleve a cabo de manera inteligente, utilizando una herramienta de apoyo tecnológico de manera de poder analizar de manera global toda la información disponible, lo que claramente no es posible a través de un proceso manual, debido a los grandes volúmenes de información involucrados. Esta herramienta permitirá respaldar la toma de decisiones respecto de que proyectos conviene finalmente estudiar.

El cambio principal deberá estar entonces en el personal que está involucrado actualmente en el proceso de vigilancia tecnológica, es decir, el Subgerente de Innovación Tecnológica y el investigador que actúa como analista. Por otro lado, también debe considerarse el cambio que esto conlleva para los demás Subgerentes y el propio Vicepresidente, quienes serán los receptores de los resultados de la mejora propuesta.

7.3 Análisis de los principios de cambio

El cambio propuesto, si bien implica una manera diferente de hacer las cosas, no implica un cambio radical al proceso de investigación y ni siquiera al proceso de selección de ideas conducentes a generar proyectos: Se mantiene a la vigilancia tecnológica para este fin, pero se mejora considerablemente su funcionamiento y utilidad. Por lo anterior se trata de un cambio de primer orden.

Todo el proceso de innovación de capacidades se mantendrá sin modificaciones y solo se rediseñará el proceso de vigilancia tecnológica. Además, no se requerirá ningún conocimiento nuevo en particular para poder usar la herramienta.

Es importante considerar los estados de ánimo del personal involucrado. Debido a que se trata de profesionales altamente capacitados y además ligados al área de la innovación, deberían comprender la necesidad del cambio y las ventajas que puede traer el mismo. Sin embargo, nunca se ha aplicado una herramienta tecnológica de este tipo, por lo que puede existir cierta incredulidad respecto de su funcionamiento sobre todo considerando que no hay familiaridad con los conceptos que hay detrás.

En cuanto al manejo del poder, la herramienta solo considera entregar información para apoyar la toma de decisiones y no es un tomador de decisiones en sí mismo, por lo que no debería generar conflictos al respecto. Además, la información pasará siempre por el Subgerente de Innovación antes de construir las propuestas finales de proyectos de investigación.

El proyecto concluye con la aplicación de un piloto de la aplicación a un caso real de búsqueda y análisis de información relevante para la compañía.

7.4 Caracterización del cambio

El cambio propuesto se introducirá a partir del analista y el Subgerente de Innovación. La idea es que ellos comprendan inicialmente como funciona la herramienta propuesta a partir de secciones de entrenamiento con algunos casos pequeños de uso y los resultados del piloto. La segunda parte consiste en expandir el uso de la herramienta hacia el resto de los usuarios, es decir, los demás Subgerentes y el Vicepresidente, ya que son ellos los que utilizarán la información generada para apoyar su toma de decisiones. En esta segunda etapa, la estrategia se centrará en mostrar los resultados que se pueden obtener con la aplicación de la herramienta, considerando tanto el piloto como las demás pruebas que se hayan realizado. Es importante mencionar que no se consideran cambios organizacionales para la implementación del cambio.

La evaluación del avance del proceso de cambio es bastante directa y consiste en medir cuantas ideas de proyecto se están generando a partir de la herramienta propuesta. Una vez que el proyecto esté implementado se considerarán otras mediciones, como por ejemplo la cantidad de ideas generadas que conducen efectivamente a la generación de nuevos negocios y la cantidad de propuestas generadas que se convierten finalmente en proyectos de investigación.

7.5 Factores críticos de éxito

Los factores críticos de éxito detectados para la implementación del proyecto son los siguientes:

- **Credibilidad de la propuesta:** El factor de éxito más importante es la credibilidad de la propuesta. Todas las personas involucradas deben creer en la propuesta. En primer lugar los ejecutores, puesto que serán los que lleven a cabo el proceso, y en segundo lugar los receptores, quienes recibirán la información que los ayudará a la toma de decisiones.

De hecho, uno de los objetivos principales del piloto realizado durante el MBE es mostrar resultados para incentivar el uso futuro de la aplicación.

- Resistencia al cambio de los ejecutores: Debido a que el personal ya está acostumbrado a trabajar de una cierta manera, puede haber resistencia a cambiar las prácticas. Para mitigar este asunto es importante contar con una herramienta tecnológica sencilla y fácil de utilizar, puesto que la utilizará personal inexperto respecto a aplicaciones tecnológicas de esta naturaleza.
- Confiabilidad de las fuentes de información: Se deben escoger fuentes de información confiables y atingentes al negocio de la compañía tanto para publicaciones científicas como para patentes.

7.6 Plan de gestión del cambio

El plan de gestión del cambio considera los siguientes puntos:

- Desarrollo del piloto del proyecto incluyendo resultados prácticos del uso de la aplicación. Esto es parte del proyecto MBE.
- Presentación de la herramienta tecnológica al Subgerente de Innovación y al analista investigador a través de pequeñas pruebas conducentes a determinar su funcionalidad de manera práctica. Para ello se considerarán cuatro sesiones de dos horas cada una. Adicionalmente, se mostrará el resultado del piloto para incentivar aún más el uso de la aplicación.
- Presentación de la aplicación a los demás Subgerentes y al Vicepresidente, enfocado en los resultados obtenidos. Para ello se considerarán tanto los resultados del piloto como los resultados de las pequeñas pruebas realizadas en el punto anterior. El enfoque de este punto estará centrado en mostrar la potencialidad de la herramienta como fuente generadora de ideas para proyectos.
- Se propondrá la utilización del proceso actual de vigilancia tecnológica para analizar la misma información que la utilizada en el piloto y así

mostrar comparativamente los resultados obtenidos. Esto último busca potenciar el convencimiento de que la herramienta propuesta funciona mejor que el proceso actual.

- Una vez aceptada la aplicación, se propondrá un piloto más extenso que considerará el uso de la herramienta para determinar varias ideas de proyectos a partir de varias ecuaciones de búsqueda de interés para la compañía. Estas búsquedas son las que ya se describieron en la sección 3.2 de los lineamientos estratégicos de la compañía. La idea es medir cuantas ideas de proyecto se han generado efectivamente y cuantas de ellas fueron aceptadas como proyectos a desarrollar. El resultado obtenido se comparará con los proyectos actuales y los de años anteriores, para comprobar si efectivamente se aumentó el número de proyectos conducentes a nuevos negocios y si aumentó el número de propuestas de proyectos en general. Con esto, se presentará el resultado definitivo de la aplicación del cambio.

CAPÍTULO 8: EVALUACIÓN DEL PROYECTO

8.1 Resultados y análisis del piloto

Para validar los resultados del procesamiento de la información, se realizó un análisis de interpretación de los tópicos en conjunto con Juan José Segura (Subgerente de Procesos Metalúrgicos de la VPIT), de manera de contar con una opinión externa y de juicio experto. Para este análisis se utilizó la información de la caracterización de los tópicos de la tabla 10 y la inspección directa de algunos de los documentos de cada tópico escogidos al azar. Además, se realizó el mismo análisis para los subtópicos del tópico 2, utilizando la información de la tabla 12. Los resultados obtenidos se muestran en las tablas 13 y 14.

Tópico	Interpretación	Cantidad de documentos
1	Propiedades, estructura química y caracterización	20.5%
2	Aleaciones y compósitos	19.8%
3	Catálisis	17.5%
4	Electroquímica y energía	14.7%
5	Materiales	12.8%
6	Recubrimientos y fenómenos interfaciales	14.8%

Tabla 13. Resultados del análisis de interpretación de tópicos.

Subtópico	Interpretación	Cantidad de documentos
1	Usos y aplicaciones	29.8%
2	Lubricantes	28.8%
3	Propiedades químicas y mecánicas	23.7%
4	Acero inoxidable	17.7%

Tabla 14. Resultados del análisis de interpretación de subtópicos.

Al respecto de los resultados obtenidos se puede comentar lo siguiente:

- La información obtenida es razonable, es decir, está dentro de cierto margen coherente esperado.

- La interpretación de tópicos debe ser realizada por un experto en la información analizada porque requiere de conocimiento específico.
- La interpretación de tópicos no puede hacerse utilizando solo la información de la caracterización de los tópicos y está sujeta al análisis de algunos documentos. Lo anterior se debe a que la relación entre las palabras a veces es compleja y puede no ser determinable directamente. La inspección de algunos documentos permitió contextualizar las palabras obtenidas en la caracterización.
- El tipo de análisis que se presenta tiene una tremenda ventaja respecto de la búsqueda tradicional, ya que permite analizar toda la información disponible y no solo un pequeño subconjunto. En las búsquedas tradicionales se suele acotar mucho las ecuaciones de búsqueda hasta obtener una cantidad de documentos que se pueda analizar manualmente, lo que limita la cantidad de información que puede recuperarse.
- La idea es realizar una intersección entre la estrategia de la empresa y la información encontrada de manera de decidir qué cosas conviene investigar, que alianzas estratégicas realizar y en definitiva no "reinventar la rueda". Independiente de la estrategia adoptada por la empresa, siempre es crucial contar con información del entorno competitivo.
- El presente proyecto se ha convertido en un incentivo a la resolución de problemas por una vía paralela a la ingeniería tradicional (ruptura de paradigma). Un ejemplo de esto es la resolución de problemas operacionales a través del análisis de datos, proyecto que se está tratando de implementar actualmente.

Adicional a la interpretación de los tópicos, se realizó una validación manual de la información obtenida. Para ello se seleccionaron 20 documentos al azar dentro de cada tópico y se comprobó si el tópico asignado era efectivamente el que se interpretó. Los resultados obtenidos se muestran en la tabla 15. Como se observa, se obtiene una alta precisión superando siempre el 70%. Los tópicos que presentan menor asertividad son los tópicos 1 y 4, que corresponden efectivamente a tópicos más ambiguos y difíciles de interpretar, tanto por las palabras que los caracterizan como por la temática menos acotada que abarcan.

Tópico 1		Tópico 2		Tópico 3		Tópico 4		Tópico 5		Tópico 6	
Docs.	T1	Docs.	T2	Docs.	T3	Docs.	T4	Docs.	T5	Docs.	T6
819	Si	998	Si	394	Si	668	No	384	Si	713	Si
24	No	621	Si	400	Si	758	Si	304	Si	611	Si
936	Si	492	Si	642	Si	517	Si	413	Si	605	Si
282	Si	742	Si	83	Si	105	No	439	Si	577	Si
407	Si	518	Si	548	Si	303	No	356	Si	180	No
117	Si	33	Si	903	Si	740	Si	568	Si	541	Si
956	No	139	Si	869	No	482	Si	310	Si	490	Si
100	Si	880	No	818	Si	807	Si	61	Si	452	Si
286	Si	901	Si	489	Si	288	No	576	Si	9	Si
1006	Si	85	Si	190	Si	492	Si	597	No	498	Si
559	No	1041	No	206	Si	35	Si	265	Si	186	Si
337	Si	476	Si	883	No	747	Si	405	Si	115	No
911	No	238	No	855	Si	280	Si	154	Si	53	Si
972	Si	502	Si	743	Si	30	Si	162	Si	72	Si
1134	Si	217	Si	765	Si	621	No	361	Si	508	Si
683	Si	503	Si	188	Si	145	Si	499	Si	21	Si
1001	Si	605	Si	985	Si	610	Si	198	Si	125	Si
8	No	196	Si	468	Si	406	Si	105	Si	803	No
189	Si	1072	Si	742	Si	32	No	115	Si	598	Si
493	Si	1052	Si	232	Si	452	Si	614	Si	613	Si
75%		85%		90%		70%		95%		85%	

Tabla 15. Resultados obtenidos en la validación de tópicos.

8.2 Definición de beneficios y costos

A pesar de que se cuenta con un plan bien definido y resultados producto de ese plan, los beneficios económicos no son medibles directamente o al menos no en el corto plazo. Para estimar una cota inferior de los beneficios económicos del proyecto se utilizará data histórica de los resultados de la VPIT en comparación con el índice de proyectos exitosos:

- Gasto anual de la VPIT: El gasto anual promedio de la VPIT para el periodo 2010-2017 fue de 1.5 MMUSD.

- Índice de proyectos exitosos: El índice de proyectos exitosos de la VPIT para el periodo 2000-2017 fue de 11%.

Los resultados de I+D dentro de una compañía siempre se ven en el largo plazo. En este sentido, a pesar de que la VPIT pueda generar solo costos para la empresa en algún año en particular, en el largo plazo debe poder justificar (como mínimo) su existencia económicamente hablando. Esto implica que un 11% de proyectos exitosos permite generar como mínimo un ingreso de 1.5 MMUSD/año en el largo plazo. Si se asume una proporcionalidad directa entre el índice de proyectos exitosos y los ingresos que genera la VPIT para Molymet, se obtendría un aumento de 0.14 MMUSD/año por cada punto porcentual de aumento en el índice de proyectos exitosos. Por otro lado, si se asume que la mejora propuesta permitiría aumentar en un 3% el índice de proyectos exitosos (Valor razonable y poco ambicioso que no permite alcanzar el 20% objetivo) se obtendría un aumento en los ingresos de 0.41 MMUSD/año.

Los costos del proyecto pueden dividirse en costos de implementación y costos de operación del proyecto. La tabla 16 muestra los costos de implementación. Como el proyecto no considera la contratación de nuevo personal, se trata en realidad de costos de oportunidad del personal necesario que ya pertenece a la empresa. Se consideró un gasto por contingencia de un 15%.

Modelamiento

PERSONAL	HH	UF/HORA	TOTAL
Vicepresidente	6	2	\$321,456
Subgerente Innovación	6	0.75	\$120,546
Subgerente Procesos Metalúrgicos	36	0.75	\$723,276
Analista	36	0.42	\$8,100,691
Ingeniero de proyecto	720	0.42	\$405,035
Contingencias			\$1,450,651
Total			\$10,716,620

Aplicación de visualización de datos

ACTIVIDADES	HH	UF/HORA	TOTAL
Configuración PC	3	1	\$80,364
Desarrollo aplicación	50	1.5	\$2,009,100
Desarrollo UI	6	1.5	\$241,092
Pruebas	12	1	\$321,456
Capacitación	6	1	\$160,728
Confección de manuales	6	1	\$160,728
Contingencias			\$446,020
Total			\$3,419,488

Tabla 16. Costos de implementación del proyecto.

La tabla 17 muestra los costos de operación del proyecto. Nuevamente se trata en su mayoría de costos de oportunidad del personal ya existente. La política TI de la empresa considera el arriendo de equipos computacionales (no hay adquisición de este tipo de equipos), por lo que debe considerarse el costo del arriendo del pc en donde se instalará la aplicación de visualización de datos. Se consideró un costo por contingencias de un 15%.

PERSONAL	HH	UF/HORA	TOTAL
Vicepresidente	6	2	\$321,456
Subgerente Procesos Metalúrgicos	6	0.75	\$120,546
Subgerente Materiales	6	0.75	\$120,546
Subgerente Innovación	24	0.75	\$482,184
Analista	360	0.42	\$4,050,346
Ingeniero de proyecto	12	0.42	\$135,012
Soporte TI	20	0.2	\$107,152
Arriendo computador			\$240,000
Contingencias			\$836,586
Total			\$6,413,827

Tabla 17. Costos operacionales del proyecto.

8.3 Flujo de caja

Considerando los costos y los beneficios descritos, la tabla 18 muestra el flujo de caja del proyecto. Se utilizó un horizonte de evaluación de 5 años por ser un proyecto tecnológico y de ingeniería de baja inversión que no necesita reemplazarse en corto plazo, pero si cuando la tecnología ofrezca herramientas más avanzadas para este tipo de modelos.

Para el cálculo de la tasa de descuento se utilizó el modelo CAPM (Pues se trata de un flujo de caja de los inversionistas):

$$t_D = R_f + \beta(E(R_m) - R_f)$$

Donde:

t_D : Tasa de descuento

R_f : Tasa libre de riesgo

$E(R_m)$: Retorno esperado sobre el portafolio de mercado

β : Coeficiente de riesgo sistemático

Para el caso en estudio se utilizaron las siguientes estimaciones para el cálculo de los parámetros:

- R_f = Se estimó como la tasa de interés promedio de 2017 para los bonos del banco central de Chile a 5 años, resultando ser 3.73%.
- $E(R_m)$ = Se estimó como el promedio de la variación anual del IPSA desde 2003 a 2017 (15 años), resultando ser 14.38%.
- β = Se obtuvo de la información de mercado de la empresa, resultando ser 0.63¹.

Con los parámetros anteriores se obtuvo una tasa de descuento de 10.44%.

¹ <http://www.msn.com/es-cl/dinero/stockdetails/fi-233.1.MOLYMET.SGO>. Consulta: 10/05/18.

	Año 0	Año 1	Año 2	Año 3	Año 4	año 5
Ingresos		\$265,634,503	\$265,634,503	\$265,634,503	\$265,634,503	\$265,634,503
Costos	\$14,136,108	\$6,413,827	\$6,413,827	\$6,413,827	\$6,413,827	\$6,413,827
Utilidad antes de impuesto		\$259,220,676	\$259,220,676	\$259,220,676	\$259,220,676	\$259,220,676
Impuesto a la renta		\$69,989,582	\$69,989,582	\$69,989,582	\$69,989,582	\$69,989,582
Utilidad después de impuesto		\$189,231,093	\$189,231,093	\$189,231,093	\$189,231,093	\$189,231,093
Flujo operacional		\$189,231,093	\$189,231,093	\$189,231,093	\$189,231,093	\$189,231,093
Flujo de caja privado	-\$14,136,108	\$189,231,093	\$189,231,093	\$189,231,093	\$189,231,093	\$189,231,093
Tasa de descuento	10.44%					
VAN	\$695,246,872					
TIR	1339%					

Tabla 18. Flujo de caja del proyecto.

A la vista de los resultados se concluye que el proyecto es económicamente viable con los supuestos utilizados. Como comentarios generales, se puede decir que no se consideró valor residual del proyecto porque no hay inversión en activos fijos y porque se cree que el modelo desarrollado debería estar completamente obsoleto dentro de 5 años, por lo que sería necesario volver a desarrollar un nuevo modelo con la nueva tecnología que estará disponible.

8.4 Análisis de sensibilidad

El supuesto más fuerte del análisis económico realizado dice relación con el aumento en el índice de proyectos exitosos. De hecho, esta misma variable es la única variable riesgosa relevante para el análisis en cuestión, pues las demás se mantienen relativamente constantes porque se trata de honorarios de trabajadores que no suelen cambiar radicalmente en el tiempo. Por lo anterior, se realizó un análisis de sensibilidad del VAN con respecto a esta variable. El resultado se muestra en la tabla 19. Se concluye que el modelo analítico debe permitir aumentar, cuando mínimo, un 0.131% el índice de proyectos exitosos para que el proyecto sea económicamente viable.

Aumento en el índice de proyectos exitosos	VAN
0%	-\$31,688,180
0.131%	\$0
1%	\$210,623,504
3%	\$695,246,872

Tabla 19. Análisis de sensibilidad del VAN respecto al aumento en el índice de proyectos exitosos.

CAPÍTULO 9: CONCLUSIONES

En este capítulo se presentan las conclusiones del trabajo realizado, incluyendo los resultados obtenidos, las lecciones aprendidas y el trabajo futuro. Se separaron en distintos componentes para abarcar de manera separada los diferentes puntos.

9.1 Metodología utilizada

La metodología de ingeniería de negocios utilizada permitió comprender de manera global el problema, ubicándolo dentro de un proceso de negocio perteneciente a la empresa y además facilitando el entendimiento de la interacción de este proceso con los demás. Así mismo permitió alinear el proceso con los objetivos estratégicos de la compañía, de manera de asegurar que se apunte en la dirección correcta. Es interesante comentar que la metodología funciona bien incluso aplicada a áreas no tan convencionales de una empresa como es Investigación y Desarrollo.

El proceso de extracción de conocimiento CRISP-DM fue una buena guía de pasos a seguir para poder desarrollar un proyecto de mejora, incluyendo un proceso de extracción de conocimiento por minería de textos, como el que aquí se presenta.

9.2 Desarrollo y resultados del proyecto

Durante del desarrollo del proyecto se obtuvieron varios resultados y se aprendieron varias cosas en función del grado de avance y las dificultades que iban apareciendo:

- Información para el modelamiento: En un principio se pensó utilizar la información de los resúmenes para el *clustering*, sin embargo, la información de los títulos resultó en tópicos mucho mejor definidos.
- *Crawler*: Se hizo necesario modelar considerando las limitaciones propias que impone la API de Scopus: Los resultados de una búsqueda solo entregan hasta 5000 resultados, por lo que se debe particionar la

búsqueda en caso de que se obtengan más documentos y solo se pueden importar 100 documentos por cada consulta, por lo que hay que hacer varias consultas para poder importar todo el *cluster* de documentos.

- Preprocesamiento: Hay que eliminar muchos signos para poder modelar de manera efectiva. Así mismo se hace necesario eliminar algunas palabras comunes y la ecuación de búsqueda utilizada ya que se repite mucho dentro de los tópicos.
- Modelamiento: Se utilizaron 5 modelos de *clustering* diferentes y se concluyó que LDA es el que entrega los mejores resultados para el caso en estudio. Además, se creó una herramienta de visualización para ayudar a la selección del número óptimo de tópicos.
- Extracción de tópicos: Se extrajeron las 10 palabras que aparecían con mayor probabilidad dentro de cada tópico de manera de poder caracterizarlo.
- Sistema de recomendación: Se diseñó un sistema de recomendación de documentos en función de las citas, de manera de contar con la mejor información disponible.
- Subtópicos: Se consideró la opción de buscar subtópicos dentro de tópicos para especificar en mayor grado la información obtenida.
- Análisis e interpretación: Se realizó un trabajo de análisis e interpretación de los tópicos obtenidos, dándole un sentido real a cada grupo. Este análisis se realizó en conjunto con Juan José Segura de manera de contar con una opinión externa y de juicio experto.
- La interpretación de tópicos debe ser realizada por un experto en la información analizada porque requiere de conocimiento específico. Además, debe considerar la inspección manual de algunos documentos con el fin de contextualizar las palabras que caracterizan a cada tópico.
- Se realizó una validación manual de la interpretación de tópicos, obteniéndose que la precisión fue siempre superior a 70% para todos los tópicos obtenidos. Los tópicos con menor precisión fueron

precisamente aquellos con palabras más ambiguas y con temáticas más amplias y por lo tanto difíciles de interpretar.

- La aplicación desarrollada es muy sencilla por lo que no debiese representar ningún desafío aprender a utilizarla.
- Prototipo funcional y resultados del piloto: El proyecto desarrollado efectivamente permite analizar los grandes volúmenes de información que se requiere y además entrega tendencias perfectamente interpretables de los datos, por lo que se vislumbra una gran utilidad. La información se exporta en archivos HTML conteniendo todo lo necesario para continuar con el análisis posterior de la información y obtener de ello las propuestas de proyectos de investigación. La idea es realizar una intersección entre la información obtenida y la estrategia empresarial con el fin de decidir acerca de los proyectos de investigación que deberían llevarse a cabo.
- Debido a que se trata de un análisis global de toda la información disponible, se obtiene una gran ventaja en comparación con las búsquedas tradicionales que suelen estar acotadas a los documentos abordables por un analista, con la consecuente limitación de información que esto conlleva.
- Imprevistos: Se tuvieron algunos imprevistos menores durante el desarrollo por errores menores o fallas en los programas, pero el más importante estuvo en el bloqueo al acceso a internet desde R por parte del área de TI de la empresa, lo que detuvo el avance del proyecto puesto que Scopus solo permite que la API funcione desde una IP de la VPIT. Para solucionarlo hubo que explicar formalmente las razones por las que se requería este uso y fue necesario configurar manualmente la conexión a través de proxy.

En cuanto a la gestión del cambio, existe un ánimo de incertidumbre por parte del personal involucrado, pero al mismo tiempo de curiosidad y de ganas de involucrarse en el proyecto porque se percibe como algo novedoso, no solo dentro del departamento, sino que también dentro de la empresa. Se espera afianzar este entusiasmo con resultados futuros de utilización de la aplicación en otras búsquedas, y además se espera incentivar el uso de soluciones

inteligentes utilizando apoyo tecnológico, como una alternativa a los proyectos de ingeniería más tradicional.

Respecto de la evaluación económica, se trata de un proyecto altamente rentable con un VAN de 695 MM\$ en un horizonte de 5 años (asumiendo 3% de aumento en el índice de proyectos exitosos) y un análisis de sensibilidad que indica que el proyecto debe permitir cuando menos aumentar en un 0.13% el índice de proyectos exitosos para justificarse económicamente hablando.

9.3 Trabajo futuro

Como trabajo futuro del proyecto se propone lo siguiente:

- Realización de evaluación comparativa del proyecto: Para ello se propone la realización de un piloto más extenso que incluya todas las ecuaciones de búsqueda que dicen relación con los objetivos estratégicos de la compañía (molibdeno y metales afines) para un año en concreto, analizar todas las propuestas de proyectos generadas y comparar con las propuestas de otros años haciendo énfasis además en las propuestas conducentes a nuevos negocios.
- Realización de evaluación real del proyecto: Consiste en la evaluación económica real de cuanto aporte genera la VPIT con el proyecto ya instaurado. Para ello habría que se plantea implementar el proyecto y utilizarlo durante 3 años, calcular el beneficio económico y los costos de todos los proyectos realizados que se hayan generado a partir de la herramienta y comprobar cuanto a aumentado el beneficio que estaría aportando la VPIT a la compañía.
- Extensión de la herramienta a más bases de datos: Extender la herramienta a más bases de datos tanto de publicaciones científicas como de patentes. Inicialmente podría extenderse hacia la base de patentes AcclaimIP, puesto que se cuenta con una membresía pagada.

CAPÍTULO 10: BIBLIOGRAFÍA

Aggarwal. C. (2015). Data mining: The textbook. Switzerland: Springer International Publishing.

Aggarwal. C. & Zhai C. (Eds.) (2012). Mining text data. London, England: Springer Science+Business Media, LLC.

Azevedo, A. & Santos, F. (2008). KDD, SEMMA and CRIPS-DM: A parallel overview. *MCCSIS'08 - IADIS Multi Conference on Computer Science and Information Systems; Proceedings of Informatics 2008 and Data Mining 2008*, 182-185.

Barros, O. (2015). Ingeniería de Negocios: Diseño Integrado de Servicios, sus Procesos y Apoyo TI. Chile: Universidad de Chile.

Berry, M. & Kogan, J. (Eds.) (2010). Text Mining: Applications and theory. United Kingdom: John Wiley & Sons, Ltd.

Blei, D., Ng, A., Jordan, M. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3, 993-1022

Blei, M. & Lafferty, J. (2007). A correlated topic model of science. *The Annals of Applied Statistics*, 1(1), 17-35.

Bramer, M. (2016). Principles of data mining. London, England: Springer-Verlag.

Chen, H., Zhang, G., Zhu, D., Lu, J. (2017). Topic-based technological forecasting based on patent data: A case study of Australian patents from 2000 to 2014. *Technological Forecasting and Social Change*, 119, 39-52.

Choi, H., Lee, W., Sohn, S. (2017). Analyzing research trends in personal information privacy using topic modeling. *Computers and Security*, 67, 244-253.

Fayyad, U., Piatetsky-Shapiro, G., Smyth, P. (1996). "From Data Mining to Knowledge Discovery in Databases". *AI Magazine* 17 (3).

Gao, S., Li, X., Yu, Z., Qin, Y., Zhang, Y. (2017). Combining paper cooperative network and topic model for expert topic analysis and extraction. *Neurocomputing*, 257, 136-143.

Grün, B. & Hornik, K. (2011). topicmodels: An R Package for Fitting Topic Models. *Journal of Statistical Software*, 40(13).

- Hax A., Wilde, D. (2003). The Delta Model – A New Framework of Strategy. *Journal of Strategic Management Education*, 1(1).
- Kwartler, T. (2017). Text mining in practice with R. United Kingdom: John Wiley & Sons, Ltd.
- Ma, W., Luo, X., Xuan, J., Xue, R., Guo, Y. (2017). Discover semantic topics in patents within a specific domain. *Journal of Web Engineering*, 16(7-8), 653-675.
- Mathew, G., Agrawal, A., Menzies, T. (2017). Trends in topics at SE conferences (1993-2013). *Proceedings - 2017 IEEE/ACM 39th International Conference on Software Engineering Companion*, 397-398.
- Moro, S., Alturas, B., Esmerado, J., Costa, C. (2017). Research trends in CISTI's unveiled through text mining. *Iberian Conference on Information Systems and Technologies*, Article number 7975765.
- Osterwalder, A. (2004). The Business Model Ontology – A proposition in A Design Science Approach. PhD thesis, University of Lausanne.
- Porter, M. (1996). "What Is Strategy?" *Harvard Business Review*, 74(6), 61-78.
- Ramage, D., Manning, C.D., Dumais, S (2011). Partially labeled topic models for interpretable text mining. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 457-465.
- SAS Enterprise Miner – SEMMA. SAS Institute 2014 [En línea]. <http://support.sas.com/documentation/cdl/en/emcs/66392/HTML/default/viewer.htm#n0pejm83csbj4n1xueveo2uoujy.htm>. Consulta: 15/08/18.
- Shafique, U. & Qaiser, H. (2014). A Comparative Study of Data Mining Process Models (KDD, CRISP-DM and SEMMA). *International Journal of Innovation and Scientific Research*, 12(1), 217-222.
- Shams, M., Baraani-Dastjerdi, A. (2017). Enriched LDA (ELDA): Combination of latent Dirichlet allocation with word co-occurrence analysis for aspect extraction. *Expert Systems with Applications*, 80, 136-146.
- Shearer, C. (2000). The CRISP-DM Model: The New Blueprint for Data Mining. *Journal of Data Warehousing*, 5, 13-22.
- Wang, C., Blei, D. (2011). Collaborative topic modeling for recommending scientific articles. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 448-456.

Yao, L., Zhang, Y., Chen, Q., Qian, H., Wei, B., Hu, Z. (2017). Mining coherent topics in documents using word embeddings and large-scale text data. *Engineering Applications of Artificial Intelligence*, 64, 432-439.