



**UNIVERSIDAD DE CHILE  
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS  
DEPARTAMENTO DE INGENIERIA INDUSTRIAL**

**DISEÑO E IMPLEMENTACIÓN DE MÉTRICAS E INDICADORES PARA APOYAR LA  
GESTIÓN DE LA PARTICIPACIÓN EN UNA COMUNIDAD VIRTUAL DE PRÁCTICA.**

**MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL**

**FELIPE IGNACIO MUÑOZ CÁRCAMO**

**PROFESOR GUÍA:  
SEBASTIÁN RÍOS PÉREZ**

**MIEMBROS DE LA COMISIÓN:  
FELIPE AGUILERA VALENZUELA  
CARLOS REVECO DÍAZ**

**SANTIAGO DE CHILE  
MAYO 2012**

RESUMEN DE LA MEMORIA  
PARA OPTAR AL TÍTULO DE  
INGENIERO CIVIL INDUSTRIAL  
POR: FELIPE MUÑOZ CÁRCAMO  
FECHA: 07/05/2012  
PROF. GUÍA: SR. SEBASTIÁN RÍOS

## **DISEÑO E IMPLEMENTACIÓN DE MÉTRICAS E INDICADORES PARA APOYAR LA GESTIÓN DE LA PARTICIPACIÓN EN UNA COMUNIDAD VIRTUAL DE PRACTICA**

El proyecto tiene por objetivo el diseño de métricas e indicadores que permitan hacer un seguimiento y control de la participación de los usuarios de una Comunidad Virtual de Práctica para apoyar la gestión de los administradores de la comunidad.

Una empresa salmonera ha implementado una Comunidad Virtual de Práctica para poder reunir a los trabajadores sin importar las distancias geográficas que los separan y compartir conocimientos y experiencias. La comunidad en estudio tiene por objetivo servir como plataforma para el surgimiento de proyectos de innovación con enfoque en la producción, por medio del aporte e interacción de todos sus miembros.

Los administradores de la comunidad han expresado que el uso de la plataforma ha resultado en mejoras para la empresa. Sin embargo, aún falta explotar todo el potencial que tiene, ya que uno de los objetivos es que la totalidad de los miembros participen y en la actualidad, según los administradores, unos pocos usuarios concentran la mayor parte de las interacciones. Además de esto, no cuentan con herramientas objetivas para hacer gestión, sino que solo su apreciación personal.

En consecuencia, se plantea como solución el desarrollo de métricas e indicadores que permitan conocer formal y objetivamente lo que ocurre dentro de la comunidad. Para diseñarla se utilizó la teoría de métricas e indicadores para conocer los niveles de participación en las secciones de la plataforma y mediante el Análisis de Redes Sociales se pudo conocer la evolución de la comunidad pero desde la teoría de grafos. Para la implementación se construyó un prototipo de Data Mart para almacenar las métricas diseñadas y se construyeron cubos OLAP para visualizarlas. Para presentar los indicadores se generaron reportes para ser consultados mediante el software JasperReports Server.

Se evaluó cualitativamente la solución mediante una encuesta efectuada a tres administradores de la comunidad y los resultados obtenidos muestran que el proyecto fue exitoso, ya que el diseño permite conocer los niveles de participación y mediante esta información se puede mejorar la gestión en la comunidad. De esto se desprende que no basta con instaurar una comunidad, sino que se requiere de herramientas adicionales para apoyar la labor de los administradores y así puedan desarrollar estrategias que permitan cumplir los objetivos de estas comunidades.

# Índice de Contenidos

1	Introducción .....	1
1.1.	Antecedentes Generales .....	1
1.2.	Justificación del trabajo .....	2
1.3.	Hipótesis de trabajo.....	4
1.4.	Objetivos .....	4
1.5.	Resultados Esperados .....	5
1.6.	Alcances del trabajo .....	5
2	Marco Conceptual.....	6
2.1.	Comunidades .....	6
2.2.	Capital Social .....	9
2.3.	Métricas e Indicadores en una Comunidad Virtual.....	9
2.4.	Repositorios de datos.....	14
2.5.	Business Intelligence.....	17
3	Metodología .....	19
3.1.	Metodología CRISP-DM.....	19
3.2.	Metodología para Diseño de Métricas e Indicadores .....	20
3.3.	Metodología para Diseño de Repositorio Centralizado de Datos .....	21
3.4.	Metodología para Evaluación Cualitativa .....	22
4	Diseño de Métricas e Indicadores.....	24
4.1	Descripción de la comunidad .....	24
4.2	Análisis de la Red Social .....	25
4.3	Diseño de Métricas e Indicadores de Participación.....	29
5	Construcción de Repositorio de Datos.....	38
5.1	Base de datos de la plataforma.....	38
5.2	Diseño de Data Mart.....	39
5.3	Proceso de Extracción, Transformación y Carga .....	41
6	Implementación de Métricas e Indicadores de Participación .....	45
6.1	Construcción de cubos OLAP .....	45
6.2	Construcción de Reportes .....	48
7	Diseño y Aplicación de Evaluación Cualitativa.....	52

7.1	Diseño de encuesta.....	52
7.2	Resultados .....	55
8	Conclusión .....	58
	Bibliografía.....	62
	Anexos.....	66
A	Descripción de la Comunidad Virtual de Práctica PIKS.....	66
B	Ejemplo esquema XML de Cubo OLAP.....	69
C	Evaluación Cuantitativa y Resultados.....	70

# Índice de Figuras

Figura 1: Grafo dirigido y no dirigido, respectivamente.....	10
Figura 2: Ejemplo de modelo estrella.....	16
Figura 3: Ejemplo de modelo constelación. ....	16
Figura 4: Evolución del grafo de la red. ....	27
Figura 5: Grafo de la comunidad según área de pertenencia. ....	29
Figura 6: Ejemplo de cálculo de Métrica PST. ....	34
Figura 7: Ejemplo de Métrica EA. ....	34
Figura 8: Modelo Relacional del Data Mart.....	40
Figura 9: Proceso ETL para las dimensiones del Data Mart.....	41
Figura 10: Proceso ETL para tablas de hechos de sección Noticias. ....	43
Figura 11: Estructura del cubo de publicaciones de la sección Noticias.....	46
Figura 12: Vista de cubo de publicaciones en sección Noticias 1.....	47
Figura 13: Vista de cubo de publicaciones en sección Noticias 2.....	48
Figura 14: Diseño de reporte mensual para las 3 secciones. ....	49
Figura 15: Ejemplo de reporte.....	50
Figura 16: Interfaz de JasperReports Server. ....	51
Figura 17: Conceptos y subconceptos a medir en la encuesta.....	53

# Índice de Tablas

Tabla 1: Diferencias entre Data Warehouse y Sistema operacional.....	15
Tabla 2: Diferencias entre Data Warehouse y Data Mart.....	15
Tabla 3: Evolución de métricas sociales. ....	27
Tabla 4: Resultado aplicación algoritmo HITS .....	28
Tabla 5: Métricas e indicadores para publicaciones 1. ....	31
Tabla 6: Métricas e indicadores para publicaciones 2. ....	31
Tabla 7: Métricas e indicadores para estados de ánimo.....	33
Tabla 8: Ejemplo de cálculo de indicador IPSTE. ....	35
Tabla 9: Ejemplo de cálculo de indicadores IPST, VPST y VIPST. ....	35
Tabla 10: Tablas de la base de datos para la extracción de métricas. ....	39
Tabla 11: Resumen de tablas del Data Mart diseñado. ....	40
Tabla 12: Conceptos y preguntas de la encuesta.....	54
Tabla 13: Resultados de la encuesta.....	55

# Capítulo 1

## Introducción

En este capítulo se presentan los antecedentes generales del proyecto y el problema que se resolverá, junto a los objetivos e hipótesis del trabajo.

### 1.1. Antecedentes Generales

Una Comunidad de Práctica corresponde a un grupo de personas que comparten preocupaciones, problemas o la pasión por temas específicos y que mediante la interacción permanente de los miembros buscan profundizar sus conocimientos y/o pericia [1]. Cuando las interacciones entre los miembros de la comunidad ocurren a través de medios electrónicos, como Internet, se habla de Comunidades Virtuales [2].

Por otra parte un Sitio Web Social es un Sitio Web que otorga la plataforma para que las personas interactúen virtualmente, ya que permite a los usuarios de la plataforma intercambiar contenidos por medio de Internet [3].

PIKS<sup>1</sup> es un Sitio Web Social, ya que es una plataforma que permite el intercambio de contenidos entre sus usuarios. Se define como *“una plataforma web que facilita las conversaciones de Innovación y el contacto de los Equipos de Trabajo...”*. Por lo tanto, los usuarios del sitio se convierten en miembros de una Comunidad Virtual de Práctica, es decir, una comunidad que se forma por las interacciones virtuales de sus miembros y por compartir contenidos específicos sobre algún tema.

Como lo explica uno de sus administradores, fue desarrollada con el objetivo de ser un apoyo en la coordinación de los trabajadores dentro de una empresa y así desarrollar proyectos de innovación con el aporte de toda la comunidad, proyectos que una vez iniciados en la plataforma sean llevados a la realidad. Con esto se busca una mejor comunicación y gestión en las organizaciones.

---

<sup>1</sup> <http://www.piks.cl>

De esta manera la plataforma representa un espacio para generar interacción entre los miembros, los cuales podrán contribuir a la comunidad sin importar las distancias geográficas.

PIKS es ofrecido a las organizaciones como un servicio, el software es adaptado según los requerimientos del cliente y acorde a esto se crea la comunidad. Dado que se vende como un servicio, la mantención de las comunidades recae en PIKS, quienes hacen la labor de administrarla y fomentar su uso.

FRIOSUR<sup>2</sup> es una de las organizaciones que ha adquirido el servicio. Esta empresa se dedica a la pesca y cultivo. Posee plantas y centros productivos en diversas zonas de Chile, desde Santiago hasta Puerto Chacabuco.

Por lo tanto la implementación de la comunidad virtual en la organización significa poder trabajar en equipos a pesar de la separación geográfica que pueda existir entre los miembros. Con esto se busca que el conocimiento y la pericia de los usuarios puedan ser transmitidos hacia otros mediante opiniones publicadas en la plataforma y de esta manera promover la innovación en la empresa.

Sin embargo, los administradores – tanto de *PIKS* como de la empresa *FRIOSUR* – no se encuentran conformes con el desarrollo y aporte que ha significado para la organización. Es por esto que en el presente trabajo se tiene como propósito realizar una contribución que permita apoyar la labor de administrar y fomentar el crecimiento de la comunidad.

## 1.2. Justificación del trabajo

El concepto de Capital Social se refiere al potencial que tiene una organización para generar valor. Dentro de este se encuentra el Capital Social Relacional, el que consiste en conocer las interacciones que ocurren dentro de la organización [4].

Parte del objetivo de la implementación de la comunidad se basa en explotar el Capital Social de la organización que conforman sus trabajadores, por lo que el uso de la plataforma está restringido solo a los trabajadores de la empresa, o sea, es una comunidad cerrada.

Estas interacciones son de gran relevancia, pues, permiten transferir parte del conocimiento tácito entre los miembros de la comunidad. El conocimiento tácito se refiere a la experiencia, la que no puede ser transferida mediante, por ejemplo, documentos, como si es posible hacerlo en el caso del conocimiento explícito. Sin

---

<sup>2</sup> <http://www.friosur.cl/>

embargo, si es viable de compartir mediante interacciones informales como lo son las conversaciones entre personas, en las cuales es más simple profundizar el entendimiento de temas complejos [1].

Dado lo anterior, se tiene que este tipo de comunidades resultan útiles para las organizaciones, ya que permiten combinar aspectos implícitos y explícitos del conocimiento al permitir que los usuarios interactúen, pues así pueden transmitirles a otros sus conocimientos ligados a la experiencia y cultura organizacional.

Es por esto que potenciar la comunidad es de importancia para los administradores, ya que lograr los objetivos por los cuales se creó la comunidad – que es explotar el Capital Social – se requiere que los usuarios participen, es decir, que interactúen con otros usuarios o que den su opinión sobre los temas que se planteen, pues, se necesita de otros para poder complementar y desarrollar la pericia de cada uno [1].

Se deben cultivar las comunidades activa y sistemáticamente para que sus miembros y la propia comunidad se beneficie, ya que sin una intención de desarrollarlas, el crecimiento de estas dependerá solamente del tiempo libre de los miembros y por lo tanto la participación probablemente será irregular [1].

PIKS en la empresa FRIOSUR ha significado mejoras y ha implicado el desarrollo de nuevos proyectos relacionados con la innovación en la producción de la empresa. A pesar de esto los encargados de la comunidad expresan que falta participación de los usuarios – el 25%<sup>3</sup> de los usuarios concentran la mayor parte de las interacciones, es decir, de los comentarios, temas, opiniones o nuevos proyectos que se plantean - lo que lleva a que no se aproveche el potencial que representa la plataforma y por lo tanto la generación de propuestas innovadoras sea baja. Si bien esta porción de usuarios que participa mayoritariamente puede ser la esperada en este tipo de comunidades, no es suficiente para los administradores, ya que con ese nivel de participación no se está logrando el objetivo de la comunidad.

Además, se tiene que no existe un mecanismo formal de medir los niveles de participación o algún aspecto similar. Las mediciones que se realizan son solamente en base a las observaciones que los administradores hacen.

De esta manera, el propósito de la comunidad – que es lograr una constante participación de todos los usuarios – no se está cumpliendo en la forma en que los administradores desean.

Para mejorar esto se propone la construcción de métricas e indicadores que muestren la situación de la comunidad, es decir, que midan cuantitativamente la

---

<sup>3</sup> Según observaciones de los administradores.

participación de los usuarios por medio de las interacciones que entre ellos exista. Con esto los administradores podrán tomar decisiones y adoptar estrategias en función de la información que le brinde la solución y así hacer una mejor gestión de la participación de la comunidad.

### 1.3. Hipótesis de trabajo

Dado los antecedentes anteriores, la hipótesis del trabajo es la siguiente:

La gestión de la participación puede mejorarse a través de la mejor visibilidad que aportaría la creación de métricas e indicadores de los niveles de participación y que permitiría tomar decisiones a los administradores de la comunidad u otros miembros clave.

### 1.4. Objetivos

#### 1.4.1 Objetivo General

Diseñar e implementar métricas e indicadores para asegurar el seguimiento y control de la participación de los usuarios de una Comunidad Virtual de Práctica para apoyar la gestión de los administradores de la comunidad.

#### 1.4.2 Objetivos Específicos

1. Realizar una revisión literaria para establecer un marco conceptual sobre: Comunidades Virtuales de Práctica y participación de usuarios; Capital Social; Análisis de Redes Sociales; Teoría de Métricas e indicadores; y Repositorios de datos y tecnologías para construirlos.
2. Diseñar, implementar y evaluar métricas e indicadores para el seguimiento y control de una Comunidad Virtual de Práctica.
3. Diseñar y construir un Repositorio Centralizado de Datos para almacenar las métricas desarrolladas.
4. Evaluar cualitativamente la solución diseñada.

## 1.5. Resultados Esperados

1. Una metodología para la creación de métricas e indicadores para la comunidad virtual PIKS.
2. Un set de métricas e indicadores para medir la participación de los usuarios.
3. Un Repositorio Centralizado de Datos.
4. Un proceso de Extracción, Transformación y Carga (ETL) de los datos de la comunidad.
5. Una encuesta para medir cualitativamente la apreciación de los administradores sobre la solución diseñada y el resultado que se obtenga de la aplicación de esta. Con esto se concluirá sobre la hipótesis del trabajo.

## 1.6. Alcances del trabajo

Desarrollar métricas e indicadores para los módulos de interacción existentes en la actualidad en la plataforma y en base a los datos de la comunidad virtual de la empresa FRÍOSUR, es decir, no se trabajará con la información de otras organizaciones que hayan implementado PIKS.

Las interacciones se medirán cuantitativamente. No se considerará la información cualitativa que poseen los datos, o sea, lo que se refiere al contenido que puedan tener las publicaciones. Esto se sustenta en el hecho de que – como se observa – las interacciones que los usuarios realizan son con el objetivo de contribuir a la comunidad y no de desviar el sentido de las conversaciones. Por otra parte, las interacciones que se medirán corresponden a la participación activa, es decir, la que se refiere a los *posts* que los usuarios realizan, ya que no hay weblogs disponibles para medir la participación pasiva.

Se contempla el desarrollo de un prototipo de repositorio centralizado de datos, el que consistirá en un Data Mart en donde se almacenarán los datos de la comunidad extraídos desde la base de datos que contiene las interacciones de los usuarios en los diversos foros de la plataforma. Este repositorio será la base sobre la que se realizarán los cálculos necesarios para hacer el seguimiento y control de la participación de la comunidad.

Finalmente la evaluación y validación de la solución se hará en base a la apreciación de los administradores, es decir, será cualitativa y según esto se concluirá sobre la hipótesis y la dirección en que deberá apuntar el trabajo futuro.

## Capítulo 2

# Marco Conceptual

Este capítulo tiene como propósito introducir a los conceptos involucrados en la realización del trabajo, por medio de una revisión literaria de los tópicos necesarios para abordar el problema planteado en el capítulo anterior y diseñar una solución.

### 2.1. Comunidades

Una comunidad puede definirse como un grupo de personas que comparten temas en común, como intereses, experiencias o necesidades. De esta manera, los miembros de la comunidad forman vínculos al relacionarse socialmente entre ellos. Este intercambio de recursos que se produce dentro de la comunidad genera un sentido de pertenencia de los miembros con la comunidad [15]. Existen varios tipos de comunidades, destacando principalmente las de interés, las de práctica y las de propósito [15][16]:

- Las comunidades de interés reúnen a personas con un interés común por un determinado tema. El objetivo de los miembros para formar parte de estas es aumentar sus conocimientos sobre el tema mediante el intercambio de información [37][38].
- Las comunidades de práctica corresponden a las que sus miembros comparten una profesión o actividad. Sus miembros tienen en común alguna preocupación, problema o pasión por un tema particular. El objetivo de la comunidad es que las personas profundicen sus conocimientos o mejoren sus prácticas mediante el intercambio de información y experiencias [1][15][37].
- Las comunidades de propósito consisten en las que sus miembros comparten un objetivo acerca de un interés en común y por lo tanto la comunidad tiene un estricto enfoque en ese interés. Sus miembros se caracterizan por poseer distintos niveles de conocimientos y puntos de vistas [37].

### 2.1.1 Comunidad de Práctica

Existen varias publicaciones que entregan caracterizaciones de una comunidad de práctica, dentro de estas se destaca la propuesta por E. Wenger [1]. El autor propone que estas comunidades están compuestas por una combinación de tres elementos fundamentales: dominio, comunidad y práctica.

- El dominio se define como la base de conocimientos y entendimientos sobre la que reside la comunidad. Permite formar un sentido de identidad común entre sus miembros, con lo que se reafirma el propósito y valor de la comunidad hacia sus miembros y de esta manera fomentar su participación.
- La comunidad es el conjunto de personas que permite generar aprendizaje. Una comunidad sólida en la que cada persona tiene un rol, fomenta las interacciones y las relaciones entre sus miembros. Con esto se genera un sentido de pertenencia que permite el aprendizaje de los miembros.
- La práctica se refiere al conjunto de ideas, información, herramientas, acciones, entre otros, que los miembros comparten con la comunidad. Es el conocimiento específico que la comunidad busca desarrollar, compartir y mantener.

La participación de los miembros en estas comunidades puede analizarse definiendo tres niveles de pertenencia: núcleo, activos y periféricos [1].

- En el núcleo se encuentran los miembros que participan activamente de las discusiones que se generan y guían el aprendizaje de la comunidad, entre otros aspectos. Estos miembros se convierten en líderes de la comunidad y ayudan a administrarla. El tamaño de este grupo es del 10 a 15 por ciento del total de miembros.
- En el grupo de activos se encuentran quienes participan ocasionalmente y en menor medida que el grupo anterior. Se conforma por el 15 a 20 por ciento de la comunidad.
- El tercer grupo corresponde a quienes rara vez participan en las discusiones, ya que se dedican principalmente a observar las interacciones que los otros miembros generan. Es el grupo de mayor tamaño.

Los administradores de la comunidad deben adoptar estrategias que incentiven a que los miembros se muevan de nivel de participación, en particular, que los miembros de la periferia participen de discusiones y dejen de ser meros observadores trasladándose a los grupos de activos o incluso al núcleo. Este procedimiento se denomina Participación Periférica Legítima [36] y describe como los miembros que se incorporan a la comunidad se convierten en expertos.

De acuerdo a la Participación Periférica Legítima los miembros nuevos deben realizar tareas simples pero que sean importantes para el objetivo de la comunidad y de esta manera vayan aprendiendo y familiarizándose con las prácticas de la comunidad, pero, además de esto, se requiere tener contacto con los expertos más de allá de solamente observarlos, es decir, para que este proceso de aprendizaje resulte se debe ser un participante activo de las prácticas sociales de la comunidad. Por lo tanto, es más que un proceso de adquisición de conocimientos y habilidades, es un proceso de participación social que crea una identidad en relación a las características de la comunidad.

En la medida que los miembros logren una Participación Periférica Legítima, la comunidad tendrá éxito, ya que existirá un dinamismo en los miembros que producirá que la participación no se concentre en unos pocos y por lo tanto el desarrollo de los objetivos de la comunidad no recaiga únicamente en los líderes o administradores.

### 2.1.2 Comunidad Virtual de Práctica

Cuando la comunidad existe a través de Internet, es decir, cuando los miembros interactúan en la comunidad mediante computadores se habla de una comunidad virtual. En particular, las Comunidades Virtuales de Práctica (VCoP por sus siglas en inglés) corresponden a comunidades en donde los miembros intercambian conocimientos y experiencias por medio de las funciones para interactuar que tengan a su disposición en la plataforma y mediante este esfuerzo colaborativo se pretende conseguir que los miembros mejoren sus prácticas, por ejemplo, dentro de la organización a la cual pertenece la comunidad [16].

Se pueden definir dos tipos de participación en una comunidad virtual: activa y pasiva [39]. La participación activa son las interacciones que se generan entre los miembros mediante los *posts* que publican en la comunidad. Mientras que la pasiva se refiere a la participación que se genera al observar lo que ocurre en la comunidad sin emitir opiniones en las discusiones.

En una comunidad virtual, la participación activa la concentrarán en mayor parte o en su totalidad los miembros que pertenecen al núcleo y al grupo de activos y la participación pasiva estará presente en los tres grupos, sin embargo, los miembros de la periferia, al no *postear*, su participación será principalmente pasiva.

Parte de la labor de desarrollar una VCoP debe considerar los problemas que los miembros tienen y que dificultan su participación en ella. Por ejemplo, es posible rescatar de diversos estudios [1][5][6][7], que los usuarios no participan por miedo a las críticas, porque sienten que sus aportes no serán relevantes para otros, por no tener la autoridad suficiente, por falta de tiempo, conocimiento o experiencia, entre otros. Y entonces, estos miembros, no se mueven de la periferia, no logran el aprendizaje que

se espera de ellos y por lo tanto la comunidad no consigue el dinamismo y participación que se necesita.

De lo anterior se concluye que es de importancia para los administradores de la comunidad desarrollar una labor sistemática que cultive la comunidad e incentive a la participación y compromiso, ya que si la participación solo depende de los miembros, entonces probablemente esta será irregular [1].

## 2.2. Capital Social

El concepto de Capital Social se refiere a la capacidad de generar valor de las personas dependiendo de la relación que entre ellas se forme [4] [10]. Es por esto que el concepto toma mayor fuerza al tratarse de organizaciones, debido a las diversas relaciones entre las personas que se producen en ella y al potencial productivo que se puede obtener de estas relaciones.

El Capital Social puede descomponerse en cuatro componentes [4]:

- El Capital Humano: *“La capacidad de producir valor de las personas que conforman una organización social”*.
- El Capital Relacional: *“Aquel valor que se genera como resultados de las interacciones entre las personas que constituyen la organización”*.
- El Capital Ambiental: *“Aquel valor que se genera como resultado del ambiente en que se producen las interacciones”*.
- El Capital Direccional: *“Aquel valor que surge del alineamiento e involucramiento de las personas con un proyecto común”*.

Parte del objetivo de haber construido la plataforma PIKS es desarrollar el capital direccional de una organización, es decir, que los miembros compartan experiencias y conocimientos para resolver quiebres y por medio de este trabajo en conjunto se generen innovaciones que se lleven a la realidad para, por ejemplo, mejorar procesos productivos de la empresa.

## 2.3. Métricas e Indicadores en una Comunidad Virtual.

Una Red Social consiste básicamente en un grupo de personas que se conectan unas con otras mediante un conjunto de relaciones con significado social. De esta manera, se tiene que una Comunidad Virtual de Práctica puede ser igualmente tratada como una Red Social, ya que esta solo se diferencia en que los miembros se relacionan

unos con otros mediante computadores conectados a una red, como lo puede ser Internet [8].

Dado esto, para medir los niveles de participación de una VCoP se requiere de los datos de la participación activa, para lo que se necesitan los datos de los *posts*, y de la participación pasiva, para lo que se necesitan los weblogs, que corresponden a los datos de las actividades realizadas por los usuarios en cada sesión.

Así se tiene que, las Comunidades Virtuales de Práctica al poder ser modeladas como una Red Social, es posible evaluarlas cuantitativamente [9]. Esto permite poder realizar mejoras en la plataforma que conduzcan a desarrollar la comunidad en función de los objetivos que los administradores se propongan.

Por otra parte, también se puede obtener información sobre la participación de una Comunidad Virtual desde la teoría de métricas e indicadores, lo que implica definir métricas e indicadores en base a esta teoría y según las características que se deseen conocer de la comunidad.

### 2.3.1 Análisis de Redes Sociales.

Una Comunidad Virtual puede ser interpretada como una Red Social, la que a la vez es modelada como un Grafo, en otras palabras, desde la teoría de grafos es posible extrapolar métricas a una comunidad virtual y cuantificar características de la comunidad para poder realizar un análisis de ésta.

Un Grafo es un conjunto de nodos que se conectan entre ellos mediante arcos, estos grafos pueden ser dirigidos o no dirigidos, lo que dependerá de si los arcos que inciden en los nodos tienen sentido o no respecto de la relación entre cada par de nodos [9].

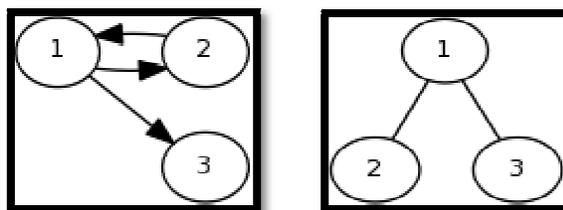


Figura 1: Grafo dirigido y no dirigido, respectivamente.  
Fuente: Elaboración Propia.

Estos conceptos de la teoría de grafos pueden ser llevados a una red social - la que representa una comunidad virtual – pues, entrega un vocabulario para expresar las propiedades de una estructura social y además permite cuantificar y medir estas propiedades [9].

Como se dijo anteriormente, el grafo se define por dos conjuntos, uno de nodos y otro de arcos. En una red social, el conjunto de nodos representará al conjunto de actores que participa en la red, los que interactúan entre ellos, mientras que los arcos representan al conjunto de relaciones que entre ellos se forma al interactuar. Entonces, la red social la forman los actores o nodos y las relaciones entre ellos representadas por los arcos [9].

Una vez modelada una Comunidad Virtual de Práctica como una Red Social, en donde los usuarios serán los actores y las relaciones serán las interacciones que en la plataforma ocurran entre los usuarios (mensajes, *posts*, entre otros), es posible utilizar la teoría de grafos para obtener métricas sociales, de aquí surge el Análisis de Redes Sociales (SNA por sus siglas en inglés) que permite cuantificar las interacciones que ocurren en la comunidad [9] [10] [11] y acorde a esto obtener conclusiones que puedan conducir a un mejoramiento. Por ejemplo, con las herramientas del SNA es posible identificar los líderes de la comunidad, los niveles de participación, encontrar redes de cooperación, entre otros aspectos [9] [11].

#### Métricas sociales<sup>4</sup>:

- **Grado:** Denotado por  $d(n_i)$ , representa el número de arcos que inciden en un determinado nodo, es decir, es el número de nodos adyacentes que posee. Si un nodo tiene grado nulo, entonces se dice que está aislado, mientras que el máximo grado de un nodo será el número total de nodos que existan en la red. Además, el grado promedio viene dado por:

$$\bar{d} = \frac{2L}{g}$$

Donde  $L$  es la cantidad de arcos existentes y  $g$  el número de nodos de la red.

- **Densidad:** Proporción entre la cantidad de arcos existentes en la red y el número máximo que pueden existir. El número máximo de arcos posibles es  $g(g - 1)/2$ , donde  $g$  es el número de nodos en la red. Entonces, se tiene que la densidad viene dada por:

$$\Delta = \frac{2L}{g(g - 1)}$$

Donde  $L$  representa la cantidad de arcos existentes.

Si la red no tiene arcos, la densidad será 0, mientras que si existen todos los posibles arcos, la densidad será 1.

- **Diámetro:** Distancia más larga existente dentro de la red. La distancia entre dos nodos se define como la menor cantidad de arcos que permiten conectarlos. Por lo tanto, el diámetro representa la distancia del par nodos que se encuentran

<sup>4</sup> Estas definiciones se basan principalmente en [9]

más lejanos entre ellos dentro del grafo. Si el diámetro es infinito, puede significar que existan infinitos nodos o que el grafo no es conexo.

- **Centralidad:** Indica el poder social de un nodo basado en cómo se conecta al resto de la red.
- **Cohesión:** El grado en que los actores están conectados entre sí por vínculos cohesivos.

## Configuración de red

Una red social puede estar configurada de diversas formas, lo que depende de a quienes van dirigidas las respuestas a los tópicos de la comunidad [27]:

- *Creator Reply:* Corresponde a cuando los comentarios que se publican en un hilo de conversación están dirigidos al usuario que lo inició. Esta configuración posee la menor densidad de arcos.
- *Last Reply:* Cuando los comentarios que se publican en un hilo están dirigidos al último usuario que hizo un *post*. Representa una densidad de arcos mediana.
- *All previous Reply:* Cada respuesta en el hilo estará dirigida a todos los usuarios que hayan publicado un *post* con anterioridad en el mismo hilo. Esta configuración representa la mayor densidad.

## Algoritmo HITS

HITS (*Hyperlink-Induced Topic Search*) [28] [30] es un algoritmo que fue creado para valorar sitios web mediante el cálculo iterativo de dos indicadores. El objetivo es poder clasificar los sitios web en base a cuantos “buenos” sitios lo apuntan y a cuantos buenos sitios “apunta” él.

Para esto se definieron dos indicadores: *hub* y *authority*. *Authority* corresponde a sitios web que contienen información relevante y buena sobre algún tema, mientras que *hub* corresponde a sitios que poseen links hacia *authorities*. Dado esto, se define como buen *authority* a un sitio que sea apuntado desde buenos *hubs* y un buen *hub* será un sitio que apunte a buenos *authorities*.

El algoritmo funciona calculando iterativamente un puntaje *hub* y un puntaje *authority* para cada sitio web y genera una clasificación de estos.

Si bien este algoritmo fue creado para rankear sitios web, la misma lógica puede aplicarse a una red social simulando que cada nodo representa un sitio web y cada arco representa un link desde un sitio a otro.

Entonces, un buen nodo *authority* será el que es apuntado desde buenos nodos *hubs* y un buen nodo *hub* será el que apunta a buenos nodos *authorities*. Una interpretación que puede darse a estos indicadores es, por ejemplo, que un nodo con *hub* alto será el que apunta a muchos nodos *authorities*, los que a su vez son apuntados por muchos otros nodos, mientras que un nodo con *authority* alto será el que es apuntado por muchos nodos *hubs*, los que a su vez apuntan a muchos otros nodos. Entonces, un nodo con *authority* alto en una red social representará a un miembro que es buena fuente de información y que contesta a otros miembros, o sea, en algunos casos podría ser considerado un experto o líder. Mientras que un nodo con puntaje *hub* alto puede interpretarse como miembro que motiva al resto, ya que constantemente está generando contenido de acuerdo a los propósitos de la comunidad [27]

## Gephi<sup>5</sup>

Es un software para realizar Análisis de Redes Sociales que posee una plataforma interactiva que permite la visualización de las redes. Permite calcular fácilmente las propiedades de una red como el grado, diámetro, densidad, entre otras métricas, y aplicar algoritmos como *HITS*. Además, posee opciones gráficas de visualización como colorear, variar tamaños de nodos, layouts, entre otros, lo que permite comprender mejor la red.

### 2.3.2 Teoría de Métricas e Indicadores

Antes de desarrollar un proyecto que involucre el diseño de métricas e indicadores se deben definir algunos conceptos que permitan realizar un trabajo formal, ordenado y que evite caer en redundancias o incongruencias. Con estos conceptos se podrá desarrollar una metodología<sup>6</sup> para desarrollar el diseño. Estas definiciones se basan en el trabajo de L. Olsina [12]:

- **Necesidad de Información:** Corresponde a definir qué tipo de información es la que se quiere obtener como resultado del proyecto.
- **Concepto medible:** Es lo que relaciona a la necesidad de información con los atributos medibles del problema.
- **Atributo medible:** Una propiedad que dependerá de la naturaleza de cada problema y que es medible o cuantificable.
- **Métrica:** Es la medición formal de un atributo y que entrega la medida del atributo, esta medida es el valor numérico o nominal que permite comparar objetivamente los atributos del problema. Por lo tanto, la métrica representa una abstracción y formalización de lo que se desea analizar.

<sup>5</sup> <https://gephi.org/>

<sup>6</sup> En esta sección se presentan brevemente algunos conceptos de la teoría, la metodología se encuentra explicada en mayor detalle en la sección 3.2

- **Indicador Elemental:** A partir de las métricas es posible diseñar indicadores elementales, los que se componen de combinaciones de una o más métricas y permiten conocer con mayor profundidad lo que ocurre en el problema o poder evaluar el sistema en estudio.
- **Indicador Global:** Se compone de indicadores elementales y sirve para evaluar desde una perspectiva más general el problema.

## 2.4. Repositorios de datos

Para almacenar los datos de un sistema existen diversas formas de hacerlo y diversos tipos de repositorios. A continuación se presentan dos estructuras, Data Warehouse y Data Mart, las que su construcción dependerá de los requerimientos que el proyecto tenga. Posteriormente se muestran las formas de modelar los datos dentro de un repositorio, ya sea Modelamiento multidimensional o Modelamiento Estrella. Luego se explica el proceso para la Extracción, Transformación y Carga de los datos en el repositorio. Y finalmente como visualizar los datos y poder obtener la información requerida.

### 2.4.1 Data Warehouse

Data Warehouse es un tipo de arquitectura para el almacenamiento de información. Se ha definido de diversas maneras: “Una colección de datos para dar soporte a la toma de decisiones de gestión que está orientada al negocio, integrada, no volátil y que varía en el tiempo” [17]; o como “Una colección de datos en forma de base de datos que almacena y ordena información que se extrae directamente de sistemas operacionales y datos externos” [18].

Las principales características de un Data Warehouse, mencionadas anteriormente, son las siguientes [13] [17]:

- Orientada al negocio/tema: El Data Warehouse se enfoca en temas importantes para el negocio o compañía.
- Integrado: Se alimenta de varias fuentes de datos. El Data Warehouse se va construyendo a medida que se van integrando múltiples fuentes de datos.
- No volátil: Los datos almacenados en el Data Warehouse rara vez deberían ser actualizados, es decir, los datos debiesen ser cargados y luego consultados por el usuario.
- No varía en el tiempo: Se refiere a mantener un registro de los datos cargados sobre el tiempo, para así tener una perspectiva histórica de los datos almacenados.

Parte del objetivo de este tipo de sistemas es que la información sea de fácil acceso y por lo tanto orientados al usuario final, lo que se logra al tener una sola fuente de datos integrada y a la granularidad, es decir, el nivel de detalle de los datos en el Data Warehouse, entre otros aspectos [17].

A continuación se presentan las principales diferencias entre este tipo de arquitectura y un sistema de bases de datos operacionales [13] [17]:

Tabla 1: Diferencias entre Data Warehouse y Sistema operacional.

Sistema Operacional	Data Warehouse
Datos actualizados constantemente. Actualizaciones no programadas y frecuentes.	Datos estáticos, rara vez modificados. Actualizaciones programadas y periódicas.
Gran nivel de detalle. Mínima redundancia	Datos agregados. Alta redundancia.
Refleja el "ahora". Soporta operaciones del día a día.	Refleja la "historia". Soporta requerimientos de información.
Muchos usuarios concurrentes.	Pocos usuarios concurrentes.
El usuario final depende de terceros para que el sistema se adapte a nuevos requerimientos.	Orientado al usuario final, el usuario es autónomo de obtener la información que requiera, por lo tanto el Data Warehouse es más flexible.

Fuente: Elaboración propia

## 2.4.2 Data Mart

Un Data Mart se asemeja a un Data Warehouse, pero a menor escala. Está diseñado para cumplir los requerimientos de información a nivel departamental y para un número limitado de usuarios [17]. El Data Mart suele satisfacer los requerimientos de información de un solo proceso de negocio [19] [20].

En la siguiente tabla se pueden observar las principales diferencias entre un Data Warehouse y un Data Mart [13]:

Tabla 2: Diferencias entre Data Warehouse y Data Mart.

Data Warehouse	Data Mart
Alcance a nivel de toda la empresa.	Alcance a nivel departamental.
Extrae datos de muchas fuentes operacionales.	Extrae datos de pocas fuentes operacionales.
Desarrollo y resultados a largo plazo	Desarrollo y resultados a corto plazo.

Fuente: Elaboración Propia.

Dadas las definiciones de Data Warehouse y Data Mart, se tienen dos enfoques para la construcción de un repositorio de datos:

- Enfoque Top Down: Diseñar el Data Warehouse para luego desplegar Data Marts que se alimentan del Data Warehouse [17].
- Enfoque Bottom Up: Se construyen Data Marts para posteriormente diseñar el Data Warehouse [18].

### 2.4.3 Modelo Estrella

El modelamiento estrella consiste en un modelamiento multidimensional pero utilizando el modelo entidad-relación, por lo tanto se requiere de un sistema de base de datos relacional. Este modelamiento se compone de en una tabla de “hechos” y varias tablas dimensionales [17].

La tabla de hechos guarda las medidas del negocio según los requerimientos de información, además, su llave primaria está compuesta por llaves foráneas que hacen referencias hacia las tablas de dimensiones. La granularidad es clave en esta tabla, ya que representa el nivel de detalle mínimo al que se quieren tener los datos. Mientras que las tablas de dimensiones almacenan las diferentes formas de representar los datos, como pueden ser el tiempo, ubicación, cliente, vendedor, entre otros. Un ejemplo de esto hay en la figura 2.

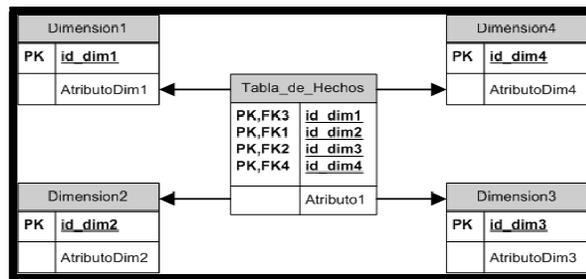


Figura 2: Ejemplo de modelo estrella.  
Fuente: Elaboración Propia.

### 2.4.4 Modelo Constelación

Este modelo es similar al modelo estrella, con la diferencia de que está compuesto de más de una tabla de hechos, las que comparten algunas o todas las dimensiones existentes y posee la restricción de que las tablas de hechos no pueden referenciar a otra tabla de hechos [31][32]. Las ventajas de este modelo por sobre el modelo estrella son que otorga una mayor flexibilidad, ya que permite reutilizar dimensiones y al poseer más de una tabla de hechos se pueden analizar aspectos relevantes con mayor facilidad.

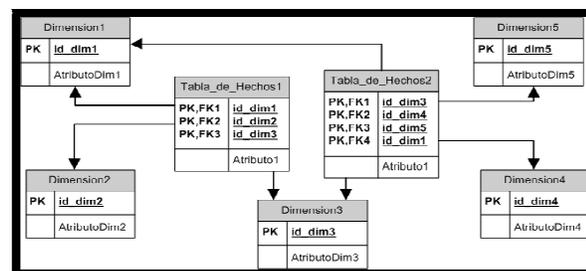


Figura 3: Ejemplo de modelo constelación.  
Fuente: Elaboración Propia.

### 2.4.6 Proceso de Extracción, Transformación y Carga

El proceso de Extracción, Transformación y Carga de datos (ETL por sus siglas en inglés) se utiliza para poblar un repositorio de datos [29] (como un Data Warehouse o un Data Mart).

Consiste en la extracción de datos de diversas fuentes operacionales, luego realizar las transformaciones necesarias para que los datos sean consistentes con el diseño del repositorio y finalmente cargar o almacenar los datos en el repositorio.

### 2.4.7 Cubos OLAP

Los cubos OLAP (On-Line Analytical Processing) representan una manera de visualizar los datos almacenados en un repositorio modelado multidimensional.

Los usuarios pueden acceder a los datos haciendo cruces entre las dimensiones existentes en el modelo y además pudiendo consultar distintos niveles de agregación.

El usuario final puede realizar las siguientes operaciones sobre el cubo [13] [22]:

- *Rollup*: Consiste en aumentar el nivel de agregación de una dimensión.
- *Drill-Down*: Disminuir el nivel de agregación.
- *Slice and Dice*: Consiste en realizar “cortes” al cubo, lo que significa aplicar filtros sobre los datos.
- *Pivot*: Realizar una reorientación del cubo sobre alguna dimensión.

## 2.5. Business Intelligence

El Business Intelligence (BI) se refiere a sistemas que combinan consolidación de información y gestión del conocimiento con herramientas analíticas para presentar la información a quienes toman decisiones [34]. En la actualidad, es reconocido como una herramienta para dirigir la efectividad e innovación en los negocios pero solo cuando los usuarios acceden a la información y la utilizan para tomar decisiones en el momento exacto [34][35].

De esta manera, se tiene que las herramientas de BI permiten extraer datos desde diversas fuentes para ser consolidadas, por ejemplo, en un Data Warehouse y posteriormente visualizar la información que se obtenga de estos datos mediante herramientas como los reportes pre-confeccionados de aplicaciones analíticas. Existen variadas suites para BI, siendo las más populares: BIRT Project de Eclipse, PENTAHO, Jaspersoft, entre otras.

A continuación se presentan las aplicaciones de la suite Jaspersoft, herramienta que se utilizará para el desarrollo del trabajo.

### 2.5.1 JasperReports Server<sup>7</sup>

Herramienta de Business Intelligence de la suite Jaspersoft que permite la visualización de información, ya sea consultando cubos OLAP o revisando reportes de cubos o bases de datos, entre otros. El usuario puede manejar las dimensiones (por ejemplo, agregando dimensiones o filtrándolas) y de esta manera adaptar las visualizaciones de la información a sus necesidades.

### 2.5.2 Jaspersoft ETL<sup>8</sup>

Es una herramienta de código libre que permite crear el proceso ETL de manera sencilla. Posee una interfaz gráfica que facilita el diseño del proceso, ocupa lenguajes Perl o Java para generar rutinas de trabajo, cabe destacar que no siempre es necesario programar, y permite integrar diversos tipos de fuentes de datos (MySQL, PostgreSQL, entre otros), entre varias otras opciones.

### 2.5.3 JasperAnalysis Schema Workbench

Herramienta que sirve para crear esquemas en formato XML de un repositorio de datos, con esto se logra generar el modelo lógico de los cubos OLAP que se construirán para posteriormente ser consultados en JasperReports Server

### 2.5.4 iReport Designer<sup>9</sup>

Software para la creación de reportes para ser visualizados en JasperReports Server. Es una herramienta gráfica para desplegar información de una base de datos, mediante consultas SQL.

---

<sup>7</sup> <http://www..com/>

<sup>8</sup> <http://www.jaspersoft.com/jasperetl>

<sup>9</sup> <http://www.jaspersoft.com/ireport-designer>

## Capítulo 3

# Metodología

La metodología para abordar el diseño de la solución será la metodología CRISP-DM en conjunto con una metodología para el Diseño de Métricas e Indicadores y una metodología para la construcción de un Repositorio de Datos.

### 3.1. Metodología CRISP-DM

La Metodología CRISP-DM consta de los siguientes pasos [13][14]:

- **Comprensión del negocio:** Consiste en determinar los objetivos del negocio, la situación actual, determinar los objetivos del proyecto y un plan de proyecto. Lo que se abordó en el Capítulo 1.
- **Comprensión de los Datos:** Consiste en describir, explorar y verificar los datos. Lo que se abordará en la sección 4.1 y en la sección 5.1.
- **Preparación de los Datos:** Consiste en seleccionar, limpiar, construir e integrar los datos. Se abordará en las secciones del capítulo 5.
- **Modelado:** Se seleccionan técnicas, se realizan pruebas y se construye el modelo. Dado que el proyecto no tiene que ver con Minería de Datos, aquí se ocupará una metodología para Diseñar Métricas e Indicadores (Sección 4.3) y una para la construcción de Repositorio de Datos (Sección 5.2).
- **Evaluación:** Consiste en evaluar los resultados y revisar el proceso. Capítulo 7.
- **Desarrollo:** Consiste en la implementación del modelo y generar planes de monitoreo para mantener el proyecto. Además de incorporar un reporte final y de cierre del proyecto.

En el capítulo 6 se presenta la implementación del proyecto y en el capítulo 8 la conclusión final con el trabajo futuro sugerido. El proyecto solo tiene alcances de prototipo, por lo tanto no se generarán planes para monitorear.

## 3.2. Metodología para Diseño de Métricas e Indicadores<sup>10</sup>

Para el Diseño de las Métricas e Indicadores, que por objetivo tienen crear un mecanismo de control y seguimiento de la Comunidad Virtual, se deberá, como primera etapa, definir la necesidad de información que tienen los administradores sobre la actividad de la comunidad. Lo que para el caso en estudio tiene que ver principalmente con la participación de los usuarios en la plataforma y por lo tanto la solución deberá contestar preguntas como por ejemplo: ¿Cuál es la participación mensual de cada usuario?, ¿Qué usuario realizó el mayor número de respuestas a otros usuarios?, ¿Qué módulo de interacción tuvo la mayor participación de los usuarios?, entre otras.

Dado lo anterior, se procederá a contestar estos requerimientos de información por medio de la creación de Métricas e Indicadores. Para esto se deben definir cuáles son los atributos medibles de la comunidad que sentarán la base para el desarrollo de la solución, por ejemplo; el número de temas creados por un usuario, el número de respuestas a otros usuarios, el número de temas creados en un módulo, entre otros.

A partir de estos atributos se diseñarán las Métricas, las que corresponden a la cuantificación de un atributo, pero que a diferencia de estos son posibles de comparar entre ellas y poder obtener conclusiones. Ejemplo de métricas son: El número de *posts* de un usuario durante un determinado período, el número de *post* creados en una categoría durante un determinado período, el número de usuarios con los cuales interactúa un determinado usuario durante un período de tiempo, entre otras.

Con las Métricas definidas será posible obtener Indicadores, ya que estos se componen de la combinación de una o más Métricas. Los Indicadores, a diferencia de las Métricas, entregan un enfoque más profundo de la situación de lo que se está midiendo. Un ejemplo de Indicador es el cociente entre el número de *post* de un usuario durante un período determinado y el número de *post* de todos los usuarios durante el mismo período. Cabe mencionar que existen indicadores elementales y globales, los elementales son los explicados anteriormente, mientras que los globales están compuestos de otros indicadores, por lo que permiten evaluar la evolución de lo que se está midiendo en términos más generales.

Por otra parte, también se deben obtener métricas sociales mediante la utilización del Análisis de Redes Sociales. Para esto se debe, en primera etapa, extraer la información de la comunidad y llevarla a la teoría de grafos. Este proceso consiste en obtener las relaciones que existen entre los usuarios y crear el grafo de la red social.

Para realizar lo anterior se utilizará el framework construido por F. Bustos en su trabajo de memoria [23]. Este framework, programado en *Java*, se conecta a una fuente

---

<sup>10</sup> Esta metodología se basa en [12]

de datos, ingresándole sus parámetros, en este caso las características de la instalación local de la base de datos de PIKS y la información de las tablas en las que se encuentran los datos de la participación de la comunidad. Configurado lo anterior, recorre los distintos foros de la plataforma y extrae las relaciones que se dan entre los usuarios mediante las publicaciones y respuestas que *postean*.

El framework entrega como *output* varios archivos con estas relaciones según las tres configuraciones de red (*creator reply*, *last reply* y *all previous reply*), los que posteriormente pueden ser cargados en algún software para procesar redes, como Gephi, y realizar los cálculos de métricas sociales que se requieran.

### 3.3. Metodología para Diseño de Repositorio Centralizado de Datos

Para el desarrollo del Repositorio Centralizado de Datos, que contendrá la información relevante para el cálculo de las Métricas e Indicadores, se utilizará un Data Mart, es decir, un repositorio diseñado para satisfacer a un número pequeño de usuarios y que almacena información desde pocas fuentes, esto pues, la solución propuesta es para ser utilizada solo por los administradores de la comunidad y además el origen de los datos extraídos proviene solo de la base de datos de las interacciones de la comunidad.

El procedimiento que se utilizará para llevar los datos que se obtienen desde la Comunidad Virtual hacia el Data Mart es el proceso *ETL*. Este proceso consiste en extraer los datos desde la fuente de estos, luego transformarlos para tener solo la información relevante para resolver el problema y finalmente realizar la carga de los datos transformados en el Data Mart.

Para la visualización de los datos cargados en el Data Mart se realizará un Modelamiento Constelación de estos. Esto consiste en representarlos mediante tablas de hechos y tablas de dimensiones. La tabla de hechos contendrá la medición de las métricas de la comunidad y las tablas de dimensiones contendrán las diversas maneras de visualizarlas.

Posteriormente se crearán cubos OLAP, los que permiten consultar las métricas almacenadas aplicando filtros y agregaciones sobre las dimensiones. Para esto se utiliza la plataforma JasperReports server.

A partir de estos cubos se generarán reportes, los que contendrán los indicadores que se definan sobre la participación de la comunidad y que podrán ser consultados mediante JasperReports server.

## 3.4. Metodología para Evaluación Cualitativa

Se diseñará una encuesta para evaluar el nivel de satisfacción de quienes utilizarán la solución propuesta. Esta encuesta permitirá registrar la apreciación de los administradores y de los demás usuarios involucrados (Gerentes, directivos u otros) en cuanto a si la solución representa una herramienta útil para realizar gestión, toma de decisiones y el fomento de la participación de los usuarios en la comunidad.

Dado lo anterior, se debe combinar una metodología para construir encuestas con conceptos como calidad de servicio, utilidad, satisfacción, percepción y expectativas.

La calidad del servicio, en este caso, se refiere a la calidad y utilidad del diseño y de la implementación de la solución, el objetivo es conocer si la percepción que tienen los usuarios acerca de la solución está al nivel de las expectativas que tenían antes de utilizarlo, a esto se le conoce como *gap* P-E [24]. En otras palabras, se deben comparar estos conceptos para comprender el nivel de satisfacción que se tiene [26] y con esto se podrá conocer si la solución cumple con lo que los administradores esperaban y por lo tanto si es útil o no para hacer gestión de la comunidad.

La metodología para diseñar encuestas o cuestionarios que se utilizará se basa en [25]. El procedimiento que se propone se denomina operacionalización que sirve para traducir la realidad al lenguaje numérico. Lo que se busca con esto es que el encuestado exprese su opinión en forma cuantitativa para medir el grado en que posee determinada variable o concepto.

Primero se debe definir el concepto que se quieren medir y que no es observable directamente (por ejemplo, calidad/utilidad). Si el concepto es complejo se deben derivar sub conceptos que sean observables y puedan medirse con mayor facilidad. Esto deberá repetirse iterativamente hasta lograr un nivel de detalle que permita medir el grado por medio de preguntas simples.

Posteriormente se deberán definir las preguntas que permitirán medir estos conceptos, para esto se recomienda presentar respuestas preestablecidas, ya que facilitan el proceso de medición y facilitan la tarea de asignar números o valores a las respuestas.

Existen diversos estándares para formular las preguntas y respuestas, por ejemplo: preguntas abiertas, preguntas cerradas de respuesta múltiple, preguntas de respuesta ordenada, preguntas de respuesta intervalar, entre varios otros. En particular, se utilizarán preguntas de respuesta ordenada y preguntas abiertas.

En resumen, el cuestionario traduce conceptos complejos a preguntas y respuestas simples, de esta manera, el encuestado podrá expresar su percepción sobre el concepto permitiendo cuantificar su opinión.

Finalmente, con los resultados de esta evaluación se concluirá sobre si la hipótesis del proyecto se cumple y por lo tanto se logró generar una herramienta que les permite a los administradores mejorar la gestión de la participación de la comunidad.

## Capítulo 4

# Diseño de Métricas e Indicadores

El propósito de este capítulo es diseñar formalmente las métricas e indicadores que permitirán conocer objetivamente los niveles de participación de los usuarios de la comunidad virtual de práctica que ha implementado la empresa FRIOSUR. Como se ha explicado en el capítulo 1, se desarrollaron con la intención de que los administradores puedan realizar un seguimiento y control de lo que ocurre y así poder hacer una mejor gestión para fomentar la participación.

Para esto, primero se explicará en qué consiste la comunidad y cómo interactúan sus usuarios, luego se analizará la situación actual de la comunidad desde el punto de vista del Análisis de Redes Sociales y finalmente se presentará el diseño de métricas e indicadores llevado a cabo.

### 4.1 Descripción de la comunidad

La comunidad PIKS tiene por objetivo fomentar la innovación en la empresa, para lograr esto se espera que los usuarios (empleados de la empresa) interactúen entre ellos en la comunidad y así compartan conocimientos que permitan resolver quiebres que los mismos usuarios detecten y den a conocer a la comunidad.

Estas interacciones entre los usuarios se denominarán participación, es decir, cuando un usuario publica algún tipo de información, como lo puede ser un quiebre detectado, y lo da a conocer a la comunidad o cuando un usuario responde, aportando datos o experiencias, a la publicación iniciada por otro usuario.

La plataforma PIKS, por la que interactúa la comunidad, se compone de diversos módulos, donde cada uno tiene un propósito definido<sup>11</sup>:

- **Noticias:** Se publican noticias de todo tipo pero siempre relacionados con la actualidad de la empresa. Cualquier usuario puede publicar y/o comentar una noticia.

---

<sup>11</sup> En el Anexo A se encuentran vistas de la plataforma para poder comprenderla mejor.

- **Bitácora:** Esta sección está destinada a publicar un quiebre que se haya detectado y que afecte a un centro de la empresa. Cualquier usuario puede publicar y/o comentar. Pero cada centro productivo de la empresa debe hacer al menos una publicación diaria en la que se plantee un quiebre. Al publicar una entrada en esta categoría el usuario debe escoger el centro al que se referirá y se espera que los demás usuarios participen dando su opinión o aportando datos al respecto, con el objetivo de compartir el conocimiento y que se logre resolver.
- **Síntesis:** Permite observar el historial de las publicaciones y el historial de los comentarios que han realizado los usuarios.
- **Iniciativas:** Cualquier usuario puede publicar una Iniciativa. La iniciativa corresponde a plantear un quiebre, pero a diferencia de la sección Bitácora, se espera que los temas sean de mayor profundidad. El objetivo es que los usuarios estudien la problemática y aporten con sus conocimientos y así lograr que ocurran innovaciones dentro de la empresa. Además, se espera que esta sección sirva para impulsar los temas, es decir, que se comience la discusión en la plataforma para que luego de haber logrado cierto grado de maduración sea trasladado a conversaciones presenciales y continúe siendo discutido.
- **Wiki:** Sección para publicar temas en mayor profundidad y documentar proyectos que estuvieron en la sección “Iniciativas”.

PIKS corresponde a una Comunidad Virtual de Práctica, ya que según la definición de estas comunidades explicado en la sección 2.2, posee un dominio: la innovación en el cultivo y exportación principalmente de salmón; una comunidad: los miembros conocen sus roles, roles que son similares a los que poseen en la organización, pero esta comunidad no es lo suficientemente solida para generar el aprendizaje deseado por sus administradores mediante los niveles de interacciones que hasta el momento ocurren; y prácticas: los miembros tienen las herramientas, información, entre otros, para compartir con los demás, ya que cada miembro tiene una labor dentro de la organización y por lo tanto el conocimiento sobre algún proceso productivo, el que puede ser compartido con los miembros para desarrollar el conocimiento específico que la comunidad tiene por objetivo, en este caso, la innovación.

## 4.2 Análisis de la Red Social

En esta sección se hará un análisis de la comunidad desde la perspectiva del Análisis de Redes Sociales, es decir, se modelará a la comunidad como una red social y así se evaluará cuantitativamente por medio de la medición de métricas sociales.

Para generar estas métricas se utilizó un framework [23] que permite transformar los datos de una comunidad en archivos con la configuración de la red social, esto es, extrayendo los datos de los usuarios y representarlos como nodos, y las interacciones

entre los usuarios como arcos que unen los nodos. De esta manera se podrá visualizar el grafo de red social y conocer la situación actual de la comunidad.

El framework está construido en lenguaje Java y su funcionamiento se basa en “recorrer” todas las actividades que ocurran en la comunidad. De esta manera va creando objetos para cada foro, tema, respuesta y usuario con que se encuentre para luego exportarlos al archivo con la configuración de la red.

Para aplicar esto en PIKS se debió adaptar el framework a la estructura de la base de datos de la comunidad, lo que en otras palabras significó modificar las consultas SQL que hace el programa a la base de datos que se conecta. Dada la estructura de la comunidad, que se explico en el punto 4.1, esta consta de tres foros: Noticias, Bitácora e Iniciativas, solamente desde estos módulos se extraerán las interacciones de los usuarios, ya que en los otros dos no ocurren interacciones.

Una vez adaptado el programa, se ejecuta y entrega tres archivos con diferentes configuraciones de la red social: *creator*, *lastReply* y *replyAll*. Cada uno permite visualizar el grafo de la red social bajo distintos conceptos. Para el caso en estudio, se utiliza la configuración *replyAll*, dado que, mediante consenso con un administrador, es el que se asemeja en mayor medida a lo que ocurre en la comunidad, esto es, que las respuestas de los temas no son solamente respuesta al creador del tema, sino que además a todos quienes han participado respondiendo anteriormente.

Por otra parte, también se decidió que el grafo de la red social es dirigido, es decir, que cada arco tiene dirección desde el usuario que origina el contenido hacia los usuarios a los que les está respondiendo.

Los datos de los *posts* que se realizan se encuentran divididos en tablas según el foro al que correspondan, mientras que las respuestas a estas publicaciones están almacenadas, sin importar el foro, en una misma tabla.

Con los parámetros definidos anteriormente (base de datos de PIKS, tablas con los datos de interacciones, configuración de red y que el grafo es dirigido) se ejecutó el framework varias veces sobre la comunidad para así obtener la información sobre diferentes períodos de tiempo acumulados. De esta manera se podrá conocer y analizar la evolución de la red desde su inicio. Posteriormente, los archivos generados fueron ejecutados con Gephi para poder visualizar la red y obtener las métricas sociales.

Tabla 3: Evolución de métricas sociales.

Período acumulado desde Abril-2010 hasta:	Arcos	Grado promedio	Diámetro	Densidad
Junio-2010	10	0.204	2	0.001
Septiembre-2010	150	3.061	5	0.016
Diciembre-2010	178	3.633	6	0.019
Marzo-2011	260	5.306	4	0.027
Junio-2011	353	7.204	4	0.037
Septiembre-2011	416	8.490	4	0.044
Diciembre-2011	465	9.490	4	0.049
Febrero-2012	471	9.612	4	0.050

Fuente: Elaboración propia.

Como se observa entonces, se ve que la red ha evolucionado positivamente, ya que el Grado promedio a aumentado constantemente, pero por otra parte la densidad es aún baja, es decir, las conexiones directas existentes entre los usuarios es pequeña comparado con el potencial que tiene la red.

Por otra parte cabe destacar que el administrador al presentársele estos resultados, concluyó que se asemeja con precisión a lo que él observa que ocurre en la realidad, ya que los usuarios que se considera que más participan, el grafo los muestra conectados a gran parte de la red, mientras que los que menos participan tienen muy pocas conexiones con otros miembros.

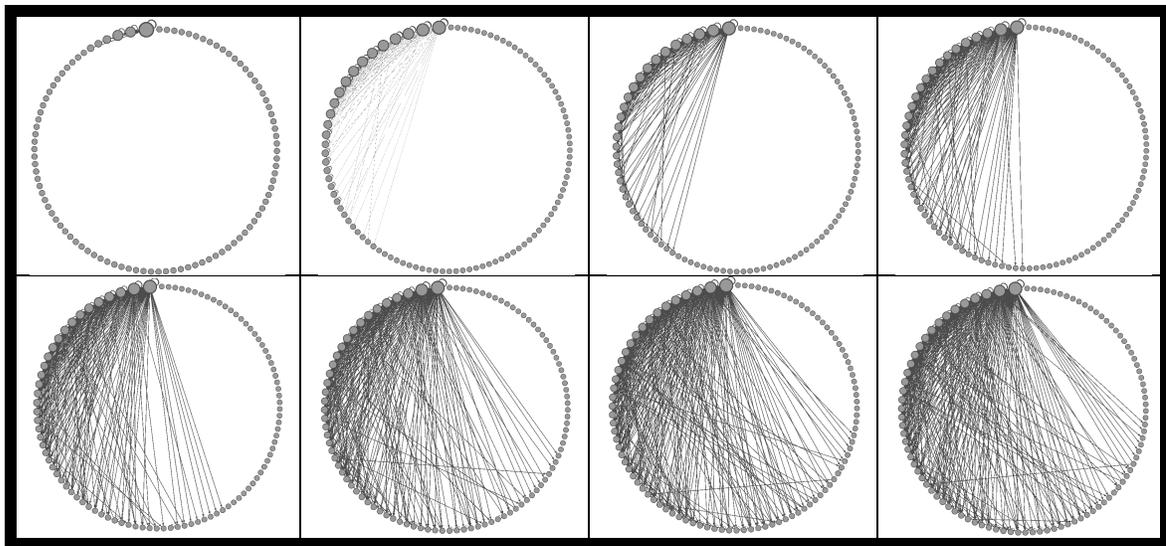


Figura 4: Evolución del grafo de la red.

Fuente: Imagen de Gephi. Elaboración propia.

En la figura 5 se observa la evolución de la comunidad en relación a las conexiones que se han generado entre los usuarios. Los períodos que se visualizan son los mencionados en la Tabla 3. Se puede apreciar cómo ha aumentado la densidad de arcos, lo que indica que la comunidad va desarrollándose con el tiempo y por lo tanto

los usuarios están conversando unos con otros. Sin embargo, aún existe una porción significativa de usuarios que nunca han participado de alguna discusión.

Utilizando Gephi también se ha aplicado el algoritmo HITS para rankear los nodos según los puntajes *hub* y *authority*. El resultado obtenido se encuentra en la Tabla 4. La tabla presenta los 10 usuarios con mayores puntajes y cabe destacar que estos 10 usuarios coinciden si se rankea por hub o por authority independientemente.

Lo anterior se debe a que suelen repetirse los miembros que participan en la mayoría de las discusiones que se producen en la comunidad. Por lo tanto, los miembros que motivan al resto (hubs) son los mismos miembros que suelen responder (authorities). Dado esto, puede concluirse que en esta lista se encuentran quienes son los expertos o líderes de la comunidad. Para comparar con la realidad se presentó estos resultados a un administrador, quien explicó que justamente varios de estos miembros son los principales interesados en que la comunidad se desarrolle, mientras que otros son expertos en los temas que se tratan en la comunidad y por lo tanto suelen participar activamente entregando sus puntos de vistas.

Tabla 4: Resultado aplicación algoritmo HITS

Usuario	Authority	Hub	In-Degree	Out-Degree	Degree
JWessel	0.07262570	0.08119658	38	52	90
FVallejos	0.04841713	0.05555556	25	45	70
LGodoy	0.03724395	0.04273504	19	20	39
PBorquez	0.03724395	0.04273504	19	22	41
RUlloa	0.03724395	0.04059829	19	31	50
LSan Martin	0.03538175	0.04059829	18	28	46
JDomke	0.03351956	0.03846154	17	16	33
CMuñoz	0.03165736	0.03632479	16	16	32
CBallesteros	0.02979516	0.03418804	15	16	31
CVignolo	0.02793296	0.03205128	14	15	29

*El número total de usuarios es de 98.*

Fuente: Elaboración propia

Para concluir el análisis de la comunidad desde la perspectiva del SNA, se ha creado un grafo agrupando a los nodos según el área de pertenencia.

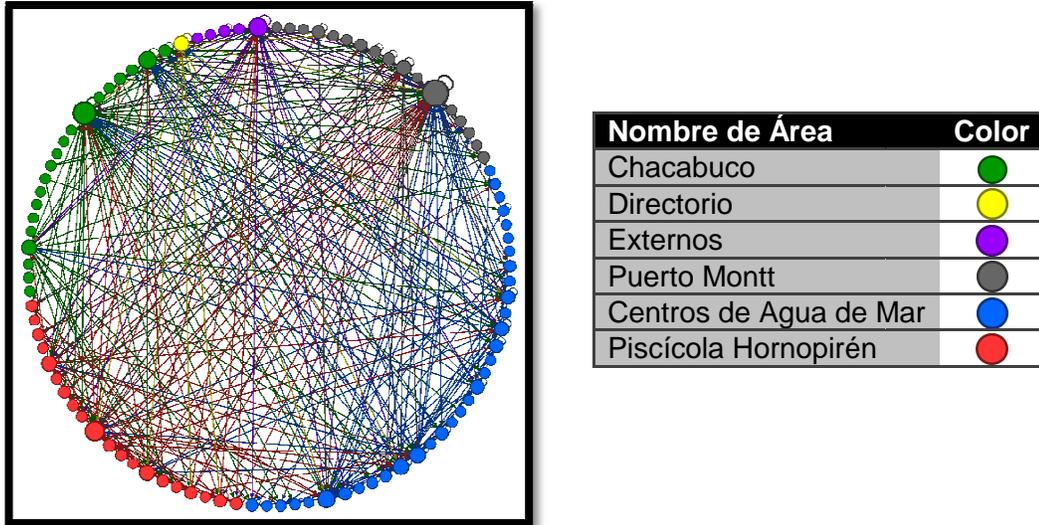


Figura 5: Grafo de la comunidad según área de pertenencia.  
Fuente: Imagen de Gephi. Elaboración propia.

El tamaño de los nodos en la figura depende del grado de cada uno. Con esto puede observarse que en general cada área tiene uno o dos miembros que son los que poseen la mayor cantidad de las interacciones y por lo tanto se vuelve a corroborar el hecho que la mayoría de los miembros no están suficientemente motivados para participar.

### 4.3 Diseño de Métricas e Indicadores de Participación

Dada la descripción de la plataforma hecha en el punto 4.1, para el diseño de métricas se considerará a las secciones Noticias, Bitácora e Iniciativas, las que básicamente son foros en que los usuarios pueden crear hilos de conversación y participar en los hilos que otros usuarios han creado dando a conocer su opinión.

Profundizando en la descripción de la comunidad, se tiene que cada usuario pertenece a una de las áreas de la empresa, estas son: Puerto Montt, Centros de Agua Mar, Piscícola Hornopirén y Chacabuco – estas representan plantas de la empresa que se ubican en distintas zonas geográficas – el Directorio de la empresa también está representado como otra área y finalmente existe el área Externos – que corresponde principalmente a administradores de la comunidad.

Así como existe la segmentación de usuarios en las distintas áreas de la empresa, también existe una segmentación en centros productivos. Estos centros productivos se refieren a lugares específicos de la empresa en los cuales se desarrollan labores de carácter operativo. Algunos de estos centros pertenecen a un área específica de la empresa pero otros no tienen una dependencia clara o fueron creados solo para efectos de ordenar las publicaciones que se realizan en la plataforma. Por ejemplo, el área

Piscícola Hornopirén tiene los centros Incubación, Producción, Administración, entre otros, mientras que los centros FRIOSUR o Transportes no son centros productivos propiamente tales, sino que existen solo para hacer referencia cuando se plantee un quiebre o innovación en la plataforma respecto de esos temas y se pueda comprender y mantener un orden de a qué parte de la empresa se está comentando. Dado lo anterior, no existe una dependencia clara de los usuarios hacia algún centro y por lo tanto no se puede hacer una segmentación de los usuarios en centro así como si ocurre con las áreas de la empresa.

La plataforma también permite conocer el estado de ánimo de cada usuario. Hay un set de 15 estados de ánimos entre los cuales el usuario puede seleccionar el que mejor lo represente en ese momento, algunos de estos estados son: Preocupación, Optimismo, Confianza, entre otros. Estos estados de ánimo se dividen en 2 grupos, en uno se ubican los que abren oportunidades y en el otro los que cierran oportunidades.

Los requerimientos de información de los administradores para mejorar la gestión de la comunidad son en relación a la participación que tienen los usuarios en la plataforma, por lo que se necesita medir las interacciones que ocurren entre ellos. Con esto se espera poder identificar quienes son los usuarios que más participan y lideran la comunidad, con qué frecuencia participan o que secciones son más utilizadas. Además, para los administradores resulta útil conocer el uso de los estados ánimos, por lo que también se medirá a qué grupo pertenecen los estados de ánimo que los usuarios utilizan en la plataforma.

Como se ha mencionado, los usuarios pueden crear hilos de conversación o comentar otros, entonces, se definirá formalmente como Publicación a la creación de hilos, mientras que a los comentarios de los hilos se les definirá como Respuestas.

A partir de lo anterior se diseñarán métricas para medir los niveles de participación y posteriormente construir indicadores acerca de estas métricas.

Siguiendo la metodología propuesta en [12], primero se debe definir la necesidad de información que se tiene, la que corresponde a conocer los niveles de participación de la comunidad. Luego, para obtener esta información se deben definir atributos medibles que permitirán satisfacer la necesidad. Estos atributos son la cantidad de Publicaciones o Respuestas que ocurren en la comunidad, ya sea en las distintas secciones (foros) de la plataforma o puede ser por usuario, área o centro existente.

Con los atributos definidos se pueden obtener métricas, las que representan la medición o cuantificación del atributo. Además, a diferencia del atributo, la métrica es comparable y por lo tanto permite analizar y concluir. De los atributos se definieron métricas en relación a períodos de tiempo, como por ejemplo, cantidad de Publicaciones de una sección durante un año o mes determinado o también la cantidad

de Respuestas hechas por un usuario en una sección durante una semana determinada.

Teniendo estas métricas se podrán crear indicadores elementales para poder evaluar y conocer la evolución de las métricas, estos indicadores son una composición de métricas, por ejemplo, que porción de las publicaciones realizadas en todas la comunidad durante un período determinado pertenecen a la sección Noticias o que porción del total de respuestas hechas en la sección Bitácora durante cierto mes corresponden a un usuario determinado.

Finalmente, también se podrán crear indicadores globales, los cuales se calculan utilizando indicadores elementales, estos permiten entender lo que ocurre en la comunidad desde una perspectiva más general, ya que evalúan a otros indicadores. Ejemplo de estos puede ser la variación de la porción de publicaciones en la sección Noticias durante determinado mes, respecto del mes anterior.

Tabla 5: Métricas e indicadores para publicaciones 1.

Necesidad de Información	Atributo	Métrica PST	Indicador Elemental IPST	Indicador Elemental VPST	Indicador Global VIPST
Conocer niveles de participación	Cantidad de publicaciones en la sección X	Semanal (Cantidad de publicaciones en la sección X durante el período Y)	Semanal (Porcentaje de publicaciones en la sección X durante el período Y respecto de todas las secciones)	Semanal (Variación de cantidad de publicaciones en la sección X durante el período Y respecto del período (Y-1))	Semanal (Variación del porcentaje de publicaciones en la sección X durante el período Y respecto de período (Y-1))
		Mensual (Cantidad de publicaciones en la sección X durante el período Y)	Mensual (Porcentaje de publicaciones en la sección X durante el período Y respecto de todas las secciones)	Mensual (Variación de cantidad de publicaciones en la sección X durante el período Y respecto del período (Y-1))	Mensual (Variación del porcentaje de publicaciones en la sección X durante el período Y respecto de período (Y-1))
		Anual (Cantidad de publicaciones en la sección X durante el período Y)	Anual (Porcentaje de publicaciones en la sección X durante el período Y respecto de todas las secciones)	Anual (Variación de cantidad de publicaciones en la sección X durante el período Y respecto del período (Y-1))	Anual (Variación del porcentaje de publicaciones en la sección X durante el período Y respecto de período (Y-1))

Fuente: Elaboración propia.

Tabla 6: Métricas e indicadores para publicaciones 2.

Necesidad de Información	Atributo	Sub atributo	Métrica PSTE	Indicador Elemental IPSTE
Conocer niveles de participación	Cantidad de publicaciones en la sección X	Cantidad de publicaciones en la sección X realizadas por la entidad Z	Semanal (Cantidad de publicaciones en la sección X durante el período Y realizadas por entidad Z)	Semanal (Porcentaje de publicaciones en la sección X durante el período Y realizadas por la entidad Z respecto de toda la sección)
			Mensual (Cantidad de publicaciones en la sección X durante el período Y realizadas por entidad Z)	Mensual (Porcentaje de publicaciones en la sección X durante el período Y realizadas por la entidad Z respecto de toda la sección)
			Anual (Cantidad de publicaciones en la sección X durante el período Y realizadas por entidad Z)	Anual (Porcentaje de publicaciones en la sección X durante el período Y realizadas por la entidad Z respecto de toda la sección)

Fuente: Elaboración propia.

Las métricas e indicadores descritos en las tablas anteriores están definidos para las publicaciones y de igual manera se definió un set para las respuestas que se originen.

En la Tabla 5 se realizaron las definiciones en base a conocer los niveles de participación, el atributo medible en este caso es la cantidad de *posts* (publicaciones o respuestas) para una sección determinada, mientras que en la Tabla 6 se desprende un sub atributo que se refiere a medir los *posts* relacionados a una sección y entidad determinadas, donde una entidad corresponde a un usuario o área, y para la sección Bitácora se suma una tercera entidad que son los centros.

Las definiciones formales de cada métrica e indicador de las tablas son las siguientes:

*Métrica PST: Métrica de [P]osts (publicaciones o respuestas) en la [S]ección X durante el período de [T]iempo Y*

*Indicador Elemental IPST: Indicador de [P]osts (publicaciones o respuestas) en la [S]ección X durante el período de [T]iempo Y respecto de todas las secciones*

$$IPST(X, Y) = \frac{Métrica\ PST(X, Y)}{\sum_x Métrica\ PST(X, Y)}$$

*Indicador Elemental VPST: Indicador de [V]ariación de [P]osts (publicaciones o respuestas) en la [S]ección X durante el período de [T]iempo Y respecto de (Y - 1)*

$$VPST(X, Y) = \frac{Métrica\ PST(X, Y)}{Métrica\ PST(X, Y - 1)}$$

*Indicador Global VIPST: Indicador de [V]ariación del [I]ndicador de [P]osts (publicaciones o respuestas) en la [S]ección X durante el período de [T]iempo Y respecto de (Y - 1)*

$$VIPST(X, Y) = \frac{Indicador\ Elemental\ IPST(X, Y)}{Indicador\ Elemental\ IPST(X, Y - 1)}$$

*Métrica PSTE: Métrica de [P]osts (publicaciones o respuestas) en la [S]ección X durante el período de [T]iempo Y realizados por la [E]ntidad W*

*Indicador Elemental IPSTE: Indicador de [P]osts (publicaciones o respuestas) en la [S]ección X durante el período de [T]iempo Y realizados por la [E]ntidad Z respecto de toda la sección*

$$IPSTE(X, Y) = \frac{Métrica\ PSTE(X, Y, Z)}{\sum_z Métrica\ PSTE(X, Y, Z)}$$

También se han definido métricas e indicadores para conocer el uso de estados de ánimo, el atributo que se medirá es la cantidad de veces que es utilizado cada tipo de estado de ánimo por los miembros de la comunidad. En la tabla siguiente se muestran las definiciones a las que se llegó.

Tabla 7: Métricas e indicadores para estados de ánimo.

Necesidad de Información	Atributo	Métrica EA	Indicador Elemental EA
Conocer uso de estados de ánimo	Cantidad de veces que se utiliza un estado de ánimo del tipo X	Semanal (Cantidad de utilización de estados de ánimo del tipo X durante el período Y)	Semanal (Porcentaje de utilización de estados de ánimo del tipo X durante el período Y respecto del total)
		Mensual (Cantidad de utilización de estados de ánimo del tipo X durante el período Y)	Mensual (Porcentaje de utilización de estados de ánimo del tipo X durante el período Y respecto del total)
		Anual (Cantidad de utilización de estados de ánimo del tipo X durante el período Y)	Anual (Porcentaje de utilización de estados de ánimo del tipo X durante el período Y respecto del total)

Fuente: Elaboración propia.

Los tipos de estados de ánimos son dos, uno corresponde a los estados que pertenecen al grupo que abren posibilidades y el otro al que cierran oportunidades. Las definiciones formales son las siguientes:

*Métrica EA: Métrica de utilización de [E]stados de [Á]nimo del tipo X durante el período [Y]*

*Indicador Elemental EA: Indicador de utilización de [E]stados de [Á]nimo del tipo X durante el período Y respecto del total de uso*

$$IEA(X, Y) = \frac{Métrica EA(X, Y)}{\sum_x Métrica EA(X, Y)}$$

Los parámetros de este diseño final de métricas e indicadores, tanto para participación como para estados de ánimo, se logró mediante consenso con un administrador. En este consenso se definió, por ejemplo, que las unidades de tiempo debían ser semana, mes y año, en particular, la semana era de importancia por ser la base sobre la que se hacen los cálculos de participación para analizar la comunidad en la actualidad. De igual manera, la decisión de medir los estados de ánimo en función del grupo al que pertenecen o que las mediciones se hagan sobre toda la comunidad y no sobre las distintas entidades, se debe a que entrega más utilidad para los administradores conocer el uso global y comparar entre los dos grupos, en vez de medir para cada uno de los quince estado de ánimo existentes o para cada usuario y área.

Con este diseño se espera que los administradores puedan comprender la situación de la comunidad, en términos de participación y a distintos niveles de agregación, y por lo tanto puedan analizarla y obtener conclusiones sobre lo que ocurre para poder tomar decisiones que fomenten un mejoramiento y mayor uso de la plataforma.

Se realizó un cálculo preliminar de algunas de estas métricas para dimensionar la situación de la comunidad. Para esto se hicieron consultas en lenguaje SQL sobre la

base de datos de la comunidad utilizando el software MySQL Workbench y posteriormente se graficaron los resultados:

- Se calculó la Métrica PST para publicaciones y respuestas en las 3 secciones y para cada mes comprendido entre Febrero-2011 y Enero-2012.

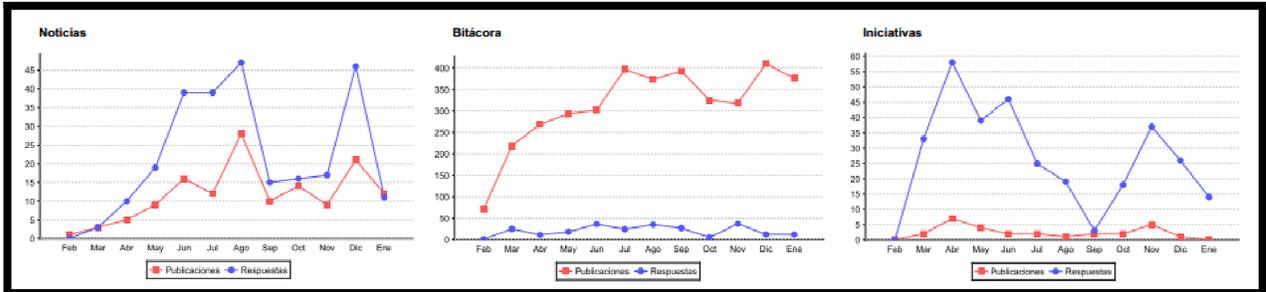


Figura 6: Ejemplo de cálculo de Métrica PST.  
Fuente: Elaboración propia.

- Métrica EA para los dos grupos de estados de ánimo durante las primeras 6 semanas del año 2012.

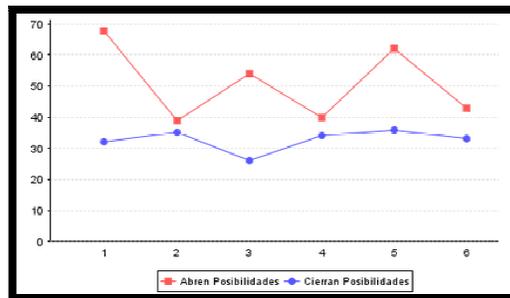


Figura 7: Ejemplo de Métrica EA.  
Fuente: Elaboración propia.

De la misma manera, también se calcularon algunos de los indicadores creados:

- Indicador IPSTE para las publicaciones y respuestas en la sección Noticias hechas durante el año 2011 por los 10 usuarios más participativos, acompañado de la métrica respectiva. El total de publicaciones realizadas fue de 136 y el de respuestas fue de 264.

Tabla 8: Ejemplo de cálculo de indicador IPSTE.

Publicaciones				Respuestas			
Ranking	Usuario	Métrica PSTE	Indicador IPSTE	Ranking	Usuario	Métrica PSTE	Indicador IPSTE
1	Joachim Wessel	22	16%	1	Joachim Wessel	88	33%
2	Nancy Ruiz	13	10%	2	Roberto Ulloa	21	8%
3	Francisco Vallejos	11	8%	3	Francisco Vallejos	21	8%
4	Roberto Ulloa	7	5%	4	Carlos Vignolo	11	4%
5	Patricio Borquez	7	5%	5	Leonel Godoy	11	4%
6	Pablo Werner	6	4%	6	Luis San Martin	10	4%
7	Horacio Lyon	5	4%	7	Patricio Borquez	9	3%
8	Leonel Godoy	5	4%	8	Claudio Muñoz	6	2%
9	Jaime Espinoza	4	3%	9	Sergio Garcia	5	2%
10	Veronica Muñoz	4	3%	10	Veronica Muñoz	5	2%

Fuente: Elaboración propia.

- Indicador IPST, VPST y VIPST para publicaciones y respuestas en las secciones Noticias y Bitácora durante los últimos 6 meses.

Tabla 9: Ejemplo de cálculo de indicadores IPST, VPST y VIPST.

Noticias								
Período	Publicaciones				Respuestas			
	Métrica PST	Indicador IPST	Indicador VPST	Indicador VIPST	Métrica PST	Indicador IPST	Indicador VPST	Indicador VIPST
Agosto-2011	28	7%	233%	233%	47	47%	121%	107%
Septiembre-2011	10	2%	36%	29%	15	33%	32%	70%
Octubre-2011	14	4%	140%	200%	16	41%	107%	124%
Noviembre-2011	9	3%	64%	75%	17	19%	106%	46%
Diciembre-2011	21	5%	233%	167%	46	55%	271%	289%
Enero-2012	12	3%	57%	60%	11	30%	24%	55%
Bitácora								
Período	Publicaciones				Respuestas			
	Métrica PST	Indicador IPST	Indicador VPST	Indicador VIPST	Métrica PST	Indicador IPST	Indicador VPST	Indicador VIPST
Agosto-2011	373	93%	94%	96%	35	35%	146%	130%
Septiembre-2011	393	97%	105%	104%	27	60%	77%	171%
Octubre-2011	325	95%	83%	98%	5	13%	19%	22%
Noviembre-2011	318	96%	98%	101%	37	41%	740%	315%
Diciembre-2011	410	95%	129%	99%	12	14%	32%	34%
Enero-2012	376	97%	92%	102%	12	32%	100%	229%

Fuente: Elaboración propia.

El mismo procedimiento se aplicó a la mayoría de las métricas e indicadores diseñados, con esto se obtuvo información objetiva de la participación de la comunidad y se realizó un análisis de la situación general.

De estos resultados es posible concluir que en general los usuarios más recurrentes, es decir, que aparecen con más publicaciones o respuestas, pertenecen en su mayoría a la administración de la comunidad o son usuarios que poseen cargos de mayor responsabilidad en la empresa. Con esto se refleja la apreciación de los

administradores, esto es, que los trabajadores de la empresa no están suficientemente motivados para participar, a pesar de ser uno de sus deberes hacerlo y en general solo participan activamente quienes están interesados en que la comunidad se desarrolle.

El indicador IPSTE, aplicado a las respuestas de las tres secciones y durante períodos anuales, muestra que los usuarios más participativos se suelen repetir en los rankings generados y siempre se tiene que componen más del 40% de las interacciones de cada sección, es decir, el 5% (de un total de 98 usuarios) generan la mayoría del contenido.

Esto es relevante, pues, debido a los objetivos de la comunidad, medir la cantidad de respuestas entrega más información que conocer la cantidad de publicaciones que se originan, ya que las respuestas indican que se están produciendo discusiones entre los usuarios y, aunque si bien pueden existir comentarios que no aporten a resolver el quiebre que se discute, sí ayuda a que la conversación avance y puedan conocerse más perspectivas para abordar el problema.

Por ejemplo, durante el 2011, en las respuestas de la sección Bitácora los usuarios de los cinco primeros lugares poseen el 43% del total de respuestas, en Iniciativas el 71% pertenece a los cinco primeros lugares y en Noticias el 57%.

Como se ha dicho, se espera que los usuarios comenten los quiebres que se publican y que se genere una conversación entre ellos, es por esto que es lógico esperar que existan más respuestas que publicaciones en cada sección.

La información obtenida no necesariamente indica que se esté cumpliendo lo anterior. Por ejemplo: en la sección Noticias, durante el año 2010, se tuvieron 61 publicaciones y 105 respuestas y, durante el año 2011, se tuvieron 136 publicaciones y 264 respuestas; en la sección Iniciativas, en el año 2011, se tuvieron 31 publicaciones y 318 respuestas y en el 2010 solo 1 publicación y 3 respuestas. Con esta información se puede concluir que se han logrado generar conversaciones en estas secciones, sobre todo en Iniciativas, donde se plantean quiebres de mayor envergadura. Sin embargo, en la sección Bitácora, en la que los trabajadores deben publicar quiebres diariamente, se tiene que cumplen con esto, pero no están interesados en comentarlos para tratar de ayudarse entre sí. Durante el año 2010 se tuvieron 1605 publicaciones y solo 255 respuestas y durante el año 2011 se tuvieron 3894 publicaciones y 272 respuestas, en consecuencia, se observa que ocurre lo contrario que en las otras dos secciones, los usuarios parecieran no estar generando conversaciones para compartir experiencias o conocimientos y por lo tanto la plataforma no estaría cumpliendo sus objetivos.

Si bien existe un nivel significativo de participación en la comunidad cuando se analiza globalmente, ya que desde el año 2011 la participación es relativamente constante semana a semana, esta participación es sostenida principalmente por la sección Bitácora, pues, en las otras dos secciones las publicaciones y respuestas son

irregulares. En particular, en la sección Iniciativas, ocurre que las discusiones se mantienen por semanas, pero durante otras las métricas caen a cero.

Entonces, como puede observarse de los resultados obtenidos del diseño de métricas e indicadores propuesto, la participación no ha crecido como los administradores desean, debido a que en algunas secciones la participación es esporádica y en general las discusiones que se producen y que generan la mayor actividad de la comunidad, se concentran en quienes son los interesados que la comunidad se desarrolle y por lo tanto se puede concluir que no se ha logrado traspasar la motivación para que los demás trabajadores participen.

## Capítulo 5

# Construcción de Repositorio de Datos

Este capítulo tiene por objetivo la construcción de un repositorio centralizado de datos, en el que se almacenen los datos relacionados con las interacciones de los usuarios en la plataforma y que permitan medir los niveles de participación de la comunidad, o sea, las métricas de participación definidas en el capítulo anterior.

Para esto, primero se explorarán los datos que se registran en la base de datos de la plataforma, posteriormente se propondrá un diseño de Data Mart siguiendo un Modelamiento Constelación, es decir, definir tablas de hechos que almacenarán las medidas que se tomen de la base de datos y las respectivas tablas de dimensiones que permitirán visualizar la información según los requerimientos de agregación que se necesiten, y finalmente se diseñará el proceso de Extracción, Transformación y Carga de datos que permitirá obtener de la base de datos de la plataforma los datos relevantes según los requerimientos de diseño del Data Mart.

### 5.1 Base de datos de la plataforma

La base de datos con todos los registros de la actividad que ocurre en la plataforma está compuesta por 29 tablas, muchas de estas tablas presentan redundancia de datos y prácticamente no están normalizadas. Dado esto, adquiere mayor relevancia la creación de un Data Mart que contenga los datos limpios para que puedan ser consultados de manera sencilla.

De las 29 tablas existentes solo se usarán 11 para poblar el Data Mart. Cada una de estas tablas contiene sus correspondientes *identificadores* y registro de fechas según corresponda, entre otros. A partir de estas tablas se extraerán los datos que serán almacenados en el Data Mart, cuyo diseño se explicará a continuación.

Tabla 10: Tablas de la base de datos para la extracción de métricas.

Nombre de la tabla	Datos que se extraerán
<i>noticia</i>	Contiene publicaciones de la sección noticias.
<i>bitacora</i>	Contiene publicaciones de la sección bitácora.
<i>iniciativas</i>	Contiene publicaciones de la sección iniciativas.
<i>comentario</i>	Contiene los comentarios que se hacen en todas las secciones de la plataforma.
<i>notificacion</i>	Contiene toda la actividad de la comunidad. De esta se extraerán los cambios de estados de ánimo.
<i>usuario</i>	Datos de cada usuario (nombre, e-mail, password, entre otros).
<i>centro</i>	A pesar del nombre, contiene los datos de las áreas.
<i>centrousuari</i>	Dependencia entre usuario y área.
<i>proyectos</i>	Listado de centros productivos de la empresa.
<i>estadoanimo</i>	Contiene los 15 estados de ánimo existentes.

Fuente: Elaboración Propia.

## 5.2 Diseño de Data Mart

Considerando las métricas e indicadores definidos en el capítulo anterior, lo que se requiere almacenar en el Data Mart son las métricas de participación, es decir, la medición de *posteos* que cada usuario ha realizado en la plataforma, ya sean publicaciones de nuevos hilos de conversación o respuestas en temas creados por otros usuarios, además de la métrica que mide el uso de las actualizaciones de estados de ánimo.

Para esto se diseñó un repositorio que consiste en 7 tablas de hechos y 5 dimensiones, dimensiones que son compartidas entre algunas tablas de hechos y por lo tanto el diseño representa un Modelo Constelación. La decisión de representar la participación de esta manera se debe a que se agruparon los atributos según su significado y así los esquemas de relación resulten más sencillos de interpretar, es decir, se utilizó la semántica de los atributos como guía para construir el Data Mart [33]. Otra alternativa era haber creado el esquema reuniendo a las medidas que tiene las mismas dimensiones en una misma tabla de hechos, pero se hubiera tenido que haber creado una dimensión extra que indicara a que foro pertenecen y otra para indicar si corresponde a una publicación o a una respuesta, si bien esto hubiera resultado en menos tablas de hechos, la implementación de la solución (cubos OLAP y reportes) hubiera sido más compleja y los esquemas serían menos intuitivos.

Se decidió que la mejor manera de almacenar las métricas de cada tabla de hechos es calcular el valor de la métrica diariamente, por lo que la unidad básica de tiempo será el día.

Tabla 11: Resumen de tablas del Data Mart diseñado.

Nombre de la tabla	Dimensiones que posee
<i>fact_publicaciones_noticias</i>	dim_tiempo, dim_usuarios, dim_areas
<i>fact_publicaciones_bitacora</i>	dim_tiempo, dim_usuarios, dim_areas, dim_centros
<i>fact_publicaciones_iniciativas</i>	dim_tiempo, dim_usuarios, dim_areas
<i>fact_respuestas_noticias</i>	dim_tiempo, dim_usuarios, dim_areas
<i>fact_respuestas_bitacora</i>	dim_tiempo, dim_usuarios, dim_areas, dim_centros
<i>fact_respuestas_iniciativas</i>	dim_tiempo, dim_usuarios, dim_areas
<i>fact_actualizaciones_estados_animo</i>	dim_tiempo, dim_usuarios, dim_areas, dimestadosanimo

Fuente: Elaboración Propia.

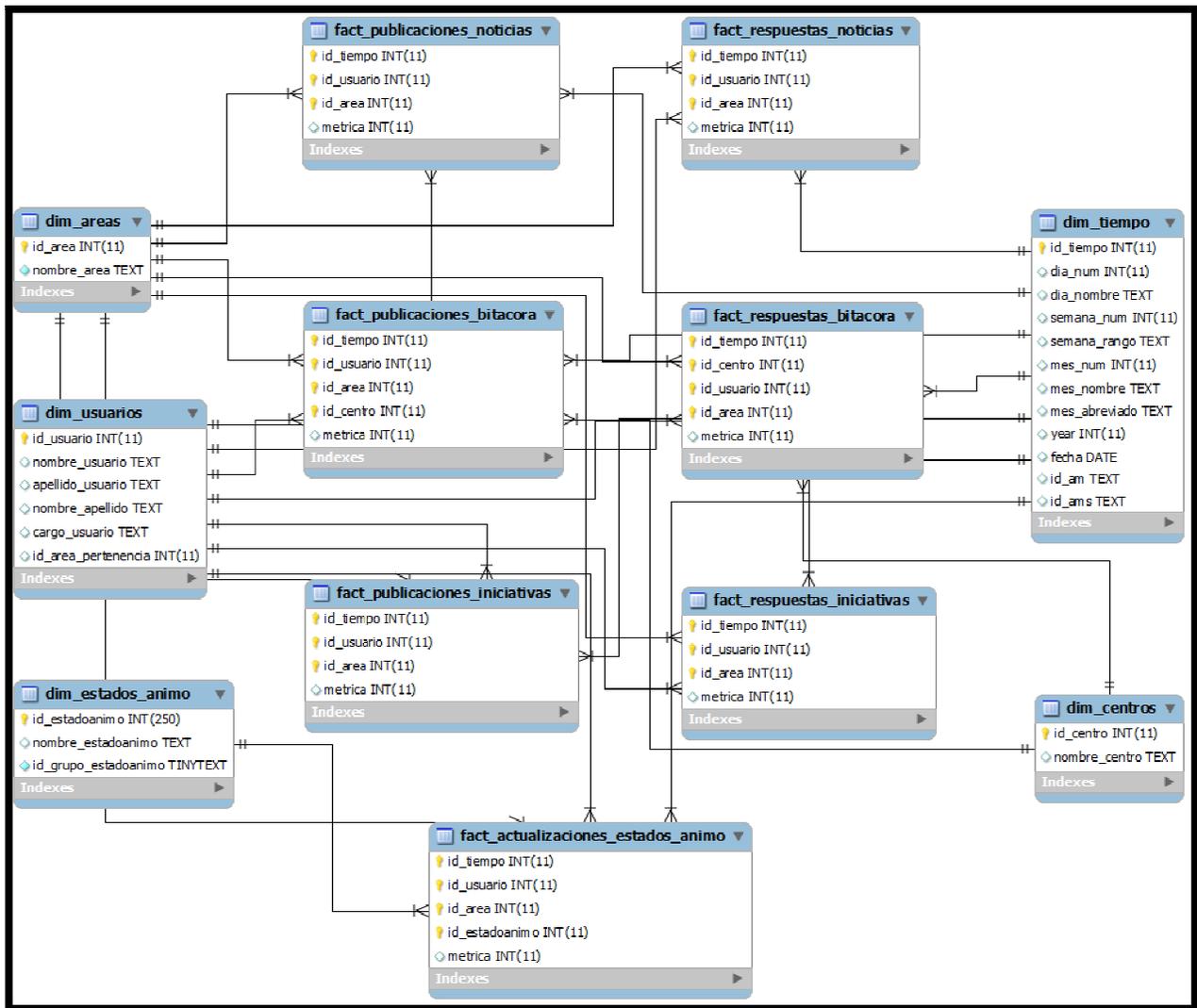


Figura 8: Modelo Relacional del Data Mart

Fuente: Elaboración Propia.

El Data Mart fue construido sobre el motor de bases de datos MySQL. Una vez poblado el repositorio, este será la base sobre la que se realizarán los cálculos necesarios para obtener las métricas e indicadores para hacer el seguimiento y control de la participación de los usuarios de la comunidad.

## 5.3 Proceso de Extracción, Transformación y Carga

Para diseñar el proceso ETL, que permitirá hacer la carga de datos en el Data Mart, se utilizó la herramienta Jaspersoft ETL, la que permite crearlo de manera sencilla, gracias a su interfaz gráfica.

### 5.3.1 Creación de Dimensiones

El proceso ETL para la carga de las dimensiones puede observarse gráficamente en la siguiente figura:

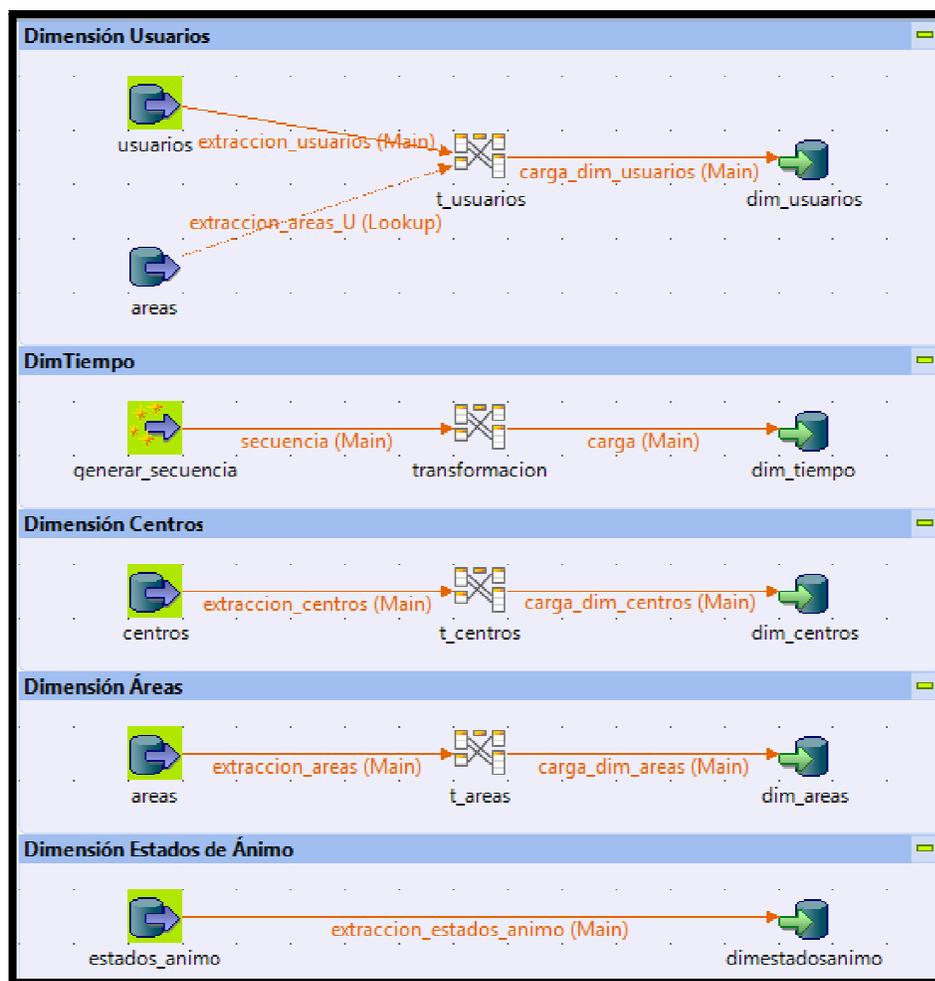


Figura 9: Proceso ETL para las dimensiones del Data Mart.

Fuente: Elaboración Propia.

**Dimensión Usuarios:** Esta dimensión contiene los datos de cada miembro de la comunidad. Para esto se extrae desde la tabla de usuarios el identificador, nombre, apellido y área. Estos datos se limpian y se cargan en la tabla *dim\_usuarios* del Data Mart y el proceso ETL termina.

**Dimensión Tiempo:** Esta dimensión representa los días en que ocurren las interacciones entre los usuarios que están almacenadas en las tablas de hechos. Por lo tanto, el atributo básico es el día.

Para lograr distintos niveles de agregación y para que la dimensión sea consistente con el diseño de las métricas, fueron creados atributos como: número de semana, mes, año, entre otros.

El proceso ETL diseñado consiste en generar una secuencia de fechas partiendo desde la creación de la comunidad. Posteriormente, se extraen los atributos de esta fecha, como la semana, mes y año, y se transforman para ser consistentes con el diseño del Data Mart. Finalmente, se cargan los registros a la tabla *dim\_tiempo* y el proceso termina. Cabe mencionar que para la transformación de algunos atributos fue necesario programar rutinas de trabajo en Java.

**Dimensión Centros:** Similar a la dimensión usuarios, se extrae el identificador y nombre de cada centro y se carga en la tabla *dim\_centros*.

**Dimensión Áreas:** Para esta dimensión solo se debe extraer el identificador y nombre del área para ser cargados en la tabla *dim\_areas*.

**Dimensión Estados de Ánimo:** Se debe extraer el identificador, nombre y grupo al que pertenece cada estado de ánimo existente y se carga en la tabla *dim\_estados\_animo*.

### 5.3.1 Creación de Tablas de hechos

Dado el diseño del Data Mart se deben construir y poblar siete tablas de hechos, las que almacenarán la medición de las métricas de cada día para las diferentes secciones de la plataforma, guardando además el usuario, área y centro, según corresponda, que se esté referenciando.

El proceso ETL diseñado para las tablas de hechos de la sección Noticias (publicaciones y respuestas) se presenta en la figura 10. Para las demás tablas los procesos son similares.

**Tabla de Hechos Publicaciones en Noticias:** Como se ha mencionado, esta tabla contiene la métrica de publicaciones diarias calculada para cada usuario en la sección Noticias. La tabla contiene cuatro atributos, un atributo para la métrica y los otros tres corresponden a las llaves: *id\_tiempo*, *id\_usuario* e *id\_area*.

El proceso ETL diseñado consiste en extraer las interacciones registradas en la base de datos de la plataforma junto con rescatar las llaves correspondientes que

identifican a la métrica, transformar estos datos según el diseño requerido por el Data Mart y almacenarlos en la tabla *fact\_publicaciones\_noticias*.

**Tabla de Hechos Respuestas en Noticias:** El llenado de esta tabla es similar a las publicaciones, la diferencia es que se conecta a la tabla donde se encuentran los comentarios que los usuarios realizan.

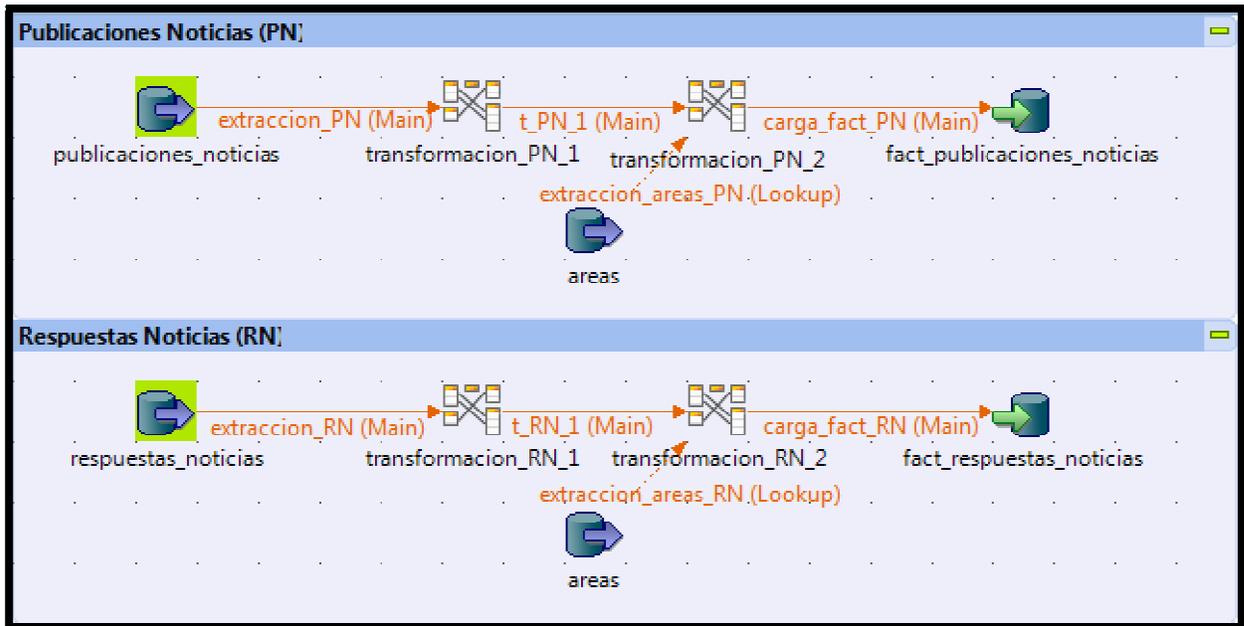


Figura 10: Proceso ETL para tablas de hechos de sección Noticias.

Fuente: Elaboración Propia.

**Tabla de Hechos Publicaciones en Iniciativas y Tabla de Hechos en Respuestas Iniciativas:** El proceso para la obtención de los datos correspondientes a estas tablas es similar a los descritos para la sección Noticias. Contienen las mismas llaves y la diferencia radica solamente en que se extraen y recolectan las métricas para otra sección de la plataforma.

**Tabla de Hechos Publicaciones en Bitácora y Tabla de Hechos Respuestas en Bitácora:** Estas tablas se diferencian de las tablas de las secciones anteriores solamente en que tienen una llave más. Esta llave corresponde al centro sobre el que el usuario ha publicado o respondido. La lógica del proceso ETL es la misma que para las otras tablas, la única diferencia apreciable es que se maneja un atributo más y por lo tanto las tablas *fact\_publicaciones\_bitacora* y *fact\_respuestas\_bitacora* contienen una columna extra.

**Tabla de Hechos para Uso de Estados de Ánimo:** Esta tabla contiene la métrica de uso diaria de cada estado de ánimo. Contiene cuatro llaves: *id\_tiempo*, *id\_usuario*, *id\_area* e *id\_estadoanimo*. El proceso ETL extrae los registros de actualización de estados de ánimo que realizan los usuarios, los transforma para ajustarse al diseño del Data Mart y posteriormente se cargan en la tabla *fact\_actualizaciones\_estados\_animo*.

Jaspersoft ETL permite realizar todos estos procesos en un mismo *Job*, de esta manera, una vez diseñadas las extracciones, transformaciones y cargas de datos y realizadas las configuraciones a las base de datos para la extracción y para la carga, se ejecuta el Job y todos los procesos son ejecutados uno tras otro.

## Capítulo 6

# Implementación de Métricas e Indicadores de Participación

El propósito de este capítulo es explicar la implementación de las métricas e indicadores de participación definidos en el capítulo 4. El objetivo es crear un sistema que permita visualizar esta información con mayor automatización en comparación a como se obtendrían en la actualidad.

Este sistema se construyó utilizando las herramientas de Jaspersoft y el Data Mart construido en el capítulo 5. Específicamente, se explicará cómo se construyeron los cubos OLAP que permiten visualizar las métricas modificando las dimensiones según los requerimientos del usuario y como se diseñaron los reportes para conocer los indicadores de participación.

### 6.1 Construcción de cubos OLAP

Los cubos OLAP permiten visualizar los datos almacenados en un repositorio (En este caso el Data Mart construido) multidimensional. El propósito principal es obtener información relevante realizando cruces entre las dimensiones del repositorio y también poder consultar los datos según los niveles de agregación que se deseen.

Para realizar esto se utilizaron los siguientes softwares: JasperAnalysis Schema Workbench para el diseño de los cubos y JasperReports Server para visualizarlos.

La herramienta JasperAnalysis Schema Workbench permite diseñar los esquemas de los cubos, mediante una interfaz gráfica se ingresan las características del cubo y luego entrega un esquema en formato XML que contiene la estructura del cubo. Para lograr esto, el software se conecta al Data Mart y se hacen las referencias a las tablas que contienen los datos del cubo que se quiere crear.

Para cada tabla de hechos del Data Mart se creará un cubo OLAP. Por lo tanto, en cada esquema se debe especificar la tabla de hechos – de la que se obtienen las

medidas del cubo, en este caso las métricas de participación – y las tablas de dimensiones – en donde se especifican los niveles de jerarquía que cada dimensión posee, por ejemplo, para la dimensión tiempo existen tres niveles: semanas, meses y años. Por lo tanto, en total existirán siete cubos OLAP: Tres para las publicaciones (Uno para cada sección), tres para las respuestas (Uno para cada sección) y uno para las actualizaciones de estados de ánimo.

Para el cubo que contendrá las publicaciones de la sección noticias, una vez ingresadas todas sus especificaciones, se obtiene la siguiente estructura<sup>12</sup>:

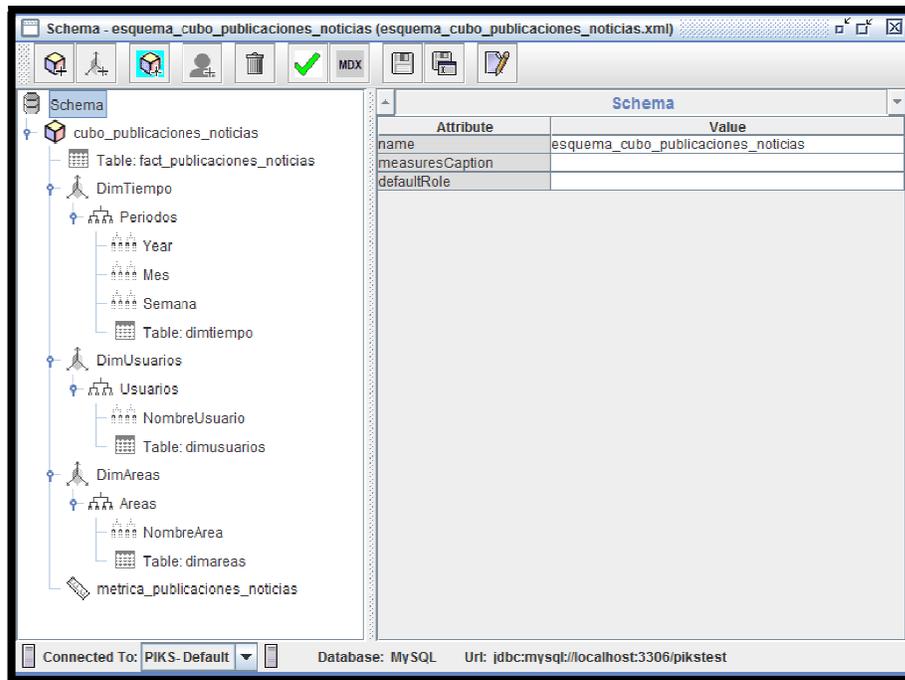


Figura 11: Estructura del cubo de publicaciones de la sección Noticias.  
Fuente: Imagen de JasperAnalysis Schema Workbench. Elaboración propia.

De esta misma manera se obtiene un esquema para el resto de las tablas de hechos del Data Mart con sus correspondientes dimensiones. En la Tabla 15 se pueden observar las tablas de hechos y las tablas de dimensiones correspondientes que debe contener cada esquema.

.Para implementar los esquemas de cubos diseñados se utiliza JasperReports Server, este software permite visualizar las consultas que se hagan a los cubos según los requerimientos de información que el usuario necesite, es decir, permite hacer cruces de dimensiones y modificar los niveles de agregación, entre otras opciones.

<sup>12</sup> Un ejemplo del esquema en formato XML de este cubo se encuentra en el Anexo B.

Los esquemas XML diseñados anteriormente deben ser cargados a JasperReports Server, en donde además se debe crear una conexión al repositorio y especificar la consulta MDX que se hará al cubo para obtener la información sobre las métricas.

Realizado lo anterior, el software será capaz de conectarse al Data Mart y, siguiendo la estructura del cubo, ejecutará la consulta ingresada, obteniéndose la siguiente vista para el caso del cubo de publicaciones en la sección Noticias:

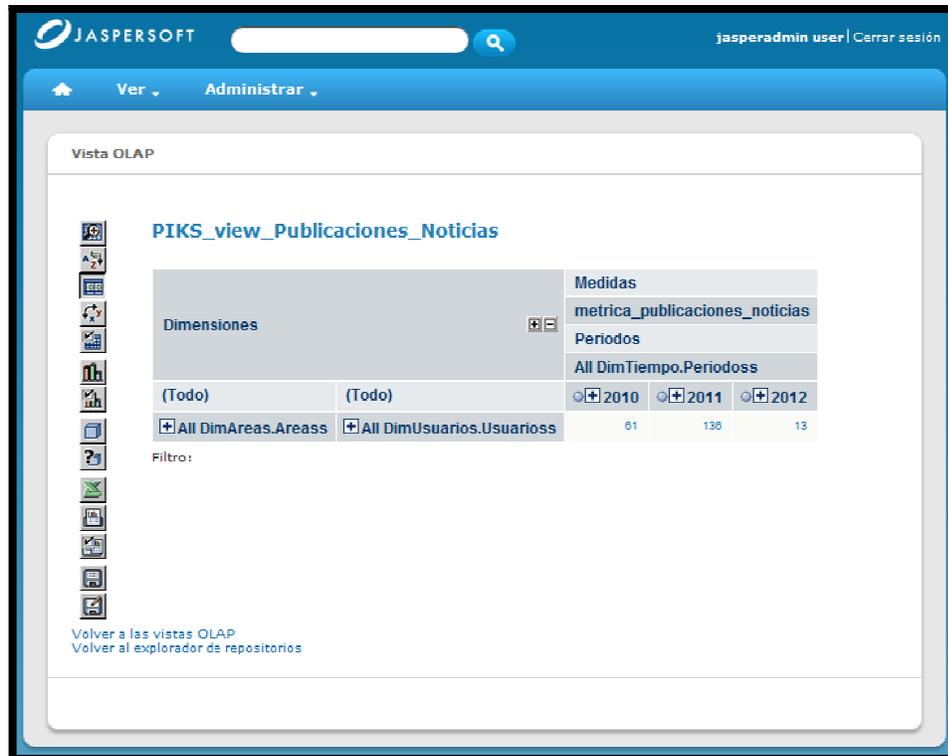


Figura 12: Vista de cubo de publicaciones en sección Noticias 1.

Fuente: Imagen de JasperReports Server. Elaboración propia.

En esta vista se obtiene la métrica de publicaciones en Noticias por año y para todos los usuarios de la plataforma. La interfaz permite al usuario desagregar el año en meses y los meses en semanas, así como desagregar por Áreas y por Usuarios o eliminar/agregar una dimensión, entre varias opciones más.

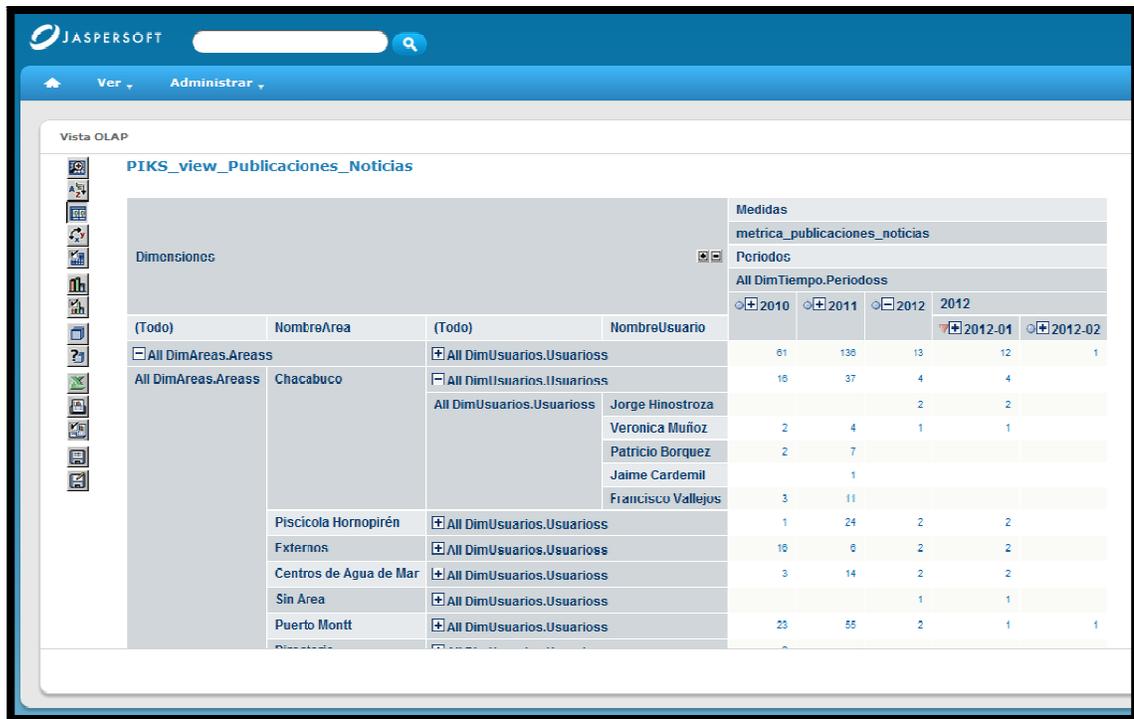


Figura 13: Vista de cubo de publicaciones en sección Noticias 2.  
Fuente: Imagen de JasperReports Server. Elaboración propia

En la anterior se observan las métricas desagregadas por meses del año 2012 y desagregadas por Áreas, además el Área Chacabuco ha sido desagregada en los Usuarios de esa Área que han publicado noticias.

La aplicación además permite ordenar la medida del cubo de mayor o menor, cambiar el orden de las dimensiones, conocer la consulta MDX exacta que se está mostrando, entre otros.

## 6.2 Construcción de Reportes

El objetivo de crear reportes sobre el Data Mart es obtener los indicadores de participación diseñados. Para esto se utilizó el software iReport Designer, este permite construir reportes que, al igual que los esquemas XML de los cubos OLAP, deben ser cargados a JasperReports Server, desde donde pueden ser ejecutados.

iReport Designer entrega las herramientas para dar formato a un reporte, crear campos de datos variables, gráficos, entre otros, los que son llenados con información del Data Mart, información que se obtiene por medio de consultas SQL.

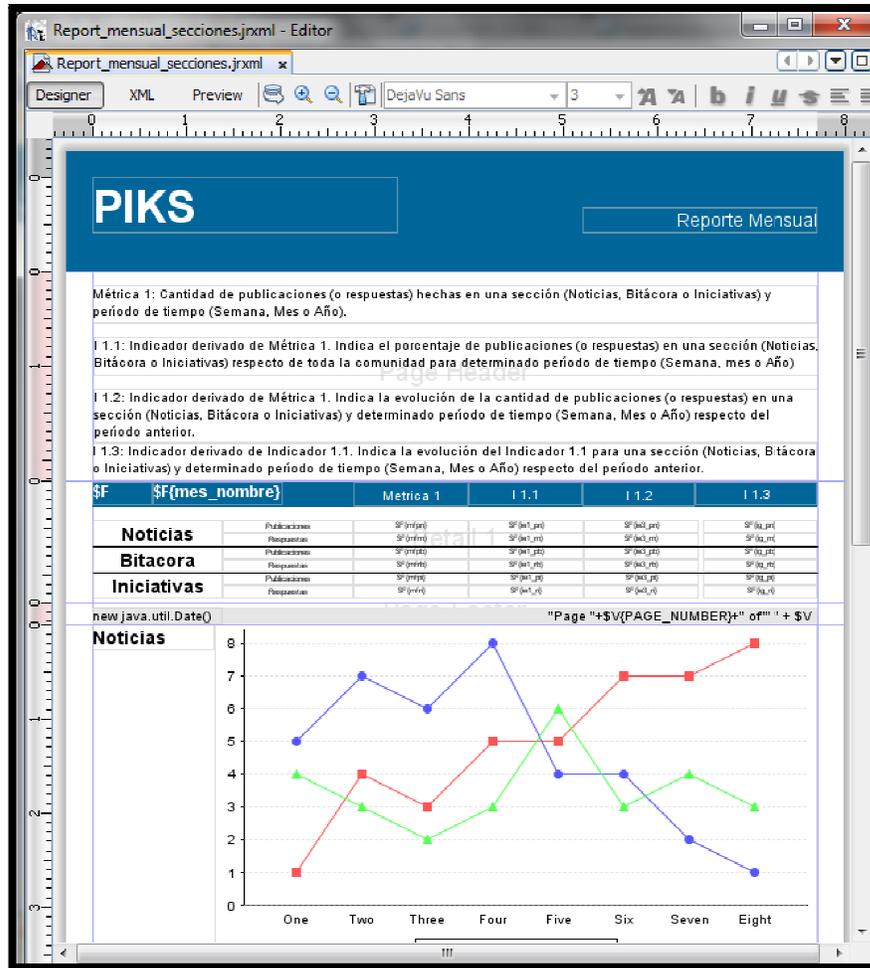


Figura 14: Diseño de reporte mensual para las 3 secciones.  
Fuente: Imagen de iReport Designer. Elaboración Propia.

En la figura anterior se puede apreciar como se le da formato a un reporte. Se crean campos con sus respectivas etiquetas que hacen referencia a las columnas del resultado de la consulta SQL que se ejecuta por detrás y también se agregan y configuran los gráficos para ser desplegados. El resultado que se obtiene puede observarse en la siguiente figura que representa una muestra de un reporte:

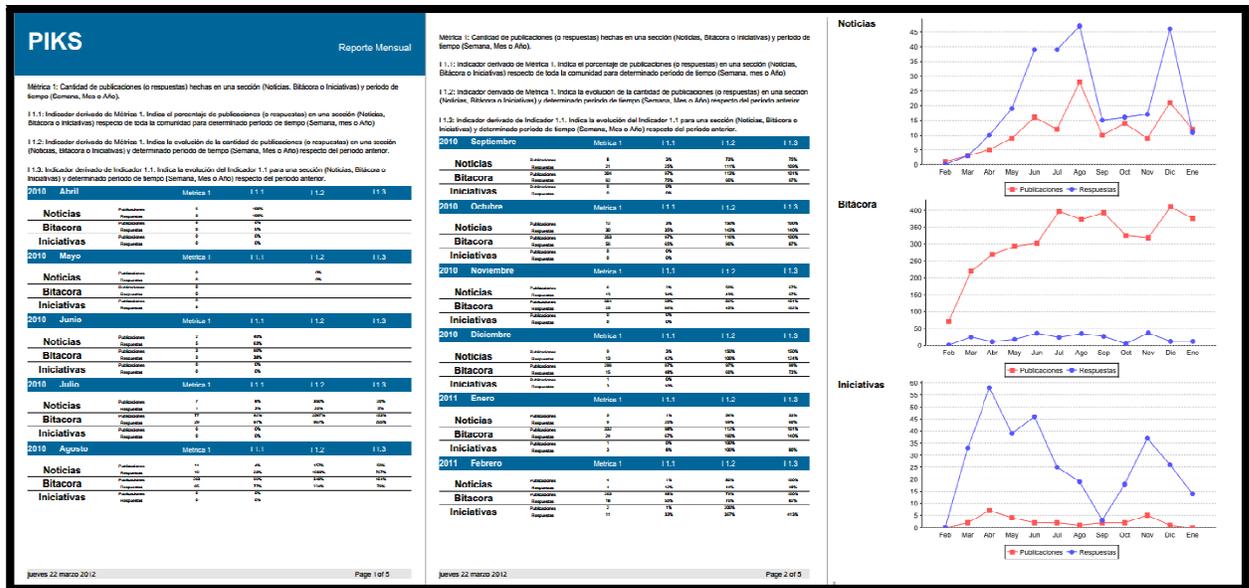


Figura 15: Ejemplo de reporte.  
Fuente: Imagen de iReport Designer. Elaboración Propia.

Como se ve en la imagen anterior, el reporte que se genera, va desplegando los campos de la consulta SQL según el formato que se le otorgó, esto permite conocer el valor de los indicadores de participación y poder concluir al respecto. Estos reportes pueden ser exportados a diversos formatos de archivos como PDF o Excel, entre otros.

Dada la cantidad de métricas e indicadores diseñados, se construyeron 27 reportes: 9 anuales, 9 mensuales y 9 semanales. Cada uno de estos tres sets de 9 reportes se compone de lo siguiente:

- 1 Reporte Anual/Mensual/Semanal para las tres secciones en conjunto con la métrica PST y los indicadores IPST, VPST y VIPST calculados para las publicaciones y para las respuestas. Este reporte incluye gráficos.
- 1 Reporte Anual/Mensual/Semanal para cada sección (por lo tanto, 3 reportes diferentes) con la métrica PST y el indicador IPST calculados para las publicaciones y para las respuestas según cada usuario (ranking).
- 1 Reporte Anual/Mensual/Semanal para cada sección (por lo tanto, 3 reportes diferentes) con la métrica PST y el indicador IPST calculados para las publicaciones y para las respuestas según cada área (ranking).
- 1 Reporte Anual/Mensual/Semanal para la sección Bitácora con la métrica PST y el indicador IPST calculados para las publicaciones y para las respuestas según cada centro (ranking).
- 1 Reporte Anual/Mensual/Semanal con la métrica EA y el indicador EA calculados. Este reporte incluye gráficos.

De esta manera, se obtiene un completo set de reportes con todas las métricas e indicadores diseñados en el capítulo 4 para ser consultados mediante JasperReports Server. La interfaz de este software es la siguiente:

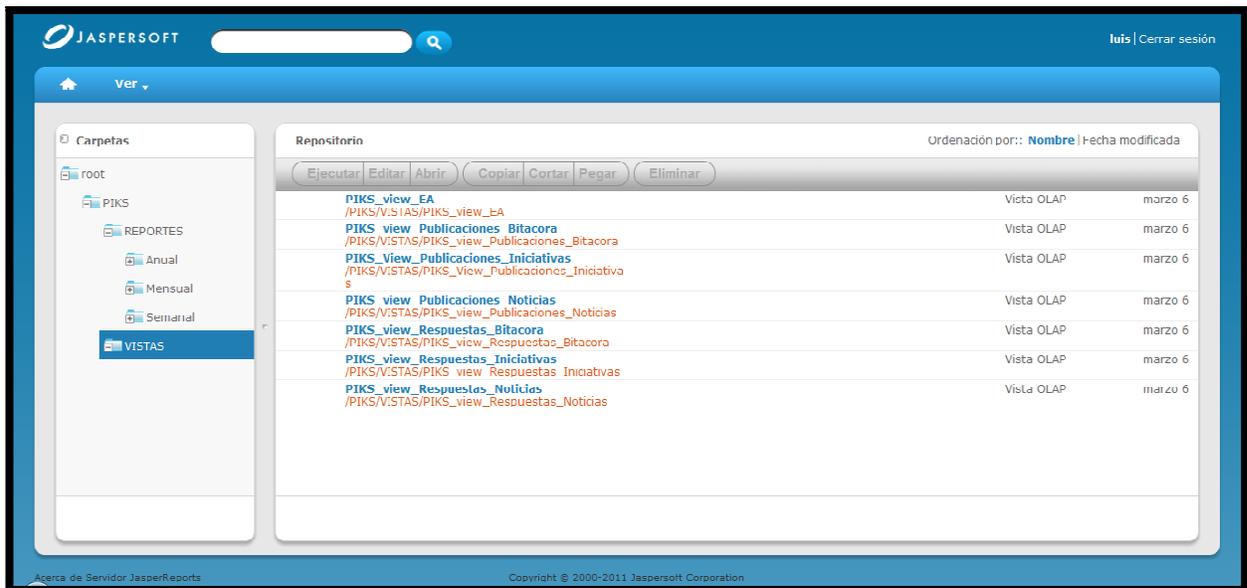


Figura 16: Interfaz de JasperReports Server.

Fuente: Elaboración propia.

Se conectó el servidor a Internet para que pueda ser probado por los administradores. La interfaz a la que se accede al conectarse es la que se observa en la figura. A la izquierda se encuentra un menú para seleccionar el tipo de recurso que se desee consultar: reportes separados según períodos de tiempo o vistas a los cubos OLAP diseñados en la sección 6.1.

Mediante el acceso a este sistema se pudo evaluar la apreciación de los administradores, proceso que se explica en el siguiente capítulo.

Dado lo anterior, la solución implementada posee una arquitectura de dos capas: En una primera capa se instaló tanto el Data Mart (almacenamiento) como el software JasperReports Server (aplicación), mientras que desde otra capa era posible acceder a la información (capa de presentación) por medio de la conexión de la primera capa a Internet.

Esta arquitectura puede ser extendida a tres capas llevando el almacenamiento y la aplicación a capas diferentes, de esta manera el sistema resulta escalable y con mayor nivel de seguridad.

## Capítulo 7

# Diseño y Aplicación de Evaluación Cualitativa

Para evaluar la utilidad de la solución diseñada se construyó una encuesta, la que permite medir cualitativamente la apreciación de los administradores. El objetivo es conocer si el diseño de las métricas e indicadores y la implementación realizada satisfacen las necesidades de los administradores en cuanto a la obtención de información sobre la participación de los usuarios de la plataforma, además de concluir sobre la hipótesis de trabajo.

### 7.1 Diseño de encuesta

El diseño de la encuesta llevada a cabo se fundamenta principalmente en [25]. El objetivo es evaluar la calidad de la solución por medio de la medición de la satisfacción que tengan los administradores acerca de esta. En otras palabras, conocer que tan útil les resulta la solución diseñada para enfrentar el problema de conocer la participación de PIKS y poder gestionarla.

En base a lo antes mencionado, se debe fragmentar lo que se desea medir, es decir, definir las diferentes dimensiones relevantes y evaluar cada una, para luego concluir sobre la evaluación general de la solución propuesta, ya que evaluar un aspecto intangible, como lo es la calidad o utilidad de un servicio, es complejo de realizar.

Para esto se deben comparar las expectativas preliminares con las actuales, con tal de comprender el nivel de satisfacción, y así, a partir de esta medición, poder concluir sobre la calidad o utilidad del servicio [26].

Entonces, se quiere conocer la utilidad de la solución, la cual está compuesta del diseño de métricas e indicadores de participación que se representan mediante un sistema construido a partir de las herramientas que ofrece Jaspersoft. Este sistema

incluye las vistas sobre cubos OLAP de las métricas almacenadas en el Data Mart y la generación de reportes con los indicadores de participación definidos.

Esta utilidad será el concepto general que se quiere conocer por medio de la encuesta. Pero como resulta complejo medirlo directamente, se debe dividir cada concepto en subconceptos hasta que puedan traducirse a preguntas y respuestas simples y que permitirán conocer el nivel de satisfacción específico que poseen los encuestados sobre cada punto. Aplicando este procedimiento, se obtienen los siguientes conceptos y subconceptos que se medirán:

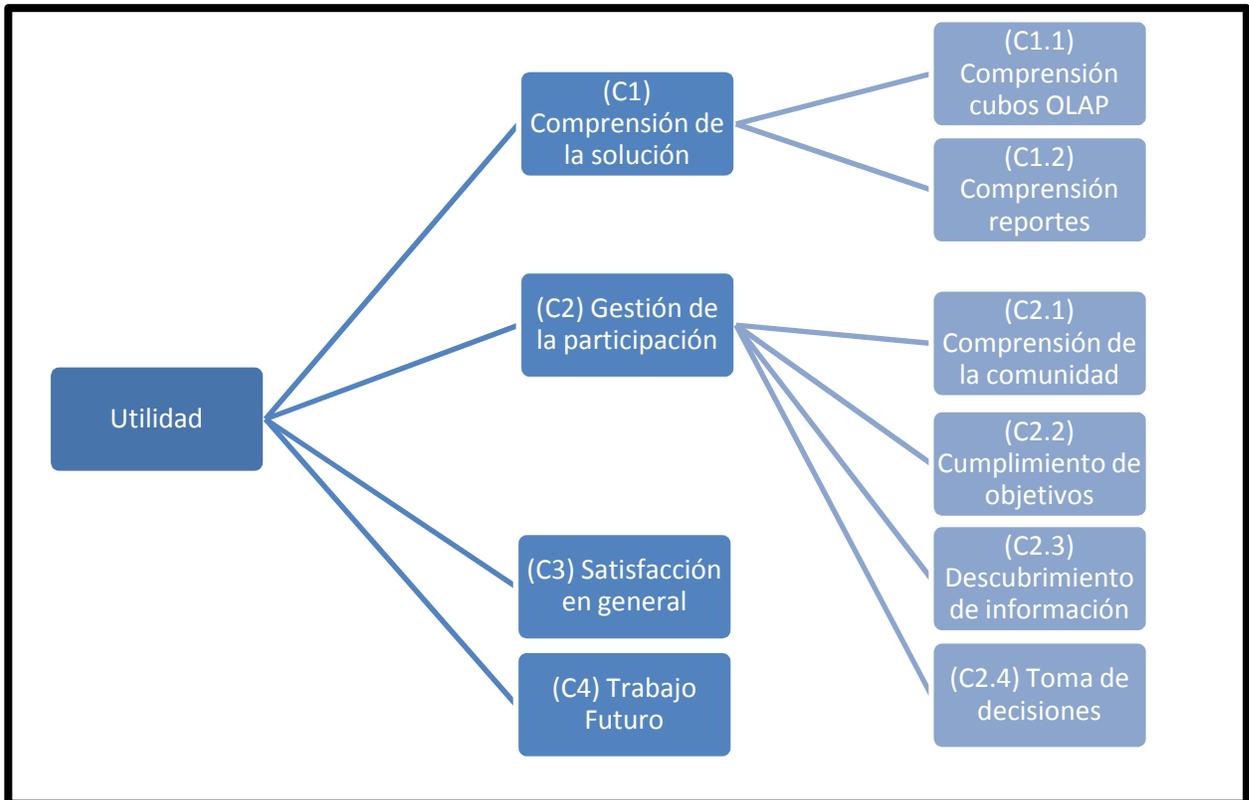


Figura 17: Conceptos y subconceptos a medir en la encuesta.

Fuente: Elaboración propia.

Como se observa, el concepto base “Utilidad” se descompone en cuatro subconceptos, dos de los cuales también se dividen, resultando finalmente en ocho subconceptos.

Teniendo definidos los conceptos a evaluar, se deben establecer preguntas que permitan cuantificar las opiniones o declaraciones de los encuestados, es por esto que las respuestas deben ser preestablecidas, pues, como se recomienda en [25] resulta mejor de esta manera, ya que facilita la tarea del encuestado y el análisis posterior de los resultados.

Dado esto, se utilizaron preguntas de respuesta ordenada, es decir, preguntas con alternativas que representan un continuo ordenado de las posibles respuestas. La idea

es que el encuestado seleccione la alternativa que exprese un mayor grado de cercanía con su opinión. Las alternativas que el encuestado podía elegir son: a) Muy en desacuerdo, b) En desacuerdo, c) Indeciso, d) De Acuerdo y e) Muy de acuerdo.

Por otra parte, al utilizar este tipo de respuestas, se recomienda que las preguntas sean redactadas como afirmaciones. Entonces, dadas las condiciones anteriores se definen las siguientes preguntas para cada concepto:

Tabla 12: Conceptos y preguntas de la encuesta.

Concepto	Preguntas
(C1.1)	Los cubos son intuitivos de manejar y de comprender.
(C1.2)	Los reportes son intuitivos, la información se despliega amigablemente.
(C2.1)	La información entregada por el sistema permite evaluar la participación de los usuarios en la comunidad.
	La información entregada por el sistema permite conocer la evolución de la participación de la comunidad
	La información entregada por el sistema permite obtener una mejor visibilidad de la participación comparado con otros métodos para hacer gestión de la comunidad que se usaran antes.
(C2.2)	La información entregada por el sistema permite conocer el cumplimiento de los objetivos de la comunidad en cuanto a la participación de los usuarios que se espera.
(C2.3)	La información entregada por el sistema permite descubrir nueva información para hacer gestión de la comunidad, como por ejemplo: Inferir conclusiones sobre la participación, desarrollar nuevas políticas/estratégicas para fomentar participación, entre otros.
(C2.4)	En base a la información y conclusiones que pueden obtenerse sobre la participación, podrían tomarse decisiones que conduzcan a una mejora de la participación en la comunidad.
	En base a la información y conclusiones que pueden obtenerse sobre la participación, podrían tomarse decisiones, que fomenten la participación, con mayor facilidad que antes.
	En base a la información y conclusiones que pueden obtenerse sobre la participación, podrían tomarse mejores decisiones que antes.
(C3)	La información entregada sobre la participación en la comunidad representa una manera formal y objetiva de visualizarla.
	La información entregada por el sistema es suficientes para iniciar cambios que conlleven a mejoras de la comunidad.
	En general, la información entregada por el sistema es útil para hacer gestión de la comunidad
(C4)	¿Cuál debería ser el siguiente paso a realizar en la plataforma PIKS en relación al sistema presentado?
	¿Cree que existen otras variables (o información relevante) que deberían considerarse para el diseño de las Métricas e Indicadores?, ¿Cuáles?

Fuente: Elaboración propia.

El set de respuestas preestablecidas se aplicó para responder a todas las preguntas, a excepción de las últimas dos, que corresponden al concepto 4, en las que no se utilizaron alternativas, sino que fueron preguntas abiertas, es decir, el encuestado

no tiene restricciones para responder. Esto se debió a que la naturaleza de las preguntas no permitía estructurarlas.

En conclusión, la aplicación de esta encuesta<sup>13</sup> permite traducir conceptos complejos relacionados con la utilidad del sistema implementado a preguntas y respuestas simples, se espera que los encuestados expresen su percepción acerca de estos conceptos y así se puedan medir sus opiniones acerca de las métricas presentadas en los cubos OLAP y los indicadores de participación presentados en los reportes. Con lo que finalmente además se concluirá sobre la hipótesis.

## 7.2 Resultados

La encuesta diseñada fue aplicada a tres administradores, donde además uno de ellos es Director de la empresa FRIOSUR y desarrollador de la plataforma, otro es Jefe del proyecto PIKS y el tercero también es programador de la plataforma. Los resultados obtenidos son los siguientes<sup>14</sup>:

Tabla 13: Resultados de la encuesta.

Pregunta	Respuesta	Muy en desacuerdo	En desacuerdo	Indeciso	De acuerdo	Muy de acuerdo
1.1				1	2	
1.2				1	2	
2.1						3
2.2					3	
2.3					1	2
2.4				1	2	
2.5						3
2.6					2	1
2.7					2	1
2.8					1	2
3.1					2	1
3.2					1	2
3.3						3
4.1	-“Desarrollar PIKS Analytics”. -“Transportarlo en aplicación”. -“Integrarlo a la plataforma como una aplicación más”.					
4.2	-“Algunos que permitan comparar evolución de usuarios y entre ellos”. -“Ratio EE.AA. (abren/cierran)”. -“Ranking de participación de usuario”. -“Cuerpo de la información al hacer clic en el número”. -“No, creo que están las correctas”.					

Fuente: Elaboración propia.

<sup>13</sup> El diseño final de la encuesta aplicada puede verse en el Anexo C.1

<sup>14</sup> Los resultados individuales se encuentran en el Anexo C.2

### 7.2.1 Análisis de resultados

La encuesta fue diseñada con el propósito de medir la utilidad que entrega la solución propuesta y para esto se definieron conceptos para evaluarlos individualmente y a partir de estos obtener una conclusión general.

**(C1) Comprensión de la solución:** Este concepto se dividió en dos sub conceptos (C1.1 y C1.2) que se evaluaron por medio de las preguntas 1.1 y 1.2, respectivamente, preguntas que obtuvieron las mismas respuestas. Los resultados indican que si bien los cubos y reportes se logran entender, falta desarrollarlos para que su despliegue sea más intuitivo, fácil de comprender y no represente una dificultad al momento de tratar de obtener información.

**(C2) Gestión de la participación:** Se dividió en los siguientes conceptos.

- **(C2.1) Comprensión de la comunidad:** Este concepto se evaluó por medio de las preguntas 2.1, 2.2 y 2.3. En general, fueron respondidas con valoraciones positivas, esto indica que la información entregada por la solución permite conocer la evolución y evaluar la participación de los miembros, y además representa una mejor manera de visualizar los niveles de participación que los métodos que se solían utilizar.
- **(C2.2) Cumplimiento de objetivos:** Se evaluó con la pregunta 2.4. Los resultados según dos encuestados indican que el sistema si sirve para medir el cumplimiento de objetivos, pero el tercero responde como indeciso. Por lo tanto, se puede deducir que falta mejorar este aspecto, lo que podría realizarse mediante un rediseño del despliegue de la información.
- **(C2.3) Descubrimiento de información:** Los resultados (pregunta 2.5) indican que la solución diseñada es muy útil para obtener conclusiones acerca de la participación y desarrollar nuevas estrategias que permitan fomentar la participación, ya que los tres encuestados respondieron con la máxima valoración. Con esto se tiene que entonces la solución si sirve para hacer gestión de la comunidad.
- **(C2.4) Toma de decisiones:** Las preguntas definidas para medir este concepto son la 2.6, 2.7 y 2.8. Las respuestas de los encuestados para estas tres preguntas fueron solamente valoraciones positivas (De acuerdo y Muy de acuerdo). Esto implica que conociendo los niveles de participación, los encuestados opinan que si se pueden tomar decisiones que fomenten el desarrollo de la comunidad y además pueden mejores decisiones y con más facilidad que antes.

Por lo tanto, para el concepto Gestión de la participación, que engloba estos 4 sub conceptos analizados, se tiene que el sistema diseñado es un aporte positivo a la gestión que deben realizar los administradores, pues,

las métricas e indicadores, entregan información valiosa sobre lo que ocurre con la participación de la comunidad.

**(C.3) Satisfacción en general:** Las tres preguntas que se refieren a este concepto (3.1, 3.2 y 3.3) fueron respondidas con valoraciones positivas por los tres encuestados. De esta manera, se tiene que: el diseño propuesto entrega información formal y objetiva, esta información es suficiente para intentar realizar cambios en la comunidad y es útil para hacer gestión. Con esto se puede deducir que entonces los encuestados se encuentran satisfechos con la solución propuesta, ya que al responder con valoraciones positivas indica que su percepción del nivel de calidad o utilidad del diseño se asemeja a las expectativas que tenían antes de conocerlo.

**(C.4) Trabajo futuro:** Este concepto tenía como propósito conocer qué pasos recomiendan continuar los encuestados y, dependiendo de sus respuestas, también podría obtenerse una conclusión sobre la solución. Las respuestas a la pregunta 4.1 sugieren que el sistema debe incorporarse formalmente a la plataforma y algunas de las respuestas a la pregunta 4.2 sugieren aplicar indicadores diseñados sobre otras métricas, con tal de conocer mejor la participación. Entonces, al sugerir que la herramienta sea formalizada e incorporada a la plataforma y que se desarrollen más métricas, implica que la propuesta si cumplió con lo esperado y por lo tanto lo lógico es que continúe su desarrollo.

Finalmente, dadas las conclusiones que se obtiene del análisis de cada concepto consultado, se puede inferir que los diseños planteados e implementados si son útiles para los administradores. Esto indica que la propuesta, al entregar una mejor visibilidad de los niveles de participación, ayuda a mejorar la labor de administrar para fomentar la participación.

## Capítulo 8

# Conclusión

Para concluir sobre el cumplimiento del objetivo general del proyecto se repasará cada paso realizado durante el desarrollo del trabajo.

El primer objetivo específico consistió en realizar una investigación sobre el estado del arte de los conceptos que permitirían abordar el problema, para esto se estudió acerca de Comunidades Virtuales de Práctica, Análisis de Redes Sociales, Teoría de grafos, Teoría para el diseño de métricas e indicadores, metodologías para diseñar repositorios de datos y las tecnologías existentes para implementar la solución que se planteara. De este objetivo se esperaba obtener una metodología para diseñar las métricas e indicadores que permitirían visualizar formalmente los niveles de participación de la comunidad, lo que consiguió y está plasmado en el capítulo 2 y 3.

El segundo objetivo específico era diseñar, implementar y evaluar métricas e indicadores para el seguimiento y control de una Comunidad Virtual de Práctica. De este objetivo se esperaba como resultado un set de métricas e indicadores para medir la participación de los miembros de la comunidad, lo que se logró exitosamente, pero cabe mencionar que aún pueden diseñarse otras métricas o indicadores acordes a nuevas necesidades de información que puedan surgir.

Para desarrollar esto se utilizó la metodología obtenida del primer objetivo específico. Se estudió y conversó con los administradores para decidir qué métricas e indicadores aportarían mayor valor para hacer gestión y se definieron formalmente. Con este resultado, y con las métricas sociales calculadas, se pudo analizar la situación actual de la comunidad y concluir que se necesita un mecanismo formal para hacer gestión, ya que la mera apreciación de los administradores sobre lo que ocurre en la comunidad suele diferir de la realidad y por lo tanto no es suficiente para realizar esa labor. Además, se pudo comprender objetivamente que la a la comunidad le falta por desarrollarse, pues, son escasos los usuarios que participan constantemente y son ellos quienes poseen la mayor parte de las interacciones existentes.

Dada la metodología utilizada para el diseño, los mismos indicadores construidos permiten evaluar la participación, ya que en el conjunto de indicadores diseñados

existen los elementales para evaluar las métricas y los globales para evaluar a otros indicadores, con esto se puede conocer cómo evoluciona la participación de la comunidad.

Para la implementación de las métricas e indicadores, y poder concluir con el segundo objetivo, se debió antes desarrollar el tercer objetivo específico.

Este objetivo consiste en diseñar y construir un Repositorio Centralizado de Datos, para esto se diseñó un Data Mart siguiendo el modelo Constelación, es decir, un grupo de tablas de hechos que comparten dimensiones. Este Data Mart fue poblado con las interacciones de los usuarios. Para ello se diseñó un proceso ETL consistente con la estructura del Data Mart que extrajo los datos relevantes, los transformó y cargó en el repositorio.

Una vez que el Data Mart se desarrolló se pudo realizar la implementación de las métricas e indicadores. Para esto se crearon esquemas de cubos OLAP para visualizar las métricas de cada tabla de hechos realizando cruces con las dimensiones y poder consultar distintos niveles de agregación. Además se construyeron reportes para desplegar el cálculo de los indicadores. Finalmente, fueron implementados en JasperReports Server para que puedan ser consultados. Con esto se concluye el segundo objetivo específico.

El último objetivo específico correspondía a evaluar la propuesta de diseño en términos cualitativos, para realizar esto se dio acceso a los administradores para conocer la implementación de la solución y luego fueron encuestados. Los resultados de la encuesta fueron favorables y permitieron concluir que la herramienta si es un aporte para la gestión de la comunidad, ya que permite tomar decisiones en base a información objetiva y formal.

Desarrollados exitosamente todos los objetivos específicos propuestos, es posible concluir sobre el objetivo general del proyecto.

El objetivo era: “Diseñar e implementar métricas e indicadores para asegurar el seguimiento y control de las interacciones de los usuarios de una Comunidad Virtual de Práctica para fomentar la generación e impacto de proyectos de innovación”

Se consiguió diseñar e implementar métricas e indicadores que midan las interacciones de los usuarios de una Comunidad Virtual de Práctica y, por la evaluación cualitativa hecha por los administradores mediante la encuesta, se puede concluir que estas métricas e indicadores sirven para asegurar el seguimiento y control de los usuarios, esto pues, los administradores expresaron que permiten comprender lo que ocurre en la comunidad al entregarles la información sobre la evolución de cada usuario y de la comunidad en general.

Entonces, al conseguirse un seguimiento formal y objetivo de la participación, y además lograr que este seguimiento sea útil para hacer gestión de la comunidad y la toma decisiones, es posible concluir que con estos resultados obtenidos los administradores podrán inducir cambios que conduzcan a mejorar la participación de cada miembro y con esto se logre fomentar una mayor generación de proyectos de innovación y un mayor impacto de cada proyecto que se plantee en términos de mejorar las discusiones que en ellos produzca.

Por otra parte, la hipótesis del trabajo planteada al comienzo fue: “La gestión de la participación puede mejorarse a través de la mejor visibilidad que aportaría la creación de métricas e indicadores de los niveles de participación y que permitiría tomar decisiones a los administradores de la comunidad u otros miembros clave.”

Como se ha mencionado, se logró medir los niveles de participación objetivamente y mediante la implementación de las métricas e indicadores se pudo mostrar esto a los administradores. Respecto de esto, los administradores expresaron su opinión en términos de la apreciación de la utilidad de la solución y se concluyó que sí son útiles al haberse obtenido como resultado de la encuesta que sus expectativas quedaron satisfechas con el diseño. A partir de esto entonces, también se pudo inferir que la solución permite mejorar la gestión que se hace en la comunidad, ya que entrega información que no se conocía, y por lo tanto la herramienta desarrollada sirve para tomar decisiones, desarrollar estrategias e inducir cambios en la comunidad que conduzcan a mejoras. Con esto, queda cumplida la hipótesis de trabajo.

Como conclusiones generales del proyecto puede mencionarse lo siguiente:

- Las observaciones que los administradores pueden hacer de la participación de la comunidad, sin contar con un mecanismo formal que la mida, no entregan suficiente información para adoptar una estrategia que permita un desarrollo óptimo de la comunidad, pues, existen aspectos que no pueden observarse a simple vista como los que entrega el análisis de redes sociales.
- Se necesita de un mecanismo que permita administrar formalmente una comunidad virtual, ya que no basta solamente con instaurarla y esperar a que lo usuarios participen, es decir, el hecho de que exista la plataforma no es suficiente incentivo a participar, por lo tanto los administradores deben recurrir a implementar sistemas que le entreguen información útil que les permita comprender el estado en el que se encuentra la comunidad que administran.
- Los administradores requieren de una herramienta como la desarrollada en este trabajo para descubrir nueva información que antes no percibían, ya que la información que pueden obtener les permite tomar decisiones en base a hechos observables objetivamente y no en base a apreciaciones subjetivas que ellos tengan del estado de la comunidad.

Finalmente, cabe destacar que el diseño desarrollado, si bien fue hecho en base a la comunidad creada por la empresa FRIOSUR, las métricas e indicadores pueden ser aplicados a cualquier otra red, ya que la metodología con la que fueron creados puede adaptarse para resolver otros problemas de medición de información.

### 8.1 Trabajo Futuro

El trabajo futuro que se dejará planteado es el siguiente:

- Los registros que se almacenan en la base de datos no guardan todas las actividades que ocurren en la plataforma, como por ejemplo, los registros de estados de ánimo con que se postea. Si esto es implementado, será posible realizar un análisis para buscar alguna correlación entre el estado de ánimo y la cantidad de *posts*.
- Al no existir weblogs de la actividad no se puede conocer la participación pasiva que originan principalmente los usuarios periféricos de la comunidad, es decir, usuarios que ingresen a la comunidad para informarse pero que no comentan, lo que también sería importante y útil de medir, ya que permitiría tomar decisiones que conduzcan a que este tipo de usuario participe activamente.
- De la encuesta puede sugerirse que el paso siguiente que debe tomarse es implementar la solución diseñada formalmente como un módulo de la plataforma.
- Por último, sería de utilidad poder conocer si la cantidad e impacto de propuestas factibles generadas por los usuarios aumenta debido a la mejor gestión de la participación dentro de la comunidad que se puede hacer por la información que entrega la solución diseñada. Para esto se debe medir el impacto que causen las decisiones que tome la administración y luego realizar un análisis que permita concluir.

# Bibliografía

- [1] Wenger, E.; McDermontt R. y Snyder, W. M. 2002. Cultivating communities of practice: a guide to managing knowledge. Harvard Business Press.
- [2] Hagel, J. 1999. Net gain: expanding markets through virtual communities. *Journal of interactive marketing* 13(1):55-65.
- [3] Kim, W.; Jeong, O. y Lee, S. 2010. On social Web sites. *Information Systems*, 35:215-236.
- [4] Vignolo, C.; Potocnjak, C. y Ramirez, A. 2003. El desarrollo como un proceso conversacional de construcción de capital social. *Ingeniería de sistemas*. 17(1).
- [5] Li, H. y Lai, V. S. 2007. Interpersonal relationship needs of virtual community participation: A FIRO perspective. *Proceedings of the Americas Conference on Information Systems*, Colorado, United States, August 9-12.
- [6] Ardichvili, A.; Page, V. y Wentling, T. 2002. Motivation and barriers to participation in virtual knowledge-sharing communities of practice. Presented at the OKLC 2002 Conference. Athens, Greece, april.
- [7] Chiu, C. M. ; Hsu, M. H. y Wang, E. 2006. Understanding knowledge sharing in virtual communities: An integration of social capital and social cognitive theories. *Decision Support Systems*. 42:1872-1888.
- [8] Wellman, B. 1996. An electronic group is virtually a social network. *Culture of the Internet*, ed.
- [9] Wasserman, S. y Faust, K. 1994. *Social Network Analysis: Methods and Applications*. Cambridge University Press.
- [10] Molina, L. 2001. *El Análisis de Redes Sociales: Una Introducción*. Editions Bellaterra, Barcelona.
- [11] Mika, P. 2006. *Social Networks and the Semantic Web*.
- [12] Olsina, L.; Papa, F. y Molina, H. 2008. How to Measure and Evaluate Web Applications in a Consistent Way. *Modelling and Implementing Web Applications*. Chapter 13th, In Springer HCIS, pp. 385-420.
- [13] Weber, R. 2010. Cátedras curso IN643 "Introducción a la Minería de Datos". Semestre otoño.

- 
- [14] Chapman, P.; Clinton, J.; Kerber, R.; Khabaza, T.; Reinartz, T.; Shearer, C. y Wirth, R. 2000. SPSS. CRISP-DM 1.0 Step-by-step data mining guide. The CRISP-DM Consortium/SPSS Inc
- [15] Ríos, S. A. y Aguilera, F. 2010. Web Intelligence on the Social Web. Advanced Techniques in Web Intelligence.
- [16] Saint-onge, H. y Wallace, D. 2003. Leveraging Communities of Practice for Strategic Advantage.
- [17] Inmon, W. 2005. Building the Data Warehouse. Cuarta edición, Wiley Publishing, Inc., Indianapolis, Indiana.
- [18] Kimball, R.; Reeves, L.; Ross, M. y Thornthwaite, W. 1998. The data Warehouse Lifecycle toolkit/ Expert Methods for designing, developing, and deploying data Warehouse. publicado por John Wiley and Sons, Inc.
- [19] Kimball, R, y Ross, M. 2002. The data Warehouse Toolkit: The complete guide to multidimensional modeling. Segunda edición, publicado por John Wiley and Sons, Inc.
- [20] Chenoweth, T.; Schuff, D. y St. Louis. R. 2003. A method for developing dimensional data marts. Communications of the ACM. 46(12):93-98.
- [21] Gyssens, M. y Lakshmanan, L. 1997. A Foundation for Multi-Dimensional Databases. Proceedings of the 23rd International Conference on Very Large Data Bases. Morgan Kaufmann Publishers Inc. San Francisco, CA, USA.
- [22] Chaudhuri, S. y Dayal, U. 1997. An Overview of DataWarehousing and OLAP Technology. 26(1).
- [23] Bustos, F. 2011. Diseño e implementación de una framework para llevar sitios web sociales a la web semántica y que facilite el Análisis de Redes Sociales. Memoria para optar al título de Ingeniero Civil Computación e Ingeniero Civil Industrial. Santiago, Universidad de Chile, Facultad de ciencias físicas y matemáticas.
- [24] Teas, R. 1993. Expectations, Performance Evaluation, and Consumers' Perceptions of Quality. Journal of Marketing. 57(4):18-34.
- [25] Asún, R. 2006. Construcción de cuestionarios y escalas: El proceso de la producción de información cuantitativa. En Canales, M. (ed) Metodologías de investigación social. Santiago de Chile.

- 
- [26] Krivobokova, O. 2009. Evaluating Customer Satisfaction as an Aspect of Quality Management. *World Academy of Science, Engineering and Technology*. 53:565-568
- [27] Alvarez, H. 2010. Detección de miembros claves en una comunidad virtual de práctica mediante análisis de redes sociales y minería de datos avanzada. Tesis para optar al grado de magíster en Gestión de Operaciones y al título de Ingeniero Civil Industrial. Santiago, Universidad de Chile, Facultad de ciencias físicas y matemáticas.
- [28] Manning, C. D.; Raghavan, P. y Schütze, H. 2009. *Introduction to Information Retrieval*. Cambridge University Press.
- [29] Kimball, R. y Caserta, J. 2004. *The data warehouse ETL Toolkit: Practical Techniques for Extracting, Cleaning, Conforming, and Delivering Data*. Wiley Publishing, Inc.
- [30] Kleinberg, J. 1999. Authoritative sources in a hyperlinked environment.
- [31] Moody, D. L y Kortink, Mark A. R. 2000. From Enterprise Models to Dimensional Models: A Methodology for Data Warehouse and Data Mart Design. *Proceedings of the Second Intl. Workshop on Design and Management of Data Warehouses, DMDW 2000, Stockholm, Sweden*.
- [32] Scheneider, M. Well-formed Data Warehouse Structures. 2003. *Design and Management of Data Warehouses 2003, Proceedings of the 5th Intl. Workshop DMDW'2003, Berlin, Germany*.
- [33] Elmasri, R. y Navathe, S. 1999. *Fundamentals of Database Systems*. Publicado por Addison-Wesley 3° ed.
- [34] Negash, S. y Gray, P. 2004. Business Intelligence. *Communication of the Association for Information Systems*. 13:177-195.
- [35] Watson, H. J. y Wixom, B. H. 2007. The Current State of Business Intelligence. *Computer*, 40(9):96–99.
- [36] Lave, J. y Wenger, E. 1996. *Situated learning: legitimate peripheral participation*. Cambridge University Press.
- [37] Carotenuto, L., et al. 1999. *Communityspace: Toward flexible support for voluntary knowledge communities*.
- [38] Henri, F. y Pudelko, F. 2003. Understanding and analysing activity and learning in virtual communities. *Journal of Computer Assisted Learning*, 19:474–487

- [39] Kooh, J., Kim, Y., Butler, B. y Bock, G. 2007. Encouraging participation in virtual communities. *Communications of the ACM*. 50(2):69-73.

# Anexos

## A Descripción de la Comunidad Virtual de Práctica PIKS

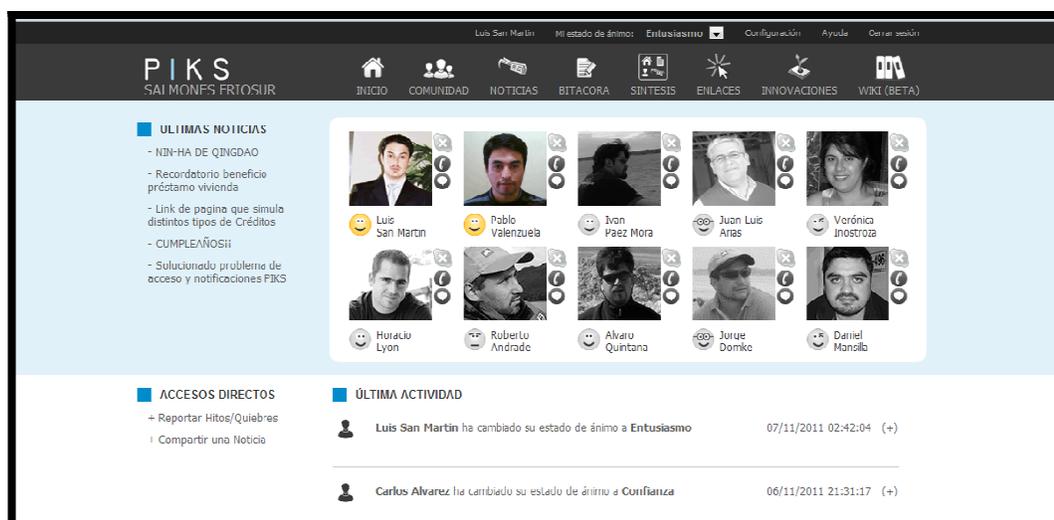
La plataforma presenta los siguientes módulos de interacción entre los usuarios:

- **Noticias:** Cualquier usuario puede publicar y/o comentar una noticia.
- **Bitácora:** Cualquier usuario puede publicar y/o comentar. Pero al publicar una entrada en esta categoría el usuario representa al centro productivo al cual pertenece.
- **Síntesis:** Se puede observar el historial de las publicaciones y el historial de los comentarios que han realizado los usuarios.
- **Iniciativas:** Cualquier usuario puede publicar una Iniciativa, en la que se plantea un quiebre y el cual se espera sea estudiado y comentado por los demás usuarios.
- **Wiki:** Sección para publicar temas en mayor profundidad y documentar proyectos que estuvieron en la sección “Iniciativas”.

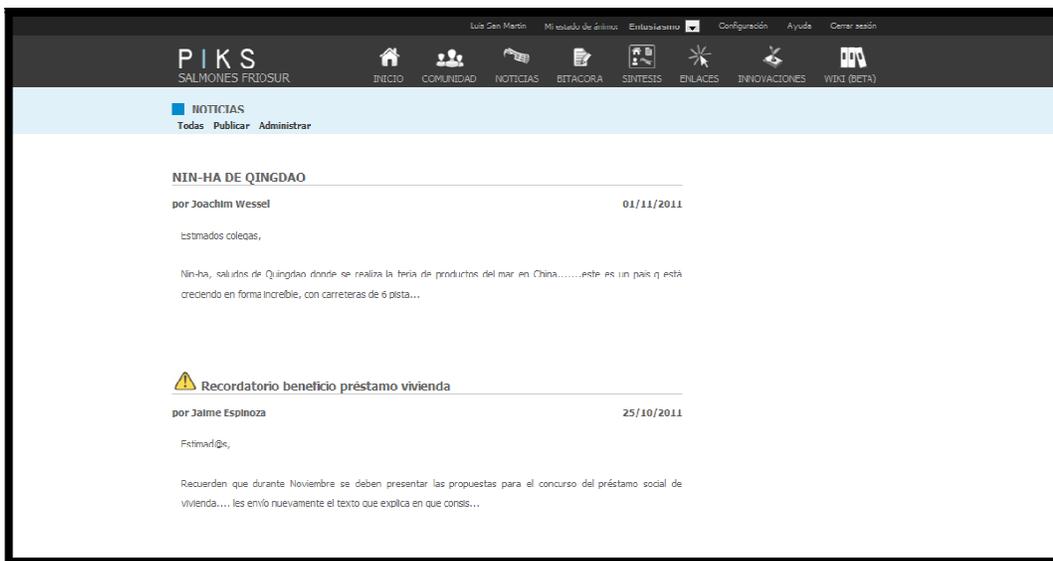
Por lo tanto se tiene que las interacciones que ocurren entre los usuarios a través de la plataforma son por medio de publicaciones y comentarios. Las publicaciones estarán implícitamente dirigidas a todos los usuarios mientras que los comentarios estarán dirigidos a quien creó la publicación e implícitamente a todos los demás.

A continuación se presentan algunas capturas de pantalla de la plataforma:

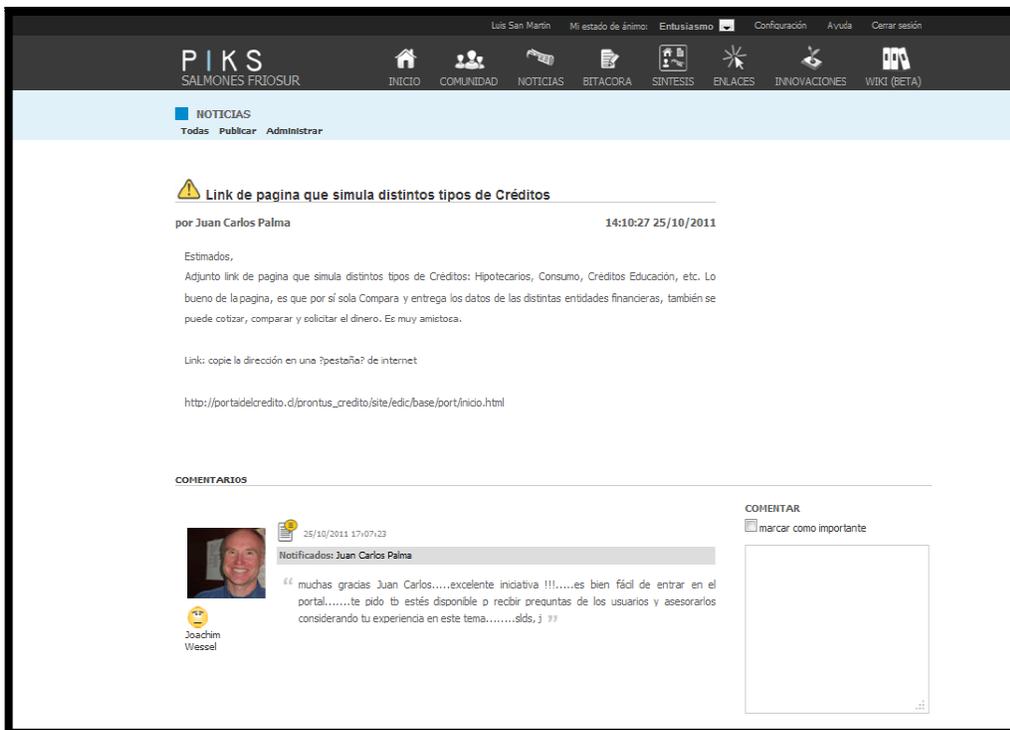
Interfaz de Inicio:



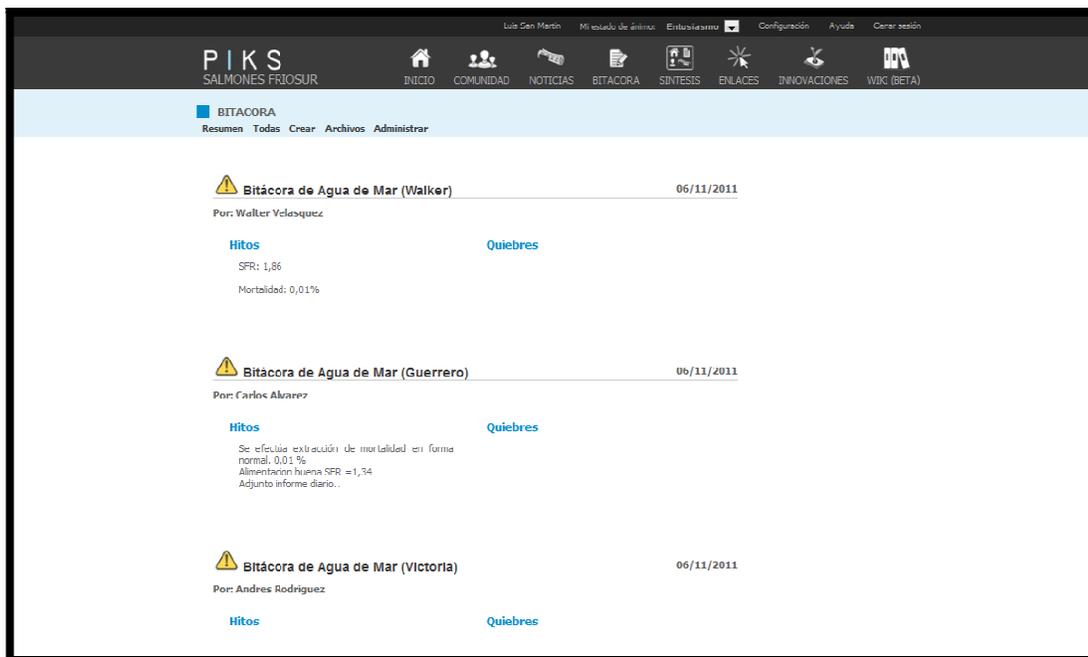
Interfaz de Noticias:



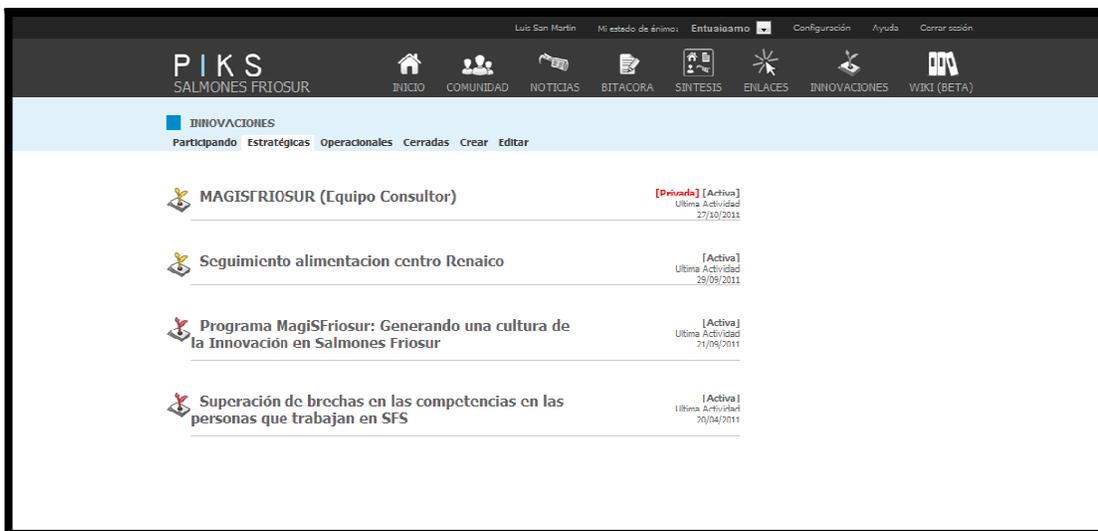
Interfaz dentro de una Noticia:



Interfaz de Bitácora:



Interfaz de Iniciativas:



## B Ejemplo esquema XML de Cubo OLAP

A continuación se presenta un esquema XML, generado por JasperAnalysis Schema Workbench, del cubo de las métricas de publicaciones en la sección Noticias:

```

<Schema name="esquema_cubo_publicaciones_noticias">
  <Cube name="cubo_publicaciones_noticias" cache="true" enabled="true">
    <Table name="fact_publicaciones_noticias">
    </Table>
    <Dimension type="StandardDimension" foreignKey="id_tiempo" name="DimTiempo">
      <Hierarchy name="Periodos" hasAll="true" primaryKey="id_tiempo">
        <Table name="dimtiempo">
        </Table>
        <Level name="Year" column="year" type="Integer" uniqueMembers="true" levelType="Regular"
hideMemberIf="Never">
        </Level>
        <Level name="Mes" column="id_am" type="String" uniqueMembers="true" levelType="Regular"
hideMemberIf="Never">
        </Level>
        <Level name="Semana" column="id_ams" type="String" uniqueMembers="false"
levelType="Regular" hideMemberIf="Never">
        </Level>
      </Hierarchy>
    </Dimension>
    <Dimension type="StandardDimension" foreignKey="id_usuario" name="DimUsuarios">
      <Hierarchy name="Usuarios" hasAll="true" primaryKey="id_usuario">
        <Table name="dimusuarios">
        </Table>
        <Level name="NombreUsuario" column="nombre_apellido" type="String" uniqueMembers="true"
levelType="Regular" hideMemberIf="Never">
        </Level>
      </Hierarchy>
    </Dimension>
    <Dimension type="StandardDimension" foreignKey="id_area" name="DimAreas">
      <Hierarchy name="Areas" hasAll="true" primaryKey="id_area">
        <Table name="dimareas">
        </Table>
        <Level name="NombreArea" column="nombre_area" type="String" uniqueMembers="true"
levelType="Regular" hideMemberIf="Never">
        </Level>
      </Hierarchy>
    </Dimension>
    <Measure name="metrica_publicaciones_noticias" column="metrica" aggregator="sum" visible="true">
    </Measure>
  </Cube>
</Schema>

```

## C Evaluación Cuantitativa y Resultados

### C.1 Encuesta aplicada

#### Encuesta sobre Métricas, Indicadores, Reportes y Cubos sobre PIKS.

**Participación:** Se ha definido a la participación de los usuarios en la comunidad, como la cantidad de *posts* que realizan en un determinado período, es decir, las interacciones que ocurren en la comunidad entre los usuarios, ya sea publicando nuevos contenidos o respondiendo a otros. La Participación se medirá sobre 3 aspectos: Cantidad de Publicaciones, Cantidad de Respuestas y Uso de Estados de Ánimo.

Para presentar la información de la participación de la comunidad se ha desarrollado un **sistema** que incluye la generación y despliegue de **reportes** de métricas e indicadores y la creación de **cubos OLAP** que permiten la visualización de las métricas según requerimientos del usuario.

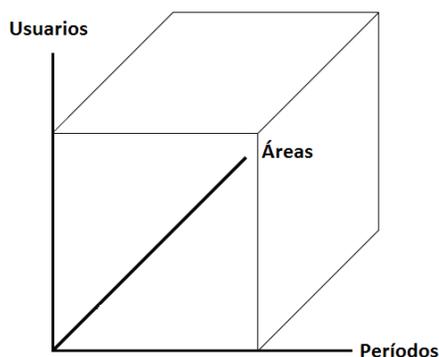
Las Métricas e Indicadores utilizados en los reportes son los siguientes:

- **Métrica 1 o Métrica SP:** Cantidad de publicaciones (o respuestas) hechas en una sección (Noticias, Bitácora o Iniciativas) y período de tiempo (Semana, Mes o Año).
  - **Indicador 1.1 o Indicador SP:** Indicador derivado de Métrica 1. Indica el porcentaje de publicaciones (o respuestas) en una sección (Noticias, Bitácora o Iniciativas) respecto de toda la comunidad para determinado período de tiempo (Semana, mes o Año).
  - **Indicador 1.2 o Indicador SPPA:** Indicador derivado de Métrica 1. Indica la evolución de la cantidad de publicaciones (o respuestas) en una sección (Noticias, Bitácora o Iniciativas) y determinado período de tiempo (Semana, Mes o Año) respecto del período anterior.
  - **Indicador 1.3 o Indicador ISP:** Indicador derivado de Indicador 1.1. Indica la evolución del Indicador 1.1 para una sección (Noticias, Bitácora o Iniciativas) y determinado período de tiempo (Semana, Mes o Año) respecto del período anterior.
- **Métrica 2 o Métrica SPU:** Cantidad de publicaciones (o respuestas) hechas en una sección (Noticias, Bitácora o Iniciativas) y período de tiempo (Semana, Mes o Año), según un usuario, área o centro.
  - **Indicador 2.1 o Indicador SPU:** Indicador derivado de Métrica 2. Indica el porcentaje de publicaciones (o respuestas) de un usuario, área o centro

en una sección (Noticias, Bitácora o Iniciativas), respecto de toda la sección en el mismo período.

- **Métrica 3 o Métrica EA:** Cantidad de actualizaciones de estados que Abren Posibilidades o Cierran Posibilidades hechas durante un período de tiempo determinado (Semana, Mes o Año)
  - **Indicador 3.1 o Indicador EA:** Indicador derivado de Métrica 3. Indica el porcentaje de actualizaciones que Abren Posibilidades hechas en un período de tiempo determinado, respecto de todas las actualizaciones hechas en el período.

Por otra parte, los cubos OLAP implementados permiten conocer las Métricas (definidas anteriormente) y manejar la forma de visualizarlas, es decir, permiten agrupar o desagrupar según Áreas, Usuarios, Centros o Períodos, entre otras opciones.



#### Parte I

**1.1** Los cubos son intuitivos de manejar y de comprender.

Muy en desacuerdo    En desacuerdo    Indeciso    De acuerdo    Muy de acuerdo

**1.2** Los reportes son intuitivos, la información se despliega amigablemente.

Muy en desacuerdo    En desacuerdo    Indeciso    De acuerdo    Muy de acuerdo

#### Parte II

**2.1** La información entregada por el **sistema** permite **evaluar la participación de los usuarios** en la comunidad.

Muy en desacuerdo    En desacuerdo    Indeciso    De acuerdo    Muy de acuerdo

**2.2** La información entregada por el **sistema** permite **conocer la evolución de la participación** de la comunidad

Muy en desacuerdo    En desacuerdo    Indeciso    De acuerdo    Muy de acuerdo

- 2.3** La información entregada por el **sistema** permite **obtener una mejor visibilidad de la participación comparado con otros métodos** para hacer gestión de la comunidad que se usaran antes.  
 Muy en desacuerdo    En desacuerdo    Indeciso    De acuerdo    Muy de acuerdo
- 2.4** La información entregada por el **sistema** permite **conocer el cumplimiento de los objetivos de la comunidad** en cuanto a la participación de los usuarios que se espera.  
 Muy en desacuerdo    En desacuerdo    Indeciso    De acuerdo    Muy de acuerdo
- 2.5** La información entregada por el **sistema** permite **descubrir nueva información para hacer gestión de la comunidad**, como por ejemplo: Inferir conclusiones sobre la participación, desarrollar nuevas políticas/estratégicas para fomentar participación, entre otros.  
 Muy en desacuerdo    En desacuerdo    Indeciso    De acuerdo    Muy de acuerdo  
 Otro: \_\_\_\_\_
- 2.6** En base a la información y conclusiones que pueden obtenerse sobre la participación, **podrían tomarse decisiones que conduzcan a una mejora** de la participación en la comunidad.  
 Muy en desacuerdo    En desacuerdo    Indeciso    De acuerdo    Muy de acuerdo
- 2.7** En base a la información y conclusiones que pueden obtenerse sobre la participación, **podrían tomarse decisiones, que fomenten la participación, con mayor facilidad** que antes.  
 Muy en desacuerdo    En desacuerdo    Indeciso    De acuerdo    Muy de acuerdo
- 2.8** En base a la información y conclusiones que pueden obtenerse sobre la participación, **podrían tomarse mejores decisiones** que antes.  
 Muy en desacuerdo    En desacuerdo    Indeciso    De acuerdo    Muy de acuerdo

### Parte III

- 3.1** La información entregada sobre la participación en la comunidad **representa una manera formal y objetiva de visualizarla**.  
 Muy en desacuerdo    En desacuerdo    Indeciso    De acuerdo    Muy de acuerdo
- 3.2** La información entregada por el **sistema** es **suficientes para iniciar cambios** que conlleven a mejoras de la comunidad.  
 Muy en desacuerdo    En desacuerdo    Indeciso    De acuerdo    Muy de acuerdo

**3.3** En general, la información entregada por el **sistema** es **útil para hacer gestión** de la comunidad  Muy en desacuerdo  En desacuerdo  Indeciso  De acuerdo  Muy de acuerdo

Parte IV

**4.1** ¿Cuál debería ser el siguiente paso a realizar en la **plataforma PIKS** en relación al **sistema** presentado?

\_\_\_\_\_

**4.2** ¿Cree que existen otras variables (o información relevante) que deberían considerarse para el diseño de las Métricas e Indicadores?, ¿Cuáles?

\_\_\_\_\_

## C.2 Resultados de encuesta

Se aplicó a 3 personas:

- Carlos Vignolo: Director de empresa FRIOSUR, Socio desarrollador de PIKS y administrador de la comunidad.
- Luis San Martín: Jefe del proyecto PIKS y administrador de la comunidad.
- Iván Páez: Programador de la plataforma y administrador de la comunidad.

Encuestado Pregunta	Carlos Vignolo	Luis San Martin	Iván Páez
1.1	Indeciso	De acuerdo	De acuerdo
1.2	Indeciso	De acuerdo	De acuerdo
2.1	Muy de acuerdo	Muy de acuerdo	Muy de acuerdo
2.2	De acuerdo	De acuerdo	De acuerdo
2.3	De acuerdo	Muy de acuerdo	Muy de acuerdo
2.4	De acuerdo	De acuerdo	Indeciso
2.5	Muy de acuerdo	Muy de acuerdo	Muy de acuerdo
2.6	De acuerdo	Muy de acuerdo	De acuerdo
2.7	De acuerdo	Muy de acuerdo	De acuerdo
2.8	De acuerdo	Muy de acuerdo	Muy de acuerdo
3.1	Muy de acuerdo	Muy de acuerdo	De acuerdo
3.2	Muy de acuerdo	Muy de acuerdo	De acuerdo
3.3	Muy de acuerdo	Muy de acuerdo	Muy de acuerdo
4.1	Desarrollar PIKS Analytics.	Transportarlo en aplicación.	Integrarlo a la plataforma como una aplicación más.
4.2	Algunos que permitan comparar evolución de usuarios y entre ellos.	-Ratio EE.AA. (abren/cierran). -Ranking de participación de usuario. -Cuerpo de la información al hacer click en el número.	No, creo que están las correctas.