

Trabajo de Grado
Wiki Semántica usando Folksonomies

Diego Torres

Directora: Alicia Díaz

27 de noviembre de 2009

Agradecimientos

Muchas son las personas que me ayudaron a concluir mi formación de grado.

En primer lugar quiero agradecerles a mis papás Denise y Eduardo por darme todo para poder desarrollarme en aquellas cosas que más me gustaron, por la educación que me dieron y por los valores que me inculcaron siempre. También hago extensible al resto de mi familia: mi hermana Denise, mis abuelos, mis tíos, primos y mi cuñado.

También a Mariana, el amor de mi vida. Compañera infatigable. Gracias por darme la fuerza, la alegría y el amor de todos los días.

A Alicia Díaz, por haber confiado en mí, en que podía hacer investigación. Gracias por tanta paciencia y dedicación en enseñarme cotidianamente . Gracias por el cariño.

A Hilda, Gustavo, Fede, Fran y Yaya. A Alicia Zingoni, por mostrarme cómo honrar la vida.

Al LIFIA por darme el espacio y las oportunidades. A los directores, a mis compañeros. A los chicos del LTF que me dieron las primeras lecciones de esta cosa tan linda que es investigar. Gracias Fede Naso por dar tanto y pedir nada. Gracias a Diego, Nando, Richard y Casco. Gracias a mis compañeros de objetos y de multimedia. Son muchos, para todos más y más gracias.

Y por supuesto a mis amigos, por la fuerza que me dieron para empezar, para seguir y para terminar. Nachito por estar siempre empujando. Emi, con quien viví muchos años de felicidad y comparto una cotidiana amistad. Lea, con quien preparé la mayoría de las materias y crecimos juntos. A los chicos de los miércoles y a sus esposas y novias.

A Nacu, por ser un maestro.

A Hala Skaf y Pascal Molli por recibirme en Nancy y ser tan generosos en todo momento.

A la universidad pública y gratuita.

Durante los últimos años, los usuarios han tenido un rol mucho más activo en la Web creando documentos de producción propia o formando partes en comunidades virtuales. Las redes sociales, los blogs, y las wikis son ejemplos de aplicaciones Web donde la colaboración y participación de los usuarios son determinantes. Dentro de este contexto, aparecen las denominadas comunidades de creación de conocimiento. Las comunidades de creación de conocimiento se destacan, entre otras cosas, por la participación de los usuarios en la creación de nuevas fuentes de información, todo esto en forma colaborativa. Cada usuario realiza pequeños aportes, los cuales se complementan con los que realizan cientos de otros usuarios. De esta forma, la comunidad en si es la autora de grandes fuentes de información y gestión de conocimiento. Las wikis semánticas son un claro ejemplo de sitios de comunidades de creación de conocimiento colaborativo, donde a pesar de la gran participación individual de cada participante, los aportes y resultados son realizados en espacios compartidos donde la comunidad toda es la creadora y el rol del individuo es dejado de lado. En este trabajo de tesis se presenta una forma de incluir espacios personales semánticos en el proceso de creación de conocimiento de una wiki semántica utilizando anotaciones semánticas personales.

Índice general

1. Introducción	3
1.1. Motivación	3
1.2. Construcción Colaborativa del Conocimiento	7
1.3. Estructura del Documento	9
2. La Web y el Conocimiento	11
2.1. Ciclo de Construcción Colaborativa de Conocimiento	11
2.2. Web Semántica	16
2.2.1. ¿Qué es la Web Semántica?	16
2.2.2. Alcances y Desafíos	25
2.3. Wikis Semánticas	25
2.4. Tipos de Wikis Semánticas	29
2.4.1. Semantic Media Wiki	30
2.4.2. OntoWiki	30
2.4.3. SemperWiki	31
2.4.4. SWooki	31
2.5. Conclusiones del Capítulo	33
3. Construcción Colaborativa de Conocimiento	37
3.1. Soporte Informático a la Construcción de Conocimiento	37
3.2. Elementos privados y compartidos	40
3.3. Externalización	40
3.4. Internalización	42
3.5. Publicación	43
3.6. Reacción	43
3.7. Conclusiones del Capítulo	44

4. Información personal en una wiki semántica	47
4.1. Wikis Semánticas como Soporte Para la Construcción de Conocimiento	47
4.1.1. Espacio compartido	48
4.1.2. Navegabilidad	52
4.1.3. Espacio privado	53
4.2. Espacio Privado y Espacio Compartido	53
4.3. Información Personal en una Wiki Semántica	54
4.3.1. El espacio personal	55
4.3.2. Anotaciones Personales	55
4.3.3. Anotaciones Semánticas Personales: Categorías e Individuos .	55
4.3.4. Reuso de conceptualizaciones compartidas	57
4.3.5. Navegabilidad	57
4.4. Conclusiones del Capítulo	58
5. P-SWOOKI	59
5.1. SWOOKI: Una wiki semántica P2P	59
5.1.1. Arquitectura	61
5.1.2. Definición Semántica	62
5.1.3. Alcance	62
5.2. P-SWOOKI	64
5.2.1. Funcionalidad	65
5.3. Navegabilidad en el espacio personal	70
5.4. Implementación	70
5.4.1. Modelo de Datos	70
5.4.2. Definición de repositorios	71
5.4.3. Arquitectura P-Swooki	77
5.5. Conclusiones del Capítulo	78
6. Evaluación	79
7. Conclusiones y Trabajo a Futuro	85
7.1. Trabajo a Futuro	86
Anexo	87

7.2. Definición semántica RDF/OWL	87
---	----

Índice de figuras

2.1. Ciclo de Creación de Conocimiento por Stahl	13
2.2. La Web Semántica en Capas	18
2.3. Representación gráfica de una terna RDF	21
2.4. Mark ups de formato en una wiki	27
2.5. Visualización del texto con formato	28
2.6. Comparación de notaciones entre una Wiki tradicional y su versión semántica	28
2.7. Semantic Media Wiki	29
2.8. OntoWiki: interacción para Ingenieros en Conocimiento	31
2.9. SemperWiki: wiki semántica personal.	32
2.10. SWooki: la primer wiki semántica P2P	32
2.11. Arquitectura general interna de una wiki	33
2.12. Arquitecturas centralizadas y distribuidas en wikis semánticas	34
3.1. Ciclo de Stahl Simplificado [DC04]	41
3.2. Externalización y Publicación con Del.icio.us	43
4.1. Publicación de un nuevo concepto en SMW	48
4.2. Publicación de un nuevo concepto en SMW con conflicto	49
4.3. Agregado de información semántica en SMW	50
4.4. Agregado de información semántica en SMW	52
4.5. Agregado de Anotaciones Personales a una Wiki Semántica	56
5.1. Pantalla en modo edición: agregado de una anotación semántica	59
5.2. Pantalla en modo normal: visualización de una anotación semántica	60
5.3. Swooki: arquitectura	61
5.4. Estructura visual SWOOKI	66

5.5. Visualización en P-SWOOKI incluyendo la información semántica personal: el ejemplo muestra que para este usuario particular la página Paris es considerada un individuo (en inglés Individual) de las categorías Ciudad y Visitadas.	68
5.6. Agregado de información personal en P-SWOOKI	69
5.7. Navegación Personalizada	70
5.8. Anotaciones Semánticas Personales: Modelo de Datos	71
5.9. Arquitectura de Sesame	72
5.10. Arquitectura de P-SWOOKI	78

“Empieza por el principio ... y sigue hasta llegar al final, allí te paras”

Lewis Carrol - Alicia en el país de las maravillas

Capítulo 1

Introducción

1.1. Motivación

Durante mucho tiempo la mayoría de los sitios publicados en internet ofrecían a los usuarios información de consulta. La producción de material estaba destinado a los editores, mientras que los visitantes (usuarios) solamente estaban habilitados para realizar la lectura de ese material. Como ejemplo de ello podemos mencionar a los sitios “.com” a fines de los 90 y principio del 2000, lo cuales brindaban información que los clientes (usuarios) podían consultar. El flujo de información en este tipo de publicaciones se realizaba en una única dirección: desde el editor hacia los lectores. La empresa o editorial generaba los documentos y por otro lado los clientes o lectores accedían a la fuente de información para consumirla. A esta etapa de la Web se la conoce como Web 1.0.

Tiempo después, comenzaron a surgir cambios en la tecnología y en las formas de organización de las actividades. La información en los sitios comienza a ser dinámica, ciertas plataformas web posibilitaban a los usuarios generar material e información para uso personal y a configurar las vistas de acuerdo a perfiles pre definidos. Ejemplos de esto fueron las alternativas a programas de escritorio pero con acceso mediante la Web: aplicaciones de correo electrónico (hotmail, yahoo e-mail), calendarios, repositorios y organizadores de archivos remotos (BSCW).

En este proceso de migración de aplicaciones originariamente de escritorio a un entorno web aparecieron también las tecnologías de soporte al trabajo colaborativo (CSCW). Estas fueron precursoras en la generación de material y su almacenamiento en servidores remotos dentro del espacio de la Web. El trabajo colaborativo incrementó el número de sistemas caracterizados para la producción de material y pensados en entornos de trabajo remotos. Sistemas de conferencias, chats, pizarras

compartidas, calendarios para grupos de trabajo son algunas de los principales sistemas en los que CSCW comenzó a darle relevancia utilizando en gran parte como soporte la Web. En este proceso, los usuarios comenzaron a tomar la web como un recurso para, además de obtener información de consulta, generar nuevas fuentes de datos y almacenarlos [ter98].

La profundidad de estos cambios dio origen a denominada Web 2.0, la cual profundizó un cambio en el rol de los usuarios. La Web comenzó a nutrirse de producciones de información realizadas por usuarios, los cuales, a diferencia de la Web 1.0, no pertenecen a fuentes clásicas de editoriales de material como pueden ser sitios de una compañía o fuentes de información como diarios o newsletters: los usuarios utilizan la Web como medio de expresión. Estas necesidades generaron una demanda hacia la Web de manera que sea posible generar contenidos fácilmente, poder administrarlos y a su vez compartirlos.

Algunos autores caracterizan a la Web 2.0 como “un conjunto de tendencias económicas, sociales y tecnológicas que colectivamente dan forma a las bases para la próxima generación en Internet - más madura, distinguida por la participación de los usuarios, abierta e interconectada” [MOtORT07].

En términos estructurales, podemos indicar que la Web 2.0 esta basada en tres pilares [VH07]:

- Información Emergente - la forma de utilizar y combinar información de varias fuentes en una representación que permita la derivación de nueva información o valor agregado, como ejemplo de ellos podemos citar a los “mash-ups”.
- Funcionalidad - componiendo aplicaciones para generar otras, en algunos casos enriquecidas con la experiencia de los usuarios en la Web o para crear aplicaciones de internet ricas (RIAs *Rich Internet Applications*).
- Socialización del contenido - mediante el uso de sitios para la creación de contenidos con un importante sustento social como Wikis, blogs y sistemas de tagging. En estos sitios los usuarios no consideran su información solamente como privada, sino que deciden compartirlas con otros usuarios y de esta forma dar sustento a una plataforma de información subyacente.

La combinación de factores que permiten a los usuarios expresarse y compartir información entre pares se encuentra presente en los denominados sitios Web 2.0. Aunque la mayoría de estos sitios presentan una combinación de características antes mencionadas, podemos diferenciarlos en familias de sitios web de acuerdo a

las posibilidades que les brindan a sus usuarios. Estas familias están formadas por las redes sociales, los sistemas de tags y bookmarking, los blogs y las wikis.

La formación de redes sociales está, principalmente determinada por la posibilidad de que cada usuario pueda combinar la agenda de contactos con una rápida comunicación entre ellos. De esta forma, los recursos que posea cada usuario en el sitio Web serán compartidos con sus “amigos” de la red social. Como la comunicación es un factor importante dentro del esquema de redes sociales, la relación entre usuarios no está sujeta simplemente a compartir recursos sino también poder comunicarse mediante mensajería instantánea, envío de e-mails y notificaciones de eventos.

Un Blog es típicamente una secuencia de textos cortos en los que las entradas aparecen en orden inverso al que fueron publicados, es decir que las publicaciones más recientes siempre aparecen al comienzo de la página. En su forma más simple, un blog consiste solamente en texto aunque en las nuevas versiones más ricas en contenido es posible agregar imágenes.

Por su parte, la creciente popularidad que han ganado los sitios de tagging, como del.icio.us¹ y [Flickr](http://www.flickr.com)², han proporcionado nuevas fuentes para el entendimiento de los intereses de los usuarios en la Web. Los sistemas de taggeo colaborativo permiten a los usuarios elegir sus propios términos como tags para describir sus recursos favoritos en la Web, obteniendo una emergente esquema de clasificación conocido como *folksonomy* [mAYGS08].

Siguiendo con este tipo de aplicaciones, una Wiki es una página Web o un conjunto de páginas (wiki pages) las cuales son provistas por un sistema wiki que ofrece manejo por parte de los usuarios, un editor Web para la edición de páginas wikis y el control de versionado de las páginas. En contraste con los Blogs, las páginas wikis son editadas, comúnmente, por múltiples usuarios. Por lo tanto, las Wikis permiten a los colaboradores poder comunicarse por medio de la edición de las páginas wikis, esto es, ya sea leyendo los cambios o agregados que realizaron otros usuarios o realizando cambios ellos mismos [FS08]. En algunos casos la edición puede realizarse sin necesidad de registración previa; como es el caso de Wikipedia³ donde la wiki es pública. Así, una wiki se puede ver como una página o colección de páginas editables las cuales no requieren que sus usuarios conozcan cómo escribir un documento utilizando lenguajes del tipo de HTML. El término “wiki” es un derivado de la palabra hawahiana “wikiwiki” que significa “rápido”. Por lo tanto, el nombre

¹<http://del.icio.us>

²<http://www.flickr.com>

³<http://www.wikipedia.com>

sugiere un medio rápido para realizar publicaciones colaborativas en la Web. Una wiki es, en efecto, una base de información orientada al usuario la cual es simple y fácil de usar; lo cual la convierte en un excelente medio para la creación y el manejo de conocimiento.

Este abanico de sistemas que facilitan a los usuarios la creación de documentos en la web, la comunicación entre diferentes usuarios y el intercambio de contenidos se traduce en un fuerte incremento en la producción de información. Este incremento sumado al intercambio de conocimiento plantea un nuevo desafío: permitir que la información sea fácilmente accesible por los usuarios y sistemas web que pueden llegar a utilizarla.

La mayoría de la información que actualmente se encuentra en la Web está destinada para ser consumida por seres humanos. Los usos típicos de la Web, hoy en día, involucra a personas buscando o haciendo uso de información, buscar o estar en contacto con otras personas o consultando catálogos de productos. Pese a todo, estas actividades no están completamente bien soportadas por herramientas de software. Además de la existencia de links entre documentos, las herramientas más valiables y casi indispensables son los motores de búsqueda.

Los motores de búsqueda basados en palabras claves como Yahoo y Google son las principales herramientas utilizadas en la Web hoy en día. Aunque es destacable el aporte que realizaron estos buscadores al desarrollo de la Web, existen algunos problemas reflejados en el uso, las cuales son enunciadas en [Av08]: gran cantidad de resultados con baja precisión, poca cantidad de respuestas, resultados sensibles al vocabulario y que los resultados son páginas web y no otro tipo de documentos.

Aunque los motores de búsqueda han ido mejorando con el correr del tiempo, la gran producción constante de documentos en la Web los ha ido superando constantemente. De todas formas, aunque la búsqueda sea exitosa, es la persona quien debe realizar la extracción del contenido específico que está buscando. Esto se debe, principalmente, a que no se encuentra una gran cantidad de soporte para la "ubicación de la información". Además, los resultados de los motores de búsqueda no son accesibles fácilmente por otras herramientas de software, sino que son herramientas aisladas.

En general, el gran problema para obtener información de la Web no decae simplemente en las tecnologías de los buscadores, sino que el significado de los elementos de la Web no es accesible por una máquina. La idea general de Tim Bernes-Lee era permitir que los documentos entendibles por los seres humanos sean aumentados con *meta-data* procesable por una máquina. De esta forma, se podrían

ofrecer sistemas inteligentes sobre la Web que permitan realizar combinaciones de diferentes fuentes de información esparcidas en la Web con cierto significado, de manera de complementar el procesamiento netamente sintáctico y obtener una ventaja de la combinación de la nueva forma de representar la información. A este desafío se lo denomina Web Semántica [BLHL01].

Como fue mencionado anteriormente, las Wikis son uno de los sitios Web que permiten generar un gran número de documentos. Una wiki combina la facilidad de publicación con la constante comunicación entre los miembros de esa Wiki. De esta forma, la wiki en si mismo genera una comunidad de interés de la cual emerge conocimiento. Para que el gerenciamiento de la información, la descripción del conocimiento y así mismo la obtención de contenido sea mejorada surgen las Wikis Semánticas [VK06, SBBK08].

Una Wiki Semántica posee las mismas características básicas de una Wiki tradicional pero incorporando al lenguaje de markup meta información para darle significado, esta meta información es comprensible por una máquina. En una Wiki semántica las páginas son anotadas con información semántica lo cual permite una mejor obtención de la información, mejorar las búsquedas y emerger representaciones de conocimiento. Esta incorporación de elementos semánticos incorpora nuevos desafíos en el desarrollo de Wikis Semánticas: por un lado preservar la simplicidad de gestar información y por el otro permitir a los usuarios manejar anotaciones semánticas para darle un mayor gerenciamiento del conocimiento que emerge de la comunidad que se genera alrededor de la Wiki. La información semántica representa la base de conocimiento compartida, la cual describe el entendimiento común entre los participantes de la comunidad wiki. La creación colaborativa de esta base de conocimiento se realiza a través de un proceso iterativo y social.

1.2. Construcción Colaborativa del Conocimiento

En una wiki semántica los usuarios participan realizando aportes en el contenido de la misma. Cada usuario tiene la capacidad y la oportunidad de agregar nuevos documentos a la wiki, editar documentos existentes (los cuales pudieron ser escritos por otros usuarios), aumentar la información un documento existente, agregar elementos de navegabilidad entre documentos, agregar información semántica y, por supuesto, consultar la información presente en la wiki. La participación en estas actividades de continua mejora del contenido de la wiki es parte de un proceso social del que cada uno de los usuarios son parte. Cada uno de los documentos de

la wiki son el resultado de constantes acciones de creación, publicación y múltiples ediciones.

Este proceso genera la construcción de conocimiento en forma colaborativa, en el cual los diferentes usuarios cooperan participando en la gesta del conocimiento. Stahl, describe un proceso de construcción colaborativa de conocimiento [Sta06] como un proceso en espiral. En este proceso el conocimiento emerge primero como conocimiento personal el cual, luego, es socializado. Este ciclo involucra cuatro etapas fundamentales: externalización, internalización, publicación y reacción. La mayoría de las wikis semánticas actuales, no incluyen dentro de sus funcionalidades esquemas de trabajo que contemplen a las cuatro etapas. Algunas Wikis Semánticas se centran en el uso personal [OVBD06], pero no permiten compartir el contenido de la misma. Otras, como por ejemplo Semantic Media Wiki [KVV06], no incluyen elementos que permiten realizar anotaciones en forma personalizada, sino que la totalidad de la información reflejada en la wiki debe ser compartida y consensuada por el resto de la comunidad.

En este trabajo de grado se presentará una forma de incluir un espacio para realizar anotaciones semánticas personales en el contexto de trabajo colaborativo de una Wiki Semántica. El trabajo de producción y gerenciamiento de conocimiento que incluye una Wiki Semántica tradicional se verá enriquecido con la incorporación de un espacio para realizar anotaciones semánticas personales. Por lo tanto, el trabajo estará basado en dos aspectos importantes: por un lado, la adaptación de ciclo de construcción colaborativa de conocimiento (que define Stahn) a una Wiki Semántica para permitir un espacio de personalización de la información semántica y por otro lado definir un formato de escritura de anotaciones semánticas personales que no atente con una de las características esenciales en las wikis: simplicidad y facilidad en el uso.

El primer aspecto, tendrá como punto de partida una Wiki Semántica definida sobre una arquitectura P2P. Cada nodo (peer) de la red P2P representará a cada usuario en la comunidad de la Wiki. Esta red permitirá sincronizar información considerada pública y además alojara espacio para la creación de anotaciones personales. Para representar la arquitectura base P2P se utilizó SWOOKI [RSMM08], la primer Wiki Semántica P2P.

Por el lado de realizar las anotaciones personales semánticas y no perder simplicidad y facilidad de uso, el enfoque está definido en alinear el enfoque de *folksonomies*. *Folksonomy* puede ser visto como una nueva forma de nombrar y en algunos casos dar significado a recursos de acuerdo al nivel de participación de usuarios en el nombrado de esos recursos. Cada usuario define en forma totalmente

personal conjuntos de etiquetas para los recursos en la web, de esta forma, si muchos usuarios categorizan a un recurso con el mismo nombre, esta categorización determina el ("nuevo") significado para ese recurso [Wal07].

Para finalizar, en este trabajo de grado se brindará un estudio y definición de una Wiki Semántica que posibilite realizar anotaciones personales semánticas categorizando información de las páginas wikis dentro del contexto de la generación de conocimiento colaborativo. A partir de estas anotaciones personales, debe ser posible favorecer el conocimiento compartido entre los miembros de la comunidad alrededor de la Wiki Semántica. Así mismo, la información semántica personal, así como la compartida por la comunidad estará expresada en lenguajes procesables por máquinas utilizados para la representación de ontologías como OWL/RDF.

1.3. Estructura del Documento

A continuación, en el capítulo 2 analizaremos el desempeño de los usuarios en la Web y el enfoque de este fenómeno manejado por la Wikis Semánticas. Luego, en el capítulo 3, estudiaremos el ciclo de creación de conocimiento en aplicaciones colaborativas. En el capítulo 4 profundizaremos en la descripción de información personal y como puede incorporarse al ciclo de creación de conocimiento para una wiki semántica analizando la combinación de esta etapa con la descripción de anotaciones personales. En el capítulo 5 se presentará P-Swooki: una solución particular a la inclusión de anotaciones personales en una Wiki Semántica. Por último, en el capítulo 6 presentaré las conclusiones y trabajo a futuro.

Capítulo 2

La Web y el Conocimiento

Este capítulo se centra en presentar las alternativas que presenta la Web, actualmente, para gerenciar y generar conocimiento. En primer lugar, encontraremos una pequeña introducción al enfoque de Stahl sobre la construcción colaborativa de conocimiento en un entorno sustentado por computadoras. Luego, enumeraremos los elementos tecnológicos actuales que soportan los enfoques de la Web Semántica, incluyendo las limitaciones y ventajas, los lenguajes asociados a la descripción de conocimiento, desafíos y estado del arte. Luego nos centraremos en las Wikis Semánticas análisis más profundo en sobre las Wikis Semánticas, diferentes alternativas en la concepción de Wikis Semánticas y sus características principales.

2.1. Ciclo de Construcción Colaborativa de Conocimiento

El gerenciamiento de la información y el manejo del conocimiento que los usuarios depositan en la Web a través de las múltiples opciones de expresión requiere un análisis para que la asistencia de las computadoras pueda ser provechoso. Las personas interactúan con otras personas. Ya sea utilizando medios electrónicos como la Web o simplemente comunicándose en forma oral. La Web provee espacios donde se establecen intercambios de pensamiento y puntos de vista en los cuales se recrea un espacio de construcción de conocimiento. El estudio de la forma en que se establecen las relaciones de conocimiento y particularmente la construcción de conocimiento se remonta a teorías anteriores a la construcción de la Web, sin embargo los procesos de aprendizaje y de construcción de conocimiento se presentan, en los últimos tiempos, en un nuevo elemento comunicacional: la Web.

En esta sección se presenta un modelo de aprendizaje como un proceso

sociocultural el cual incorpora múltiples faces distinguibles que constituyen diversos ciclos de construcción de conocimiento social y personal. Un elemento para destacar es el proceso de relación de las ideas individuales dentro de un proceso sociocultural. Este modelo de construcción colaborativa de conocimiento, desarrollado por Stahl en [Sta06], incorpora ideas desde varias teorías pedagógicas y de comprensión con la esperanza de proveer un marco conceptual para el diseño de software de aprendizaje-colaborativo, especialmente para ambientes de construcción colaborativa de conocimiento basado principalmente en la vertiente pedagógica constructivista. El modelo presentado por Stahl es un intento para comprender el aprendizaje como un proceso social incorporando múltiples facetas distinguibles que constituyen el ciclo de construcción de conocimiento personal y social. La característica cíclica de este proceso permite ir incrementando la complejidad de las consultas basándose en una creciente capacidad de comprensión.

Diagrama de Creación de Conocimiento Personal y Social

La descripción de este ciclo de construcción de conocimiento puede ser explicada mediante un gráfico donde se distinguen principalmente etapas: la creación de conocimiento en el espacio personal y la creación de conocimiento en el espacio social. Para ello, Stahl, gráficamente representa los procesos de transformación con flechas mientras que los rectángulos representan a lo producido por esas transformaciones (figura 2.1). En particular este diagrama representa la de idea de un proceso secuencial en el cual las relaciones entre los elementos pueden tomar una gran variedad e infinita complejidad de formas. Aunque la identificación de elementos es arbitraria e incompleta, el diagrama puede servir como punto de partida para establecer una discusión sobre una teoría cognitiva para la construcción de conocimiento. En particular, conocimiento gestionado por computadoras.

El diagrama intenta modelar la constitución mutua de la construcción de conocimiento personal y social como un proceso de aprendizaje. El análisis comienza en el extremo inferior izquierdo, donde se muestra el ciclo de comprensión personal. El resto del diagrama muestra como nuestras creencias personales pueden ser articuladas en lenguaje y así ingresar dentro de un “misterioso” proceso social en el cual aparecen otras personas y nuestra cultura compartida. Esta cultura, a su vez, ingresa en nuestro fuero personal de entendimiento, nuestras estructuras de pensamiento, intereses y diversas influencias. Stahl señala que la cognición personal y la actividad social pueden ser separadas solamente en forma artificial, dentro de un modelo como éste, diseñado para el análisis. Según Hegel [Heg67], esta es

la naturaleza de una relación de sujetos mutuamente constituidos: ninguno puede existir sin el otro, pero es interesante distinguir puntos característicos para realizar el análisis.

Espacio Personal de Construcción de Conocimiento

Heidegger [Hei] y Schön [Sch83] argumentan que el aprendizaje se inicia en base a entendimiento (conocimiento) tácito previo. Algunas formas de desglose en el planeamiento o en nuestras actividades cotidianas se hacen en base a este entendimiento tácito. La red de significación por la cual le damos sentido a nuestro mundo es despedazada y debe ser reparada. La resolución del problema se produce a través de roer la naturaleza del problema en diferentes piezas de nuestro entendimiento. Debemos poder reparar nuestro entendimiento explicando las implicancias de nuestros entendimientos y resolviendo conflictos o completando vacíos (reinterpretando nuestra estructura de significación) hasta arribar a una nueva comprensión. Esto típicamente involucra cierta interacción con el mundo: desde nuestra experiencia con artefactos (elementos) como nuestras propias herramientas y representaciones simbólicas. Mas ambicioso, nuestro entendimiento debe someterse a un cambio fundamental de conceptos. Si tenemos éxito y el problema desaparece, esta nueva comprensión gradualmente se instala para convertirse en nuestro nuevo entendimiento tácito y para proveer el punto inicial para futuros entendimientos y aprendizajes.

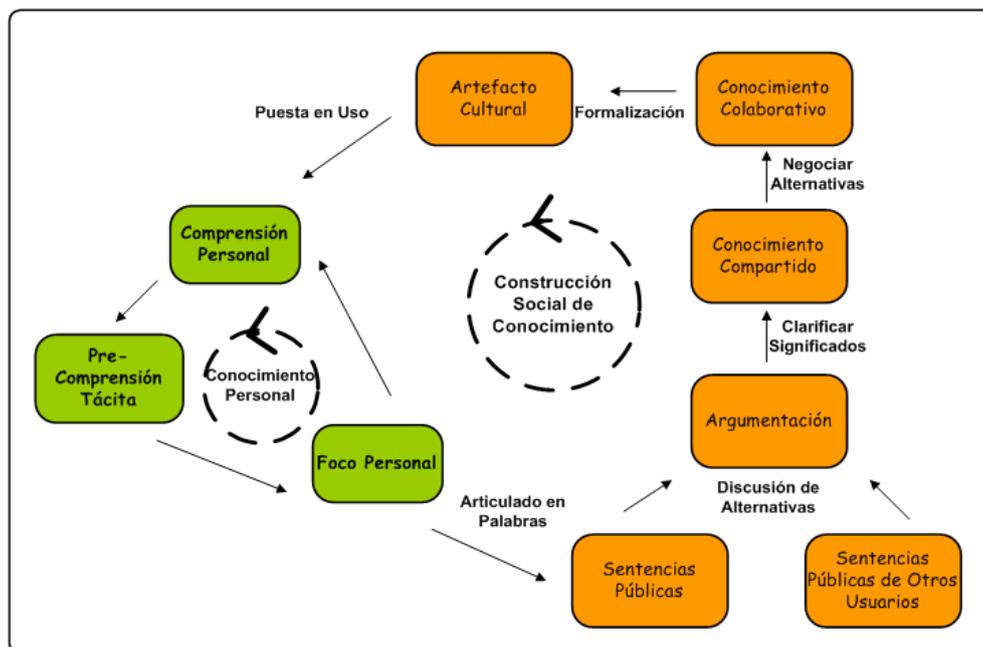


Figura 2.1: Ciclo de Creación de Conocimiento por Stahl

El proceso de interpretación, que parece ser llevado a cabo en un nivel de pensamiento individual, es también un proceso social esencial. La red de significancia personal en última instancia posee sus orígenes en el lenguaje personal y la cultura. La interpretación se ubica entre el lenguaje, la historia, la cultura, las estructuras sociales y políticas. Nuestra capacidad y estructuras de reflexión interna por sí mismas se originan en nuestras interacciones sociales previas. Nuestras interpretaciones personales o voces son una consolidación de muchas perspectivas y voces o géneros de otros que hemos conocido. Sin embargo, el contexto social y el origen social están escondidos porque ellos han sido incorporados dentro del espacio de pre entendimientos tácitos del individuo. No siempre posible resolver el carácter problemático de nuestro entendimiento personal en forma interna, particularmente cuando este es provocado por otras personas. Entonces debemos ingresar dentro de un proceso social explícito y crear nuevos significados en forma colaborativa. Para realizar esto, típicamente, articulamos nuestros credos iniciales en palabras y los expresamos nosotros mismos en declaraciones públicas.

Espacio Social de Construcción de Conocimiento

Al realizar las declaraciones públicas ingresamos en la parte más grande de la Figura 2.1. Aquí, podemos basarnos o suplementar los ciclos de aprendizaje individual de diferentes individuos (personas). Esto sucede cuando las creencias personales de alguna persona son articuladas en palabras y estas declaraciones públicas son tomadas en un contexto social y discutidas desde las múltiples perspectivas de los diferentes participantes. La discusión consiste en justificar argumentaciones provistas desde diferentes puntos de vista. El intercambio puede, gradualmente, converger en un entendimiento compartido obtenido desde una clarificación de diferencias en interpretaciones y terminologías. Si la comunicación es relativamente libre de intenciones ocultas, luchas de poder y prejuicios; los argumentos y clarificaciones puede resultar en acuerdos o por lo menos en entendimiento mutuo. Si la negociación de las diferentes perspectivas se traduce en aceptar el resultado común, entonces este resultado es aceptado como conocimiento. De esta forma, colaboración y comunicación sin distorsiones median entre creencia personal y conocimiento aceptado.

El hecho que el conocimiento sea un producto de la comunicación social no significa que este sea infundado o arbitrario. El medio del conocimiento (el lenguaje) esta fundado en las experiencias de vida de cada uno de los individuos, en nuestro sentido de racionalidad, en los patrones de interacción de las comunidades de

comunicación, en tradiciones culturales, y en el gran *background* de conocimiento que está implícitamente aceptado en cada acto de entendimiento o aceptación. Además, el proceso de comunicación que finaliza en conocimiento, incorpora argumentación que puede presentar evidencia empírica y deducción lógica a partir de otro conocimiento establecido. De todas formas, como el conocimiento nunca es absoluto siempre esta sujeto a la posibilidad de futuros cuestionamientos, reinterpretaciones y renegociaciones.

Los discursos personales que son resultado de discusiones, argumentación y clarificación de un lenguaje compartido dan consistencia al denominado proceso de comunicación. Para el proceso de comunicación se puede distinguir diferentes niveles: contenido proposicional, perspectiva, interacción social, reparación de malentendidos, connotaciones latentes, etc. Este proceso de lenguaje y análisis es negociado por el grupo público y transforma su conocimiento colaborativo. El entendimiento resultante existe solamente en la comunicación pública en la que toma lugar, aunque pueda ser incorporado subsecuentemente por cada uno de los participantes dentro de un proceso de aprendizaje individual.

Resumiendo, podemos considerar el proceso de creación de conocimiento que plantea Stahl como dos ciclos destacados de creación, el personal y el social. Podemos expresar una simplificación del proceso en cuatro etapas [DC04][NT95] :

Internalización Como parte del proceso individual de aprendizaje. La persona toma elementos del espacio social y los contrapone con sus propias creencias y redes de significancia. Luego de realizar diversos tipos de alineación y reestructuración de sus elementos, la nueva información (tal vez transformada) pasa a ser parte de sus contexto de conocimiento personal.

Externalización Donde se da lugar para que el conocimiento se transforma de tácito en explícito. Esta es una actividad personal privada que se lleva a cabo de manera aislada. Es en esta etapa donde los conocimientos tácitos son articulados en palabras que pueden ser expresadas en un ámbito público.

Publicación Implica hacer público el conocimiento, transfiriéndolo desde el plano de conocimiento individual al plano de conocimiento social.

Reacción Es el acto por el cual se responde a una nueva contribución dentro del ámbito social. La reacción en sí es una nueva contribución que torna más compleja la contribución inicial. El conocimiento generado es producto de diversos puntos de vista y lo hace más rico.

2.2. Web Semántica

A pesar que el ciclo de construcción de conocimiento está basado en interacciones entre personas dentro de un contexto social y cultural abierto, el estudio de este trabajo de grado está orientado a la interacción delimitada por las posibilidades que presenta la Web Semántica. El propósito de esta sección es analizar los objetivos que plantea la Web Semántica, cuáles son las tecnologías actuales y qué desafíos delimitan los próximos pasos.

2.2.1. ¿Qué es la Web Semántica?

La forma en que se utiliza la Web y la calidad de información reflejada en ella han sufrido alteraciones en los últimos tiempos. Las páginas Web modificaron la forma de interacción con los usuarios y esa modificación en la interacción planteó nuevas necesidades y dejó de lado elementos en desuso. La información y el uso de la Web es diferente en la actualidad que hace diez o veinte años atrás. En un principio los sitios Web ofrecían información que solamente podía ser consultada por los usuarios lectores. Con el correr del tiempo, los viejos usuarios lectores se transformaron en editores de grandes cantidades de contenido publicado en un espacio común: la Web.

El término “semántica” esta referido al estudio del significado o los cambios en el significado. Dentro del contexto de la Web Semántica, “semántica” está relacionado con la búsqueda o descubrimiento del significado de la información contenida en la web, no solamente por personas, sino también por máquinas. Según Passin [Pas04], la Web Semántica se enfoca en que computadoras (software) al igual que personas, puedan leer, buscar, interpretar y usar información diseminada en la extensión de la Web para obtener resultados útiles para los usuarios. De la misma forma, hacer que la información que esté en cualquier lugar de la Web sea accesible y comprensible tanto para computadoras como para personas.

La información diseminada en la Web está definida para ser manipulada por personas. Las personas buscan información y tratan de combinarla, establecen contactos con otras personas y consultan catálogos de productos. Estas actividades no están completamente bien soportadas por herramientas de software. Además de la existencia de links entre documentos, las herramientas más valiables y casi indispensables son los motores de búsqueda.

Los motores de búsqueda basados en palabras claves como Yahoo y Google son las principales herramientas utilizadas en la Web hoy en día. Aunque es destacable el aporte que realizaron estos buscadores al desarrollo de la Web, existen algunos

problemas reflejados en el uso, las cuales son enunciadas en [Av08]:

Gran cantidad de resultados, baja precisión Aunque el número de respuestas incluya aquellas páginas con un alto nivel de relevancia, el hecho de incluir otras 20.000.000 respuestas con un grado medio de relevancia puede generar un exceso de mala información. “Demasiado” puede ser tan malo como lo sería “poco”.

Poca cantidad o sin respuestas Originado cuando no obtenemos una gran cantidad de respuestas o las respuestas obtenidas no son de carácter relevante. Aunque este caso generalmente es poco probable en los buscadores actuales, en algunas ocasiones ocurre.

Los resultados son sumamente sensibles al vocabulario A menudo nuestras palabras claves de búsqueda no nos retornan los resultados esperados. En estos casos, los documentos originales utilizan una terminología diferente a nuestra consulta original. Esto es insatisfactorio, ya que en términos semánticos, consultas similares deberían retornar resultados similares.

Los resultados son Páginas Web Si necesitamos información que se encuentra dispersa en diferentes documentos de la Web, debemos realizar diferentes consultas para recolectar aquellos documentos pertinentes y luego manualmente obtener las piezas de información y componerlas juntas en un mismo documento.

Aunque los motores de búsqueda han ido mejorando con el correr del tiempo, la gran producción constante de documentos en la Web los ha ido superando constantemente. De todas formas, aunque la búsqueda sea exitosa, es la persona quien debe realizar la extracción del contenido específico que está buscando. Esto se debe, principalmente, a que no se encuentra una gran cantidad de soporte para la “ubicación de la información”. Además, los resultados de los motores de búsqueda no son accesibles fácilmente por otras herramientas de software, sino que son herramientas aisladas.

En general, el gran problema para obtener información de la Web no decae simplemente en las tecnologías de los buscadores, sino que el significado de los elementos de la Web no es accesible por una máquina. Según Tim Bernes-Lee en su artículo “The Semantic Web” [BLHL01], la Web Semántica no es una Web separada, pero sí una extensión de la actual, en la cual la información posee un significado bien definido de manera que computadoras y personas puedan

trabajar en cooperación. Con esta idea, aquellos documentos que solamente podían comprenderse por personas serán aumentados con *meta-data* que puede procesarse por una computadora. De esta forma, se podrían ofrecer sistemas inteligentes sobre la Web que permitan realizar combinaciones de diferentes fuentes de información esparcidas en la Web con cierto significado, de manera de complementar el procesamiento netamente sintáctico y obtener una ventaja de la combinación de la nueva forma de representar la información. A este desafío se lo denomina Web Semántica.

Las diferentes definiciones sobre la Web Semántica representan, en cierta forma, a un conjunto de ideas y tendencias más a que a una simple descripción tecnológica. A la meta-información agregada a los documentos accesibles vía un navegador web se le suman diferentes servicios que puedan procesarlos y obtener resultados también procesables por una máquina. Esta cadena de servicios e información comprensible tanto por máquinas como por personas, genera una red de manejo de conocimiento muy compleja. Varios servicios automatizados pueden ayudar a las personas a obtener y precisar información legible por una computadora. Estos procesos pueden recrear sistemas de manejo de conocimiento a través de varios servicios especializados de razonamiento sumamente útiles.

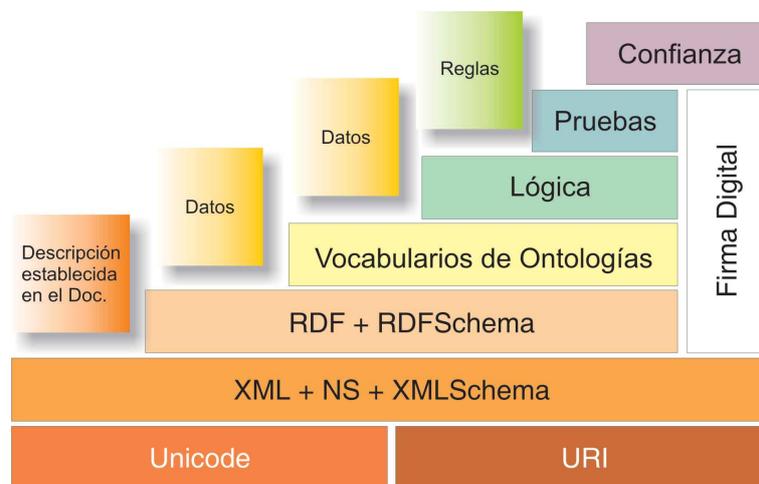


Figura 2.2: La Web Semántica en Capas

El desarrollo de la Web Semántica se define en etapas o capas. Cada etapa se encuentra definida en función de la o las capas que la anteceden. La justificación pragmática de este enfoque está basada en la facilidad de obtener consenso en base a pequeños pasos que tratan de encontrar el objetivo final analizando la totalidad en una sola toma. En la figura 2.2 esquematiza esta arquitectura en capas. Las bases están dadas por URIs y Unicode. Luego, podemos encontrar la capa de interoperabilidad sintáctica en la forma de XML y definiendo los lenguajes RDF

y RDF schema (RDFS). Sobre estas se define un vocabulario utilizando ontologías y luego, las últimas capas son las de lógica, prueba y confianza.

URI y Unicode

Una URI (Universal Resource Identifier) es una cadena de texto formateada que se utiliza para identificar recursos abstractos o físicos de forma única. Una URI puede ser utilizada como un localizador, un nombre o ambos. Existe subconjuntos de URIs: URL (Universal Resource Locator), las cuales identifican recursos por medio de su principal vía de acceso y URN (Universal resource name) las cuales son requeridas para permanecer únicas y persistir aun en los casos en los que el recurso no se encuentre disponible. Como ejemplos de estas la URL `http://lifa.info.unlp.edu.ar/staff/diegotorres.html` identifica la ubicación de la cual una página web puede ser obtenida, por su parte al URN `urn:isbn:3-543-34532-3` identifica a un libro utilizando su ISBN. UNICODE ofrece un único número por cada carácter, independientemente sobre que plataforma, programa o lenguaje se esté utilizando. De esta forma se evitan los conflictos entre diferentes codificaciones, permitiendo un abstraerse de la forma de representación de los datos y recursos.

XML

Es un lenguaje basado en etiquetas que permite estructurar documentos. Está aceptado como estándar para el intercambio de información estructurada en la Web pero no permite establecer el significado de la información en si. Es un lenguaje para información semi-estructurada utilizado para resolver problemas de integración de datos ya que permite una flexible codificación y representación de la información utilizando meta-información para describir la estructura de los datos. Una de las ventajas de XML, en comparación con HTML, es la posibilidad de definir nuevas etiquetas (tags). Por ejemplo, en XML es posible definir el tag `<nombre>`, el cual, de alguna forma, acarrea cierta semántica, sin embargo para una computadora el tag `<nombre>` no es diferente al tag HTML `<h1>`. Con este ejemplo queremos destacar que aunque el XML posea una gran ventaja en le descripción de esquemas de un documento, posee tanta descripción semántica como HTML. La información semántica que puede presentar XML no es comprensible por una máquina, solamente por personas. Sin embargo, XML resuelve varios problemas que anteriormente han sido muy difíciles de resolver mediante HTML: integración e intercambio de datos. Estas características han hecho que sea una de las bases fundacionales de la Web Semántica.

RDF

Sobre la capa de XML está definido RDF. RDF (Resource Description Framework) es un lenguaje estándar definido por World Wide Consortium (W3C)¹ para definir y estandarizar el uso de metadata. RDF es un modelo de datos básico, similar al modelo de entidad-relación, para escribir sentencias simples sobre recursos en la Web[Av08]. No es justamente un modelo XML, pero está basado en él, por eso se encuentra en la capa superior a XML. RDF se destaca por ser un modelo de datos formales para datos comprensibles por computadoras utilizado para proveer descripciones estándar de recursos Web. A su vez, RDF permite definir un modelo para definir y crear relaciones entre los recursos. Un recurso puede ser cualquier cosa, por ejemplo una persona, una canción o una página web.

Con RDF es posible utilizar primitivas previamente modeladas para expresar semántica de los datos de un documento, sin hacer ningún tipo de asunciones previos sobre la estructura del documento. En RDF es posible definir a un recurso como cualquier objeto que pueda ser identificable en forma unívoca utilizando una URI. Los recursos a su vez pueden tener propiedades asociados a ellos. Las propiedades están definidas en tipos y cada tipo posee sus correspondientes valores. Los tipos en cada propiedad expresan la relación de valores asociados con recursos.

La estructura básica de un documento RDF es utilizando ternas en la forma de sujeto, predicado y objeto. El **sujeto** es el recurso en si, es una cosa del universo sobre la que deseamos hablar y está identificada por una URL; el **predicado** estado por una propiedad que posee asociado un tipo, también está identificado por una URL; por último, el **objeto** es el valor que respeta el tipo de relación que indica el predicado. A modo de ejemplo supongamos que deseamos representar en RDF la siguiente sentencia:

John Lennon es el compositor de Imagina

Podríamos codificarlo de la siguiente manera

(<http://musician.com/lennon>, <http://www.mydomain.com/song-composer>,
#Imagina)

De esta manera estaríamos abstrayendo en la primer URL el sujeto, en este caso John Lennon, luego el predicado sería la acción de componer el cual es expresado por la identificación en la segunda URL (song-composer) y por ultimo el objeto el cual es codificado en este caso como un literal (#Imagina).

¹<http://www.w3.org/>

Gráficamente, podría traducirse en un grafo donde los nodos representan al Sujeto y al Objeto, mientras que las aristas representan a los predicados. Para el ejemplo dado se vería como muestra la figura 2.3

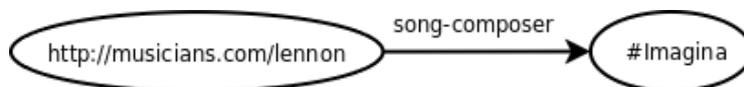


Figura 2.3: Representación gráfica de una terna RDF

Finalizando, la escritura completa de la terna del ejemplo utilizando RDF quedaría de la siguiente forma.

```

<?xml version="1.0" encoding="UTF-16"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
        xmlns:mydomain="http://www.mydomain.com/my-rdf-ns#"
  <rdf:Description rdf:about="http://musician.com/lennon">
    <mydomain:song-composer rdf:resource="#Imagina"/>
  </rdf:Description>
</rdf:RDF>

```

En la primer línea indicamos que utilizamos XML. El elemento `rdf:Description` crea una sentencia acerca del recurso `http://musician.com/lennon`. Dentro de esta descripción la propiedad esta definida como un tag y el contenido es el valor de la propiedad.

En conjunto con RDF, existe **RDF-Schema**. **RDF-Schema** provee un sistema de tipos para RDF. De esta forma, es posible crear un modelo de objetos para la cual la información del RDF es referenciada y mediante este modelo es posible describir el significado de los elementos.

Básicamente RDF-Schema provee un meta lenguaje para definir clases, propiedades y valores. El concepto de clases es similar al de programación orientada a objetos. Una clase es la definición de una estructura que compartida por un conjunto de elementos, a su vez es posible organizarlas mediante relaciones de herencia. Por su parte, las propiedades pueden verse como atributos que poseen los elementos de una clase. Las propiedades RDF puede heredarse y a su vez definir restricciones de dominio y rango. Por ultimo, los valores son los elementos básicos sobre los cuales se van construyendo estos modelos de objetos. Una de los elementos destacables al uso de RDF-Schema, es que permite una definición ontológica básica, de esta forma utilizando RDF y RDF schema, se pueden realizar consultas basadas en las relaciones fuertes que tienen los elementos representados en el dominio.

Finalizando, RDF provee los fundamentos para representar y procesar metadata. RDF se complementa con XML ya que el primero provee interoperabilidad semántica y el segundo sintáctica. Además, RDF-Schema, es un lenguaje ontológico primitivo, ofrece ciertas primitivas para definir elementos con significado. Existen lenguajes de consulta que trabajan sobre RDF y RDF-Schema como Sparql, el cual detallaremos luego.

Ontologías

La capa de representación de los datos con mayor nivel de abstracción esta dada por la descripción **Vocabulario de Ontologías**. Esta capa, basada en las inferiores, construye la base ontológica de la Web. Según Gruber [Gru93], "una ontología es una definición explícita de una conceptualización", la cual, según Borst[Bor97], debe poseer una especificación formal.

Para un dominio particular, una ontología representa un lenguaje más rico para proveer mayores restricciones en el tipo de los recursos y las propiedades. Comparada con una taxonomía, las ontologías mejoran la semántica de los términos proveyendo relaciones entre los términos del vocabulario. Las ontologías están expresadas normalmente en lenguajes lógicos, lo que permite realizar ciertas distinciones de significancia a través de las clases, propiedades y relaciones.

Para representar ontologías existen diferentes lenguajes, el que se ha destacado dentro de la Web Semántica es OWL. OWL (Web Ontology Lenguaje) es una extensión del vocabulario RDF y RDF-Schema (RDF/S), definido para contrastar las limitaciones de estos últimos. OWL mejora la expresividad de RDF proveyendo poderosos mecanismos para definir conceptualizaciones complejas y formalmente describir la semántica de clases y propiedades utilizadas en la descripción de recursos en la Web. Veamos algunas de las limitaciones de RDF-Schema[Car07, Av08].

En RDF/S no es posible definir equivalencias para propiedades. Esto es importante para establecer equivalencias de conceptos ontológicos, ya que posibilita alinear definiciones realizadas por grupos diferentes.

RDF/S no permite definir unicidad de propiedades. RDF/S permite definir el conjunto de valores que puede tener una propiedad, pero no permite indicar que ese conjunto debe ser cerrado. Como ejemplo, para la propiedad genero aplicada a una persona, podríamos indicar que permita ser valuada con "femenino" o "masculino" solamente.

RDF/S no permite definir disyunción de clases. Podríamos querer indicar

que dos clases son disjuntas. Por ejemplo, podemos decir que la clase Hombre y Mujer son disjuntas, pero en RDF/S solamente podemos indicar que Hombre es subclase de Persona y Mujer también.

RDF/S no permite expresar la unión o la intersección de clases. Esto permite la creación de nuevas clases que están compuestas por otras clases.

Características especiales de las propiedades. RDF/S no permite definir características especiales para las propiedades como decir que una propiedad es "transitiva", "única" o "posee inversa".

El conjunto de requerimientos para un lenguaje ontológico parecería ser, entonces, inalcanzable: soporte eficiente de razonamiento y una conveniente expresividad para un lenguaje que sea tan poderoso como combinar RDF Schema con una lógica completa. Estos requerimientos dieron lugar al Web Ontology Working Group de W3C para definir OWL como tres diferentes sublenguajes, cada uno dirigido a cumplir diferentes aspectos de esta serie de requisitos tan exigentes [OWL04]:.

OWL Full OWL Full está dirigido a usuarios que quieren máxima expresividad y libertad sintáctica de RDF sin garantías computacionales. Por ejemplo, en OWL Full una clase puede ser considerada simultáneamente como una colección de clases individuales y como una clase individual propiamente dicha. OWL Full permite una ontología para aumentar el significado del vocabulario preestablecido (RDF o OWL). Es poco probable que cualquier software de razonamiento sea capaz de obtener un razonamiento completo para cada característica de OWL Full.

OWL DL OWL DL está diseñado para aquellos usuarios que quieren la máxima expresividad conservando completitud computacional (se garantiza que todas las conclusiones sean computables), y resolubilidad (todos los cálculos se resolverán en un tiempo finito). OWL DL incluye todas las construcciones del lenguaje de OWL, pero sólo pueden ser usados bajo ciertas restricciones (por ejemplo, mientras una clase puede ser una subclase de otras muchas clases, una clase no puede ser una instancia de otra). OWL DL es denominado de esta forma debido a su correspondencia con la lógica de descripción (Description Logics, en inglés), un campo de investigación que estudia la lógica que compone la base formal de OWL.

OWL Lite OWL Lite está diseñado para aquellos usuarios que necesitan principalmente una clasificación jerárquica y restricciones simples. Por ejemplo, a la vez

que admite restricciones de cardinalidad, sólo permite establecer valores cardinales de 0 ó 1. Debería ser más sencillo proporcionar herramientas de soporte a OWL Lite que a sus parientes con mayor nivel de expresividad, y OWL Lite proporciona una ruta rápida de migración para tesauros y otras taxonomías. OWL Lite tiene también una menor complejidad formal que OWL DL.

Lógica, Pruebas y Confianza

El propósito de esta capa es ofrecer una funcionalidad similar a la que obtenemos con lógicas de primer orden. La idea es definir un estado de principios lógicos de forma tal que las computadoras estén en condiciones de razonar por inferencia utilizando esos principios. Por ejemplo, una Universidad desea saber qué estudiante posee un promedio superior a 8 para luego darle un premio. Un programa lógico realizaría la siguiente deducción: "Andrés tiene un promedio superior a 8, por lo tanto Andrés es meritorio de recibir un premio".

Los motores de inferencia, también llamados razonadores, son programas que derivan nuevos hechos o asociaciones a partir de información existente. Inferencia y reglas de inferencia permiten que se deriven nuevos datos a partir de datos ya existentes. Estas nuevas piezas de información pueden ser agregadas a la base de información existente, agrandándola en descripción.

Estas últimas tres capas en la definición de la Web Semántica conciben el hecho de utilizar razonadores que obteniendo información descrita en ontologías (utilizando OWL) y las capas subsiguientes, se puedan tomar ciertas decisiones y dar respuestas a deducciones lógicas. De esta forma, las aplicaciones web semánticas puedan dar pruebas de sus conclusiones.

Por ejemplo en OWL es posible realizar inferencias basadas en las asociaciones representadas en el modelo, lo que en principio significarían relaciones transitivas. Existen varias aplicaciones que actúan como razonadores, como por ejemplo Jena², Fact++³ y Hoolet⁴.

Confianza es el tope de la arquitectura de la Web Semántica. Esta capa provee servicios para autenticar la identidad y confiabilidad sobre los datos y los servicios. La idea es permitir que las personas pregunten sobre la confiabilidad de la información que pueden obtener de la Web. Actualmente, esta capa aun no ha sido muy explorada. Algunos escenarios que puedan llegar a suscitar en

²<http://jena.sourceforge.net>

³<http://owl.man.ac.uk/factplusplus/>

⁴<http://owl.man.ac.uk/hoolet>

esta capa podrían ser que una persona diga: "yo creo en la información que proviene de <http://www.informacionverdadera.com> pero no creo en la información de <http://www.informacionfalsa.com>".

2.2.2. Alcances y Desafíos

Aunque es bastante el trabajo realizado para solventar el desafío que propone la Web Semántica, no es descabellado indicar que la propuesta está dando sus primeros pasos. Esto no quita la basta cantidad de servicios que actualmente se están proponiendo en la web. Las áreas mas destacadas en las aplicaciones web semánticas están relacionadas con manejo de conocimiento (mayoritariamente en intranets), Integración de Datos (Boeing, Verizon y otras empresas trabajan en ello), *e-science* en particular ciencias biológicas, convergencia en grids semánticos, personalización, búsquedas semánticas en gran escala (aplicadas al WEB), movilidad y *context-aware*, inferencia y alineación RDF, evolución de ontologías, entre otras.

2.3. Wikis Semánticas

Una Wiki Semántica puede definirse como una Wiki tradicional que incorpora la posibilidad de agregar información semántica a fin de darle significado al contenido de manera que este sea comprensible por una máquina. En una Wiki semántica las páginas son anotadas con información semántica lo cual permite una mejor obtención de la información, mejorar las búsquedas y emerger representaciones de conocimiento. Esta incorporación de elementos semánticos incorpora nuevos desafíos en el desarrollo de Wikis Semánticas: por un lado preservar la simplicidad de gestar información y por el otro permitir a los usuarios manejar anotaciones semánticas para darle un mayor gerenciamiento del conocimiento que emerge de la comunidad que se genera alrededor de la Wiki. La información semántica representa la base de conocimiento compartida , la cual describe el entendimiento común entre los participantes de la comunidad wiki. La creación de esta base de conocimiento colaborativo se realiza a través de un proceso iterativo y social.

Normalmente, las wikis (sean semánticas o no) presentan dos vistas básicas. Una vista para realizar consultas, leer contenido y navegar por la wiki y otra vista en la que se realizan los cambios o se agrega nueva información a una página particular. Básicamente, el usuario que utiliza la wiki navega por ella y si lo cree necesario puede elegir cambiar a un modo edición para modificar el contenido. El modo edición, por lo general, presenta una vista esquemática del contenido de la página wiki que se

está editando y permite agregar o quitar información modificando texto plano, el cual contiene ciertas etiquetas para darle formato.

Al conjunto de etiquetas utilizadas para dar formato al texto contenido en una página wiki se lo denomina lenguaje de mark up. El lenguaje de mark up más usado actualmente en wikis es el que utiliza media wiki y que se hizo popular por el sitio Wikipedia. Aunque el lenguaje de mark up de media wiki no es el único, el resto de los lenguajes no difieren mucho de este ya sea en prestaciones como en la forma de utilizarlo.

Cuando un usuario accede a la vista de edición de una página wiki, visualiza en texto plano el contenido de la misma. Es decir, visualiza el texto completo complementando por aquellos elementos de markup que le dan formato a la información de la página. La sintáxis básica para formatear texto puede resumirse en:

Negrita Se encierra el texto que se desea visualizar con estilo negrita entre tres comillas simples, por ejemplo "**Texto en Negrita**".

Cursiva Utilizando solamente dos comillas simples para encerrar el texto, por ejemplo: "Texto en Cursiva".

Links Internos En este caso se utilizan dobles corchetes y entre ellos se especifica el título de una página de la wiki, por ejemplo `[[La Plata]]` para indicar un link a la página de la wiki intitulada "La Plata".

Links Externos De la misma forma que los links internos, pero en este caso entre los corchetes se coloca una url y luego el título de la página. Para indicar, por ejemplo, un link externo a la página del buscador Google se debe escribir `[[http://www.google.com Google]]`.

Títulos Los títulos se especifican con el doble signo de igualdad, por ejemplo para agregar el título "Título" se debe escribir `== Título ==`.

Las figuras 2.4 y 2.5 muestran el modo edición y el modo visualización respectivamente de una página wiki donde se utilizaron los mark ups de formato básicos.

En el caso de las wikis semánticas, el conjunto de etiquetas básicas de formato que provee la wiki, son enriquecidas con etiquetas de carácter semántico. En la figura 2.6 muestra una comparación entre la edición de una página en una wiki y en la versión semántica de la misma wiki. En la parte izquierda de la figura, se describe

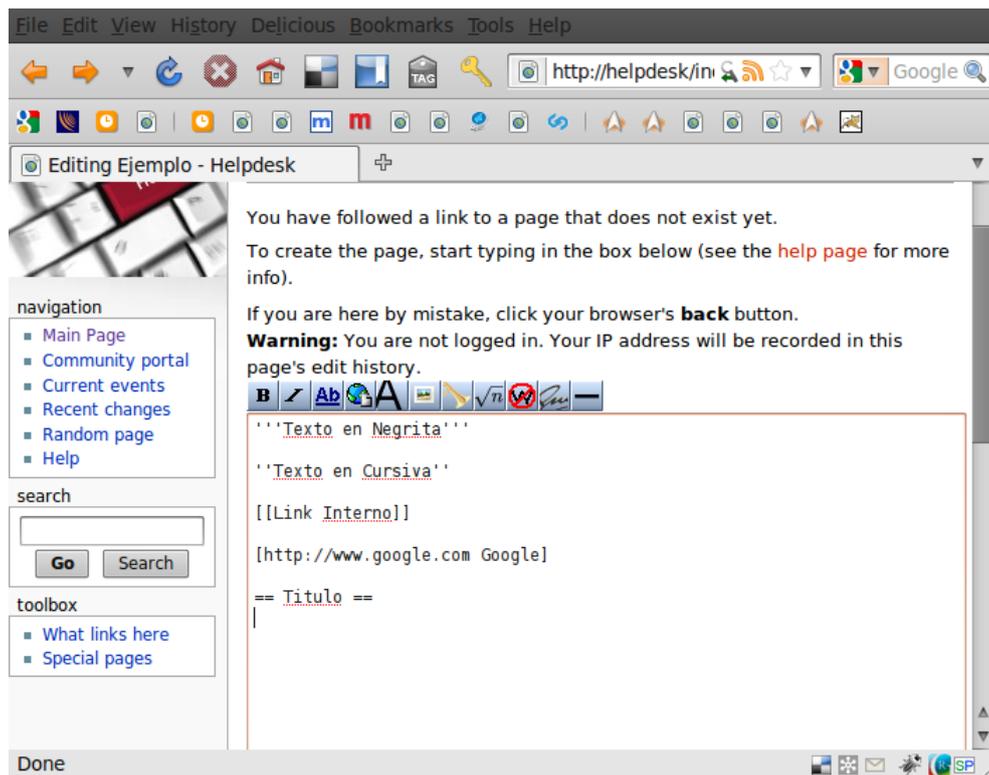


Figura 2.4: Mark ups de formato en una wiki

un link interno a la página de la wiki titulada “Buenos Aires”, este link es solamente navegacional, y no se establece ninguna información extra entre estas dos páginas. Por su parte, en la parte derecha de la figura, se ve que el link a la página “Buenos Aires” posee la etiqueta “has_capital”. De esta forma la relación de navegabilidad entre las dos páginas se ve aumentada por un tipo en esa relación. Ahora, se puede indicar que la página “Argentina” tiene como capital a “Buenos Aires”.

Por si mismas, las Wikis Semánticas son una excelente herramienta para crear modelos de conocimiento en forma colaborativa. Basadas en descripciones en lenguaje natural, ya sea emergente o existente, de modelos de conceptos o individuos, modelos de conocimiento formales pueden ser creados sucesivamente.

A medida que se incorpora información en las wikis semánticas, la misma va engrosando una base de conocimiento definiendo una ontología propia de la wiki. La forma de crear la ontología puede ser diversa, en algunos casos son creadas por expertos en el dominio, en otros casos emerge a través de la participación de los usuarios y en otros son mixtas ya que combinan las alternativas anteriores. Una alternativa que combina a las anteriores puede ocurrir cuando las descripciones en lenguaje natural son ser creadas en primera instancia por expertos en el dominio, por ejemplo biólogos, y luego formalizados en colaboración con ingenieros especializados en conocimiento (*knowledge engineers*).

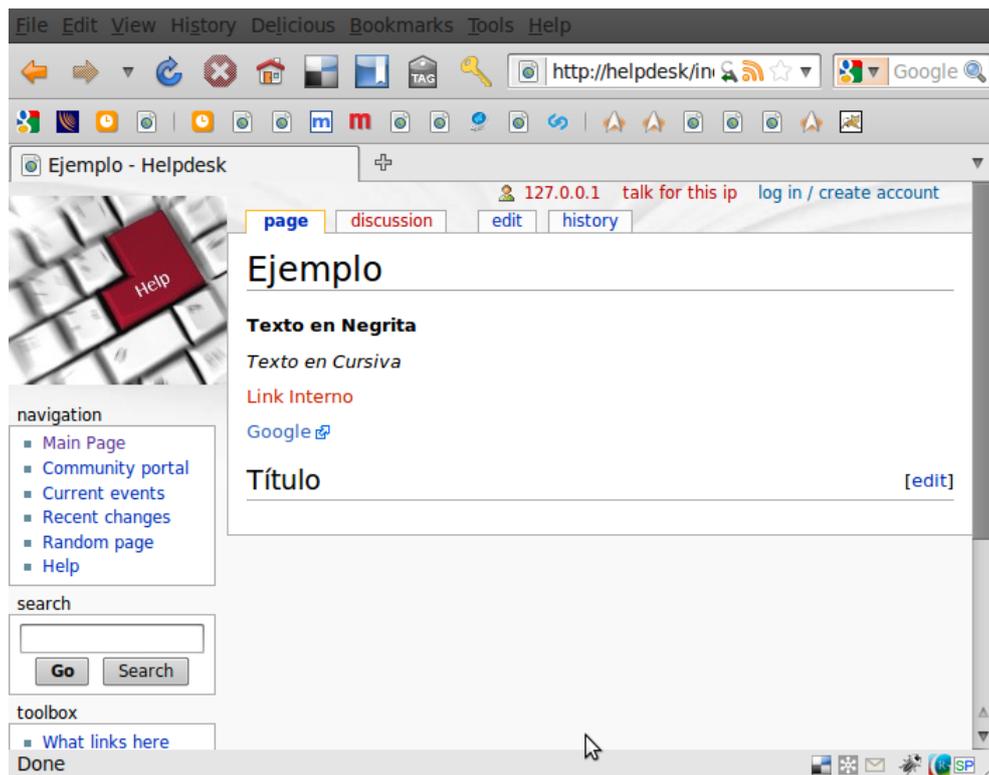


Figura 2.5: Visualización del texto con formato



Figura 2.6: Comparación de notaciones entre una Wiki tradicional y su versión semántica

Aunque la información semántica está acotada solamente a aquellos fragmentos de información que fueron etiquetados utilizando mark ups semánticos, la ontología resultante es sumamente importante. Definidos en términos de la Web Semántica, la definición ontológica de la wiki es procesable por máquinas lo cual permite realizar inferencias automatizadas del conocimiento volcado en la wiki. Estas inferencias pueden ser utilizadas directamente para nutrir de si misma a la wiki, por ejemplo incorporando páginas que retornen consultas sobre la información contenida o haciendo externo el conocimiento que expresó la comunidad reunida en torno a la wiki semántica accediendo a la información con una API de consultas. Muchas de las Wikis Semánticas poseen la facilidad de exportar el modelo de conocimiento a formatos como RDF u OWL.

2.4. Tipos de Wikis Semánticas

La diversidad de escenarios y posibles usuarios generó una gran variedad de alternativas en las Wikis Semánticas. Algunas Wikis Semánticas se enfocan en el contenido de las páginas y la posibilidad de agregar anotaciones semánticas como un valor agregado, se dice de este tipo que son "software social con soporte semántico". En este tipo de wikis, la definición ontológica emerge a partir del agregado de anotaciones semánticas en el contenido de la wiki. El hecho de que sea una ontología emergente se debe a que no se especifica una ontología ad hoc al momento inicial de interactuar con la wiki, sino que son los propios usuarios y sus acciones quienes van determinando los elementos que van formando parte de la ontología y cómo son sus relaciones.

Por otro lado existen wikis semánticas donde la anotaciones semánticas se encuentran en primer plano siendo aún más importante que el contenido en si, estas son catalogadas como del tipo "Web Semántica con soporte social". Además, existen algunas alternativas de wikis semánticas personales, como es el caso de IkeWiki.

A continuación se presentará un conjunto de wikis semánticas que permiten ejemplificar las cualidades generales de este tipo de aplicaciones sociales. De cada una de ellas se podrán observar si están enfocadas al contenido o a las anotaciones semánticas que poseen y sus cualidades particulares más destacadas.

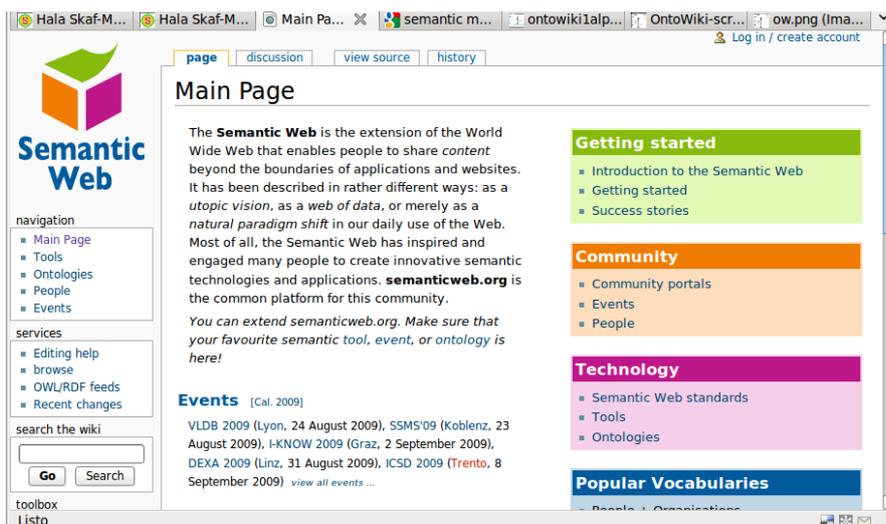


Figura 2.7: Semantic Media Wiki

2.4.1. Semantic Media Wiki

Semantic Media wiki (también conocida como Semantic Wikipedia) es uno de los casos más famosos de wikis semánticas (figura 2.7. Semantic Media Wiki (SMW) es la versión semántica de Media Wiki, el motor que utiliza Wikipedia. Según sus creadores[KVV06], SMW introduce una extensión en el sistema de markups de Mediawiki que permite anotaciones semánticas. En términos generales, el objetivo de SMW es reusar conocimiento, teniendo páginas consistentes y permitiendo a los usuarios buscar y compartir contenido de la wiki. SMW define a cada página como un concepto y extiende el conjunto de palabras claves de MediaWiki con Categorías, Relaciones y Atributos. Estas tres nuevas formas de anotar amplían el lenguaje de markup y son utilizadas en la vista de edición de una página wiki. Por ejemplo, para anotar a una página wiki como miembro de una categoría "Conferencia" se debe escribir en la vista de Edición el texto `[[Category: Conferencia]]`. SMW incluye una funcionalidad para exportar a RDF/OWL la cual esta basada en diversas reglas de mapeo: los artículos son individuos OWL, las categorías se transforman en clases y las relaciones se mapean a propiedades entre artículos (páginas wikis).

2.4.2. OntoWiki

OntoWiki está pensada para sustentar una ingeniería de conocimiento ágil[ADR06] y propone la integración de ternas RDF dentro del texto de la wiki utilizando una sintaxis especial. OntoWiki pertenece a la familia "Web Semántica socialmente habilitada". Visualmente, OntoWiki posee un ambiente mediante el cual se puede navegar a través de diferentes definiciones ontológicas. Por lo tanto, es posible navegar a través de conceptos, clases e instancias. Las facilidades de una wiki, en este caso están orientadas a facilitar la edición ontológica. En el modo de edición es posible agregar propiedades, valores y también clases a la base de conocimiento preestablecida.

De acuerdo a sus creadores, el foco principal de OntoWiki es proveer interacción social a los desarrolladores de bases de conocimiento 2.8. Esto es sustentado por sistemas de seguimiento de cambios, comentarios, ranqueo de información teniendo en cuenta calidad, presentación y originalidad, y diversas estadísticas sobre las publicaciones. Es importante destacar que los usuarios pueden discutir y calificar las publicaciones que realizan otros usuarios como así tambien sus propias publicaciones. Estas funcionalidades permiten la evolución progresiva de la definición de la ontología ya que los comentarios pueden tomarse como correcciones y los usuarios con un alto rating en sus publicaciones pueden ser tomados como referentes dentro

de la comunidad.

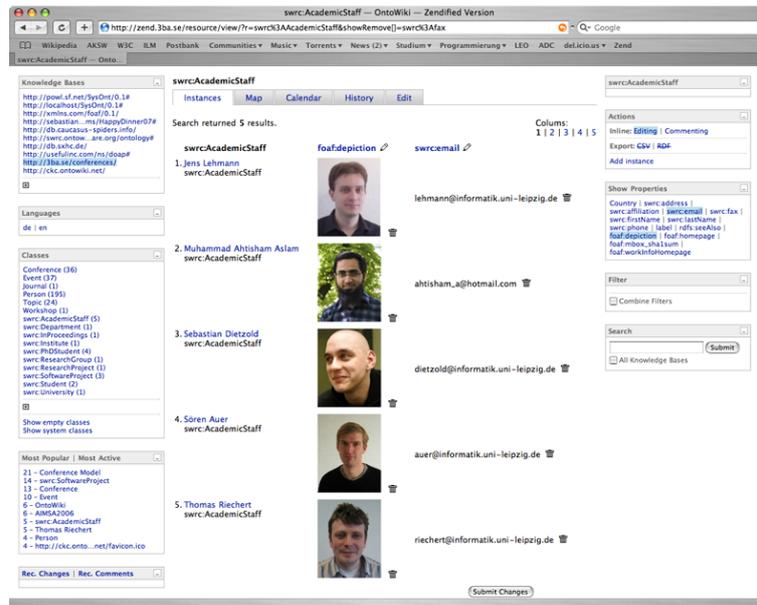


Figura 2.8: OntoWiki: interacción para Ingenieros en Conocimiento

2.4.3. SemperWiki

SemperWiki (Semantic Personal Wiki) es la implementación de una wiki semántica personal. Esta wiki está pensada como ayuda para categorizar y anotar elementos a nivel de escritorio de usuario. En esta wiki no hay separación de vistas para editar y navegar (figura 2.9). Una página puede consistir en texto normal, links a otras páginas o sitios web y tener anotaciones semánticas [Ore05]. En SemperWiki la información se guarda en formato RDF. El conjunto de ternas que SemperWiki guarda en un modelo RDF puede ser compartido directamente con otros usuarios. SemperWiki fue diseñada como un servidor de información personal que se complementa con las facilidades de realizar anotaciones semánticas.

2.4.4. Swooki

Swooki[RSMM08] es la primer wiki semántica peer to peer (P2P). Swooki está basada en Wooki[WUM07], un sistema wiki p2p. Swooki integra las tecnologías de la web semántica con la misma filosofía que Semantic Media Wiki. Las anotaciones semánticas están embebidas en el texto con un lenguaje de markups especial, como por ejemplo lo son los links tipados, que luego son visualizados como hechos (en la figura 2.10 se indican como facts). En Swooki una ontología formal emerge durante la edición de las páginas wikis. La característica P2P, permite formar una red de

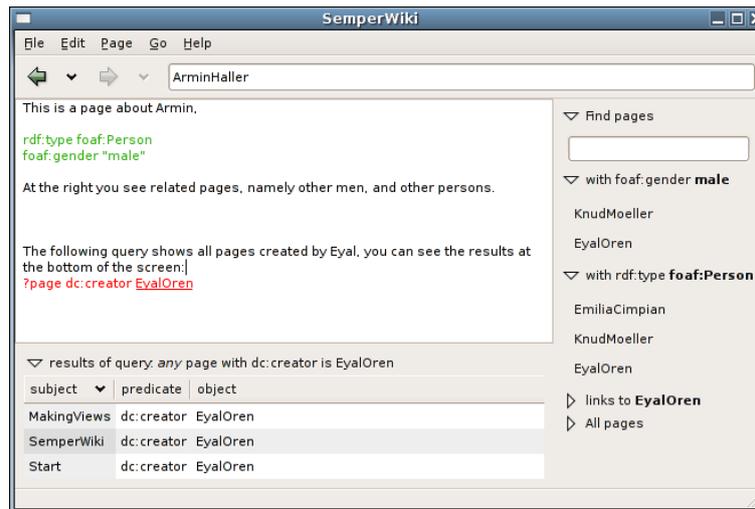


Figura 2.9: SemperWiki: wiki semántica personal.

servidores wikis ad-hoc, seleccionando con que pares desea sincronizar y favorecer la colaboración masiva. En este tipo de arquitecturas dos comunidades pueden generar páginas y relaciones en forma aislada y luego, sincronizarse para dar forma a una nueva wiki semántica permitiendo, entre otras cosas, balancear la carga en las consultas efectuadas.



Figura 2.10: SWiki: la primer wiki semántica P2P

Arquitecturas de Wikis Semánticas

Una Wiki Semántica esta conformada, al menos por los siguientes componentes: una interfaz de usuario, un parseador, un analizador de información y un repositorio de información. Los usuarios pueden navegar, editar y consultar páginas utilizando la interfaz de usuario. Cuando un usuario edita una pagina, la interfaz de usuario

notifica al parseador. El parseador, por su parte, analiza el texto y extrae las anotaciones y los links. Toda la información, como texto, anotaciones, etc, se almacena en un repositorio semántico. Desde los datos en el repositorio, el analizador computa los conjuntos de páginas que se relacionan con la página actual que está siendo visualizada por el usuario. Las consultas se realizan sobre el repositorio y los resultados se visualizan en la interfaz de usuario [OBD06]. La figura 2.11 muestra esta organización.

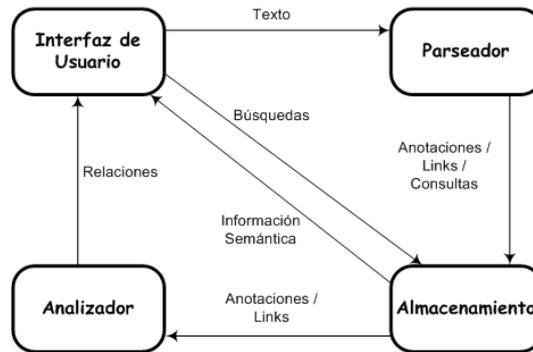


Figura 2.11: Arquitectura general interna de una wiki

Además de la organización interna, las wikis semánticas pueden definirse en centralizadas o distribuidas. En las wikis semánticas centralizadas toda la lógica que se detalló anteriormente se encuentra en un único servidor al que acceden todos los usuarios. Por su parte, existen wikis distribuidas donde la carga de las páginas está balanceada entre diferentes servidores conectados entre sí 2.12. En este último caso, los servidores deben garantizar el acceso a la información sin importar a que servidor se conecte un usuario. Más adelante analizaremos en detalle el caso de una wiki P2P.

2.5. Conclusiones del Capítulo

En este capítulo analizamos la Web como contexto para la creación de conocimiento colaborativo propuesto por Stahl. En esta visión constructivista, cada persona participa dentro de este ciclo en dos contextos marcados especialmente para el análisis. En primera instancia se considera el fuero personal donde se establecen reglas para la incorporación de nueva información, adecuación de esa información y reestructuración de las redes de significancia personal. Luego la capacidad de hacer la transformación del conocimiento tácito en una forma explícita haciendo uso del lenguaje para luego poder exponerlo en un entorno social. En la etapa social entran en juego diferentes expresiones lingüísticas provenientes de varios

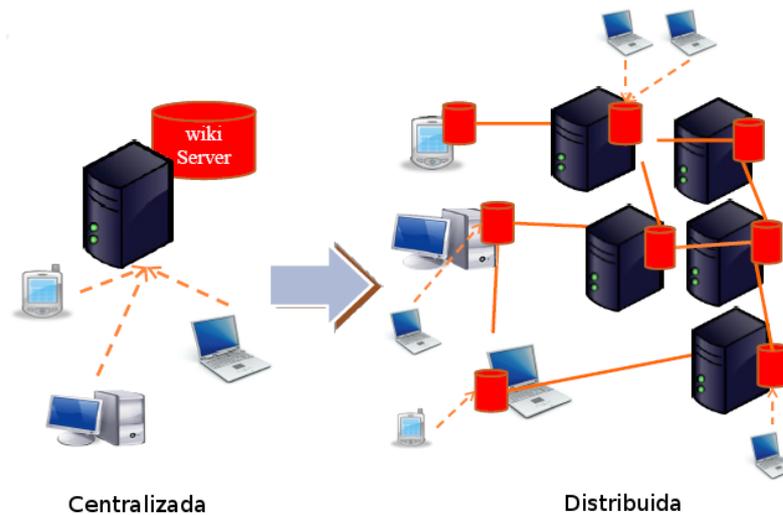


Figura 2.12: Arquitecturas centralizadas y distribuidas en wikis semánticas

individuos las cuales son discutidas dentro del contexto de la comunidad en la que fueron publicadas. La comunidad discute los nuevos postulados e intenta adecuarlos, cuestionarlos o refutarlos.

Las Wikis Semánticas, por sus características, son contextos en los cuales existe una comunidad que coopera para construir conocimiento. En ellas se combina la facilidad para generar documentos, una idea de comunidad implícita en su participación y la ayuda de las tecnologías que acompañan a la Web Semántica. Podemos ver a las wikis semánticas como un nexo entre la construcción colaborativa de conocimiento y las aplicaciones de razonamiento y pruebas que se ubican en las capas superiores de la arquitectura en capas que describen a la Web Semántica. Las Wikis Semánticas aumentan la descripción de la información de las páginas con meta información procesable por máquinas y permiten, en general, representación de los elementos presentes en lenguajes como RDF/OWL. Como corolario, presentamos posibles arquitecturas para sistemas de wikis semánticas.

Para finalizar, en este capítulo se describió en forma general tres elementos que acompañan a la construcción de wikis semánticas: el proceso de creación de conocimiento, las alternativas tecnológicas propuestas en la Web Semántica y las alternativas más relevantes en wikis semánticas.

Pretendemos analizar la propuesta de Stahl dentro del contexto de la Web Semántica, sin embargo para ello necesitamos analizar la forma general que tiene esta propuesta de concepción de la Web. El escenario propuesto por este trabajo de grado, pone el foco en la construcción colaborativa de conocimiento sobre wikis

semánticas. La arquitectura de la web semántica posiciona a las Wikis en las capas de aplicaciones, donde combinan elementos de las capas de representación de datos, utilizando lenguajes como RDF u OWL con las capas inmediatas superiores: Lógica y Pruebas.

Existen diferentes enfoques para la utilización de wikis semánticas. Algunas como Semantic Media Wiki se encuentra dentro de la familia de aplicaciones sociales que se adecuan a la web semántica, mientras que otras como OntoWiki se establecen como aplicaciones semánticas a las que se les agrega la opción de trabajar en un marco social. También podemos destacar aquellas pensadas para utilización personal, como SemperWiki.

El contexto general merece una discusión en la forma en que el usuario participa dentro del ciclo de creación de conocimiento utilizando wikis semánticas. Cada una de las alternativas propuestas, desde el esquema general de la definición de Web Semántica hasta la presentación de las diferentes alternativas wikis semánticas proponen analizar la posibilidad de adaptar el desarrollo personal dentro del ciclo de conocimiento utilizando wikis semánticas.

Capítulo 3

Construcción Colaborativa de Conocimiento

Este capítulo explora el modelo de construcción colaborativa de conocimiento desde el enfoque informático. Dentro de este modelo analizaremos los diferentes estadios. El rol del usuario como individuo y como miembro de un grupo en el que se crea y gestiona conocimiento. Analizaremos diferentes estadios en la gestión de conocimiento: Publicación, Externalización, Internalización y Reacción.

3.1. Soporte Informático a la Construcción de Conocimiento

El modelo presentado por Sthal simplifica algunas transiciones entre las diferentes etapas o la complejidad de las acciones que se encuentran involucradas con el fin de poder realizar un enfoque conceptual. Algunas de estas etapas pueden sustentarse o representarse utilizando herramientas informáticas. Dependiendo de la complejidad de la etapa, la transformación en elementos computacionales pueden tener diferentes niveles de complejidad.

Definiendo una secuencia de etapas típicas de la construcción colaborativa de conocimiento, el diagrama de la figura 2.1 sugiere un conjunto de puntos de enfoque donde el soporte dado por computadoras es deseable. Esto permite definir un marco conceptual para el diseño, uso y evaluación de un ambiente de creación de conocimiento.

Un ambiente de construcción de conocimiento debería ir más allá que un sistema de un único propósito, como podría ser un simple foro de discusión, y permitir

más de una fase del proceso de construcción de conocimiento social. Esto incluye mantener un registro del conocimiento que se va creando, a diferencia de un chat común, grupo de noticias y sistemas listas que eliminan las publicaciones luego de cierto periodo de tiempo. Por lo tanto deben ser creados en forma sincrónica, con tecnologías colaborativas persistentes y que puede ser accesible desde la Web.

Un ambiente de construcción de conocimiento debe soportar por lo menos varias etapas de la construcción de conocimiento. Debe asistir a las personas a expresar sus credos, discutirlos con otros, diferenciar sus propias perspectivas, criticar y explicar reclamos, negociar acuerdos y entendimientos compartidos y formular conocimiento en una representación final.

Por ello, se intentan corresponder un conjunto de herramientas informáticas para cada fase del ciclo social de construcción de conocimiento, sin embargo no es sencillo establecer un soporte computacional para la etapa cognitiva personal. Las creencias personales son articuladas como sentencias públicas antes que ellas puedan interactuar con el soporte informático.

Como punto de partida, los ambientes informáticos destinados a la educación (CSCL: Computer Support for Cooperative Learning) pueden facilitar el proceso de articular ideas y preservarlas en una forma correcta. Un editor de textos o simplemente un procesador de texto es una instancia sencilla de esto. Algunos ambientes de creación de conocimiento han intentado introducir técnicas para facilitar el inicio de la construcción del conocimiento. Por ejemplo, para generar que una persona realice el aporte inicial, un editor podría comenzar con la siguiente frase: “Yo creo que porque ”. Otro enfoque podría ser el hecho de incluir un editor para un resumen o un área para brainstorming.

Las declaraciones públicas de una persona frente a las de otras personas El soporte informático puede representar, también, las diferentes perspectivas desde las cuales las declaraciones públicas de una persona frente a las de otras emergen. Las perspectivas son más generales que las sentencias individuales en si, porque una persona puede ofrecer sentencias para múltiples perspectivas y diferentes personas pueden estar de acuerdo en una perspectiva común. Las perspectivas pueden relacionarse con otra, por ejemplo, derivando desde una perspectiva común que han compartido. La representación computacional de enfoques puede hacer explícito la importante relación entre perspectivas personales y del grupo, así como también proveen el significado de equipos individuales y colaborativos para articular sus propias perspectivas en un ambiente de construcción de conocimiento.

Un ambiente de creación de conocimiento que soporte perspectivas debe

proveer comparación de perspectivas sobre las cuales los usuarios puedan ver y contrastar alternativas y adoptar o adaptar ideas desde los puntos de vista de otras personas. Una comparación de perspectivas agrega ideas de varias perspectivas, sean individuales o grupales y les permite que sean fácilmente comparadas. Esta es, una fuente importante de unión de ideas que llevan a fomentar la convergencia de pensamientos y compartir interpretaciones.

El elemento mas común en los ambientes de construcción de conocimiento actuales son los foros de discusión. Los foros de discusión son sistemas de comunicación interactiva asincrónicos que le permiten a los usuarios responder notas que otros han postado. Típicamente, existe un hilo de respuestas a notas, con un árbol con opiniones divergentes.

Un requerimiento importante para construir un grupo de conocimiento es el establecimiento de conocimiento compartido. Esto puede ser fortalecido clarificando el significado de los términos más importantes usados en varios reclamos recurrentes. Un glosario sobre las discusiones puede hacer explícito la forma en que cada usuario entiende el termino que está en uso.

Tal vez una de las fases mas delicadas en la construcción de conocimiento es la negociación. El poder de la autoridad establecida resiste la negociación al cambio. El soporte de negociación informático tiende, por naturaleza, a hacer explícito los factores dentro del proceso de negociación. Esto puede ser nocivo para el proceso sutil de persuasión si no se hace con sensibilidad. Por otro lado, la negociación es crítica para ayudar a múltiples perspectivas a converger en un conocimiento compartido. El soporte informático puede proveer herramientas sumamente útiles.

La acumulación de conocimiento compartido negociado resulta en el establecimiento de una perspectiva de grupo. Como la alternativa individual y las perspectivas del equipo, la perspectiva de grupo puede ser representada en el ambiente de creación de conocimiento. El contenido de la perspectiva de grupo puede ser heredada por los individuos y las perspectivas de equipo porque ahora esta aceptada por ellos. Los individuos pueden entonces construir sobre este conocimiento compartido dentro de sus propias perspectivas e incluso comenzar a criticarlas y comenzar el ciclo completo otra vez.

El conocimiento compartido no es la fase final del proceso en el ciclo de construcción de conocimiento social. El conocimiento puede ser más formal aún. Aunque haya sido expresado explícitamente en un lenguaje formal dentro de la base de conocimiento del grupo, todavía puede ser representado en otro sistema simbólico o combinado con otro sistema de conocimiento aún más comprensible. Por ejemplo,

en la investigación académica, el conocimiento es incorporado asistiendo a una clase en la universidad, por medio de una presentación en una conferencia, un artículo en un journal y o mediante libros.

3.2. Elementos privados y compartidos

Las actividades mencionadas por Sthal como parte del ciclo de desarrollo de conocimiento colaborativo, pueden relacionarse con elementos propuestos por la informática.

Sin desconocer que uno de los elementos más fuertes en la construcción de conocimiento esta dado por la interacción social y de los individuos como parte de un grupo que se desarrolla dentro de actividades y reglas erguidas por la sociedad, en la informática existen componentes de software que pueden relacionarse con algunas de las actividades características del ciclo de construcción colaborativa de conocimiento.

El ciclo de construcción colaborativa de conocimiento puede simplificarse en función de las cuatro etapas (internalización, externalización, publicación y reacción) que presentaron [DC04][NT95] (figura 3.1). En la descripción del ciclo de construcción de conocimiento de Diaz et al., el proceso de reacción involucra todas las actividades que dan lugar en el espacio social a partir de las diferentes argumentaciones de los usuarios hasta lograr nuevas que describan el conocimiento del grupo; la publicación y la internalización son acciones que vinculan el plano personal con el grupal y por último, la externalización está descripta dentro del espacio personal.

A continuación se presentará un análisis sobre las características del soporte informático que está presente en las cuatro etapas del ciclo de construcción de conocimiento. En estas cuatro etapas se basa el resto del trabajo de grado.

3.3. Externalización

La externalización radica en el hecho de articular con palabras o mediante el lenguaje, significado que se encuentra en un plano tácito a uno explícito.

Además de ser una actividad ajena a la interacción (inmediata) social, la externalización se realiza en el fuero privado de los individuos. Sin embargo, el soporte informático puede ayudar a que la actividad de articular se realice de una forma directa. Por ejemplo, el momento en el que un usuario de un foro realiza un proceso de externalización se produce cuando escribe la nota que desea enviar; o en

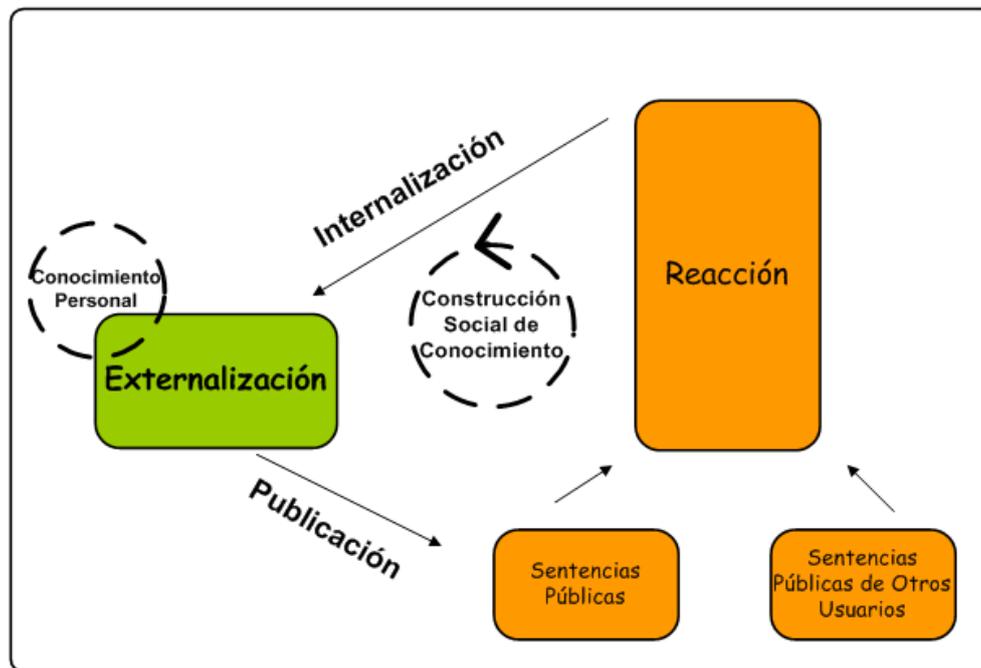


Figura 3.1: Ciclo de Stahl Simplificado [DC04]

el caso de una wiki cuando el usuario escribe en la ventana de edición.

En este punto importan las capacidades expresividad que posee el artefacto informático, ya que en función de esta cualidad será más o menos complicado para el individuo expresar la articulación del significado.

El proceso de externalización se encuentra en una etapa anterior al de publicación, sin embargo están fuertemente unidos puesto que la publicación se realizará en función de los términos en que se realizó el último proceso de externalización. Retomando el ejemplo de los foros de discusión, en estos, las opiniones personales están relacionadas con publicaciones que han realizado otros usuarios anteriormente. Las herramientas de expresión que posee cada participante del foro, están sujetas a las alternativas que posee el sistema para escribir la opinión. Además de escribir el texto de la consulta al foro, existen variantes para agregar colores, emoticones o imágenes.

Las alternativas expresivas que posee cada usuario en un foro contemplan, por ejemplo, la posibilidad de indicar que su opinión se realiza en contraposición a otra enviada con anterioridad por otro usuario, o incluir una firma que identifique sus mensajes o componer, o en base a emoticones, sus estados de ánimo. Sin embargo, en caso que desee expresar en un mismo mensaje una opinión sobre más de un hilo de conversación no lo podrá hacer directamente, sino que tendrá que describir en palabras la relación con otro comentario que se encuentra en algún hilo diferente en el mismo foro.

Es claro que el proceso de externalización sigue presente en todo momento, sin embargo el ejemplo anterior ilustra que el nivel de expresividad que posee el soporte informático puede facilitar o hacer más compleja el proceso de articular conocimiento tácito en explícito.

3.4. Internalización

Como parte del proceso individual de aprendizaje. La persona toma elementos del espacio social y los contrapone con sus propias creencias y redes de significancia. Luego de realizar diversos tipos de alineamiento y reestructuración de sus elementos, la nueva información (tal vez transformada) pasa a ser parte de sus contexto de conocimiento personal.

En la mayoría de las aplicaciones groupware, este proceso no se encuentra contemplado ya que el concepto de participación se sintetiza a la actividad de cada usuario en el espacio social y compartido.

Un acercamiento a la actividad de internalización soportada por computadoras están dados por los espacios privados de almacenamiento de información. En estos espacios los usuarios poseen elementos que pueden ser parte del dominio grupal, pero solamente visibles y manipulables por ellos mismos. En del.icio.us, por ejemplo, cualquier usuario puede navegar por los bookmarks compartidos por los usuarios e incorporar esas páginas a su repositorio privado de páginas. En esta acción, el usuario puede adaptar la página que encontró en el directorio compartido a los bookmarks que posee y hacer la categorización como privada. Es decir, ningún otro usuario puede ver a la nueva página dentro de los bookmarks del primer usuario.

La internalización, no solamente contempla la posibilidad de tomar elementos del espacio social para hacerlos parte del personal, sino que esa incorporación lleva consigo la necesidad de establecer nuevas relaciones entre los conceptos y los credos anteriores a la incorporación. Lamentablemente, la reorganización a este nivel no se encuentra actualmente fuertemente soportada por aplicaciones informáticas. Gran parte del proceso de re-estructuración de los conceptos y credos, pertenecen a un plano de conciencia personal, es decir, los realiza el usuario internamente.

La idea de soportar computacionalmente la internalización, exige formas de representación adecuada, alternativas de reorganización, análisis de la información y practicidad para el acceso a los recursos , entre otras cosas.

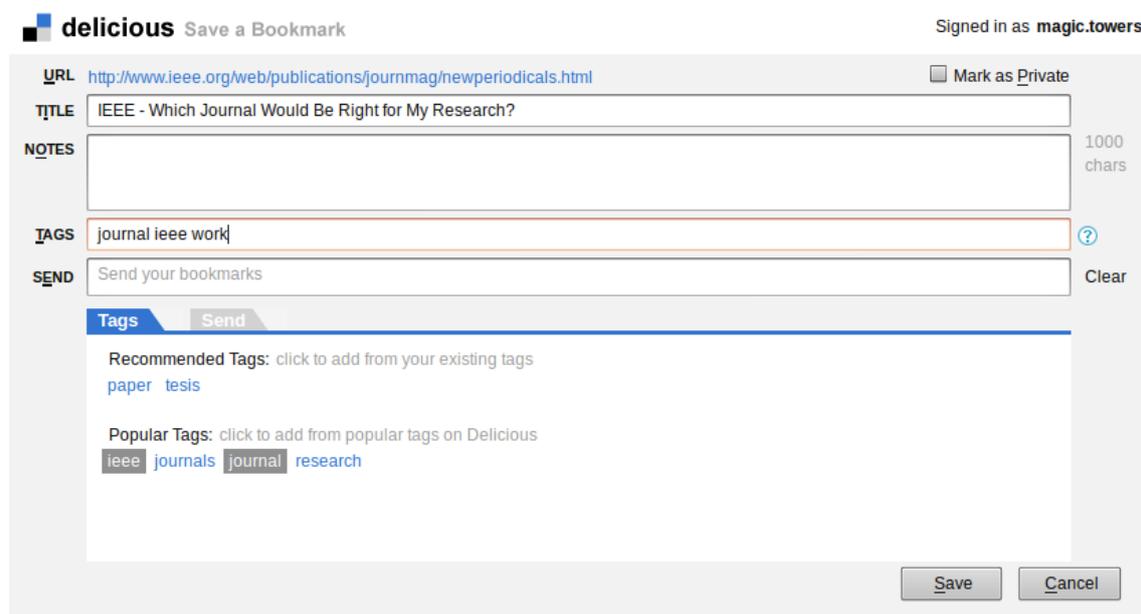
3.5. Publicación

La publicación implica hacer público el conocimiento (personal), transfiriéndolo desde el plano del conocimiento individual al plano del conocimiento social.

Está sujeto a realizarse inmediatamente después del proceso de externalización. En el ejemplo del foro, podría considerarse que se realiza esta actividad en el momento que el usuario decide presionar en el botón que envía la publicación.

En los sistemas informáticos actuales, el proceso de publicación se encuentra presente con la salvedad que los mismos sistemas definen la forma y los términos en los que se articula la externalización.

En un foro la publicación se realiza enviando la opinión sobre otro post realizado. En un sistema de negociación, la expresión de un punto de vista o hacer efectiva una votación es una forma de publicar. En la figura 3.2 podemos ver el momento en que un usuario decide indicar qué bookmarks quiere utilizar para describir al sitio de IEEE. La escritura de las palabras *journal ieee* y *work* son el resultado del proceso de externalización, cuando el usuario presione el botón *Save*, realizará finalmente el proceso de publicación.



The screenshot shows the 'delicious Save a Bookmark' interface. At the top, it says 'Signed in as magic.towers'. The main form has the following fields: 'URL' with the value 'http://www.ieee.org/web/publications/journmag/newperiodicals.html', 'TITLE' with 'IEEE - Which Journal Would Be Right for My Research?', 'NOTES' (empty, with a '1000 chars' limit), 'TAGS' with 'journal ieee work', and 'SEND' with 'Send your bookmarks'. Below the form, there are two sections: 'Recommended Tags' with 'paper' and 'tesis' (clickable), and 'Popular Tags' with 'ieee', 'journals', 'journal', and 'research' (clickable). At the bottom right, there are 'Save' and 'Cancel' buttons.

Figura 3.2: Externalización y Publicación con Del.icio.us

3.6. Reacción

Es el acto por el cual se responde a una nueva contribución dentro del ámbito social. La reacción en sí es una nueva contribución que torna más compleja la

contribución inicial. Se genera nuevo conocimiento que es producto de diferentes puntos de vista.

Como indicamos anteriormente, el soporte informático para el ámbito social dentro del ciclo colaborativo de construcción de conocimiento es el más desarrollado. La mayoría de los sistemas presentan herramientas que sustentan la reacción, veamos algunos ejemplos:

En los foros de discusión, un nuevo aporte a un hilo existente es una reacción del grupo (realizado por un individuo particular.) Para un blog, los comentarios agregados a un nuevo artículo pueden ser considerados elementos de un proceso de reacción. En el sitio youtube, existen los videos que son respuestas a otros videos. Esto también puede ser considerado reacción.

Lo interesante de la reacción en si misma, y relacionada con el resto de los procesos que son parte del ciclo de construcción colaborativa de conocimiento, radica en que los nuevos aportes son parte de un proceso que engloba los pasos anteriores: internalización, externalización y publicación.

Si tomamos el ejemplo del foro de discusión, cuando un usuario responde a un post realizado por otro, en primera instancia realiza un proceso por el cual toma los elementos significativos el post en cuestión, interpreta esa información y la adecua a sus estructuras internas. Luego de este proceso arduo, articula en palabras el resultado de incorporar la nueva información a su estructura y esa articulación, luego, la publica en el espacio personal como respuesta.

La riqueza de los nuevos puntos de vista y de la forma en que se genera el conocimiento consensuado por el grupo, depende de las capacidades de realizar las etapas anteriores.

3.7. Conclusiones del Capítulo

En este capítulo se pudo analizar la forma en que se le da sustento informático al ciclo de desarrollo colaborativo de conocimiento en términos generales y que características deben poseer.

Luego analizamos, en función de aplicaciones groupware conocidas, la presencia de los cuatro estadios: internalización, externalización, publicación y reacción. Como corolario, las actividades de internalización y externalización parecen no estar incorporadas a los sistemas de manejo de conocimiento para los que los usuarios deben adaptarse a las alternativas que les ofrecen.

Queda planteado para los próximos capítulos cómo se adecua este ciclo de

construcción de conocimiento a las wikis semánticas.

Capítulo 4

Información personal en una wiki semántica

En este capítulo se detallará cómo se soporta la construcción colaborativa de conocimiento en una wiki semántica, luego una comparación entre las implementaciones más populares de wikis semánticas y la forma en que utilizan los espacios privados y compartidos. En la sección 4.2 analizaré los lineamientos para definir un modelo de espacio personal y compartido. En la sección 4.3 presentaré un enfoque para incluir anotaciones personales en una wiki semántica como medio para lograr los espacios de conocimiento personal y compartido y para finalizar, en la sección 4.4 las conclusiones del capítulo.

4.1. Wikis Semánticas como Soporte Para la Construcción de Conocimiento

El enfoque presentado en el capítulo 3 esquematiza el ciclo de construcción colaborativa de conocimiento en dos grupos de actividades: aquellas centradas en el espacio personal (internalización y externalización) y aquellas que se destacan en el espacio social (publicación y reacción).

Como fue indicado en el primer capítulo, no existen wikis semánticas que presenten la conjunción de los dos espacios para la construcción de conocimiento. Existe un enfoque de creación de conocimiento personal en SemperWiki[OVBD06], sin embargo esta simplemente articula el espacio personal y no existe posibilidad de trasladar ese conocimiento a un ámbito compartido. Por otra parte, Semantic Media Wiki, OntoWiki, SWiki y otras wikis semánticas actuales presentan el enfoque opuesto: solamente manejo de conocimiento en un ámbito compartido.

Durante esta sección analizaremos las alternativas dentro del espacio compartido basadas en el ejemplo de Semantic Web¹, un sitio destinado a agrupar información relacionada a la web semántica que utiliza como motor a Semantic Media Wiki (SMW) y esquematizaremos las necesidades de incluir un espacio privado para manejar conocimiento semántico dentro de una wiki semántica.

4.1.1. Espacio compartido

Publicación

Las publicación en SMW queda definida en las acciones de agregar una nueva página, o editar información de una página existente. Para agregar información semántica a una página se requiere editar la misma. Recordemos que en SMW la forma de agregar información semántica es utilizando un lenguaje de markups.

Agregar una nueva página determina agregar un nuevo Concepto a la base de conocimiento de SMW. La forma de agregar un nuevo concepto es utilizando el cuadro de búsqueda. Si el concepto buscado se encuentra, entonces el sistema guía al usuario a la página existente, caso contrario le permite agregar la nueva página. La figura 4.1 ilustra el resultado de una búsqueda que no posee páginas asociadas y la posibilidad de crear la página.

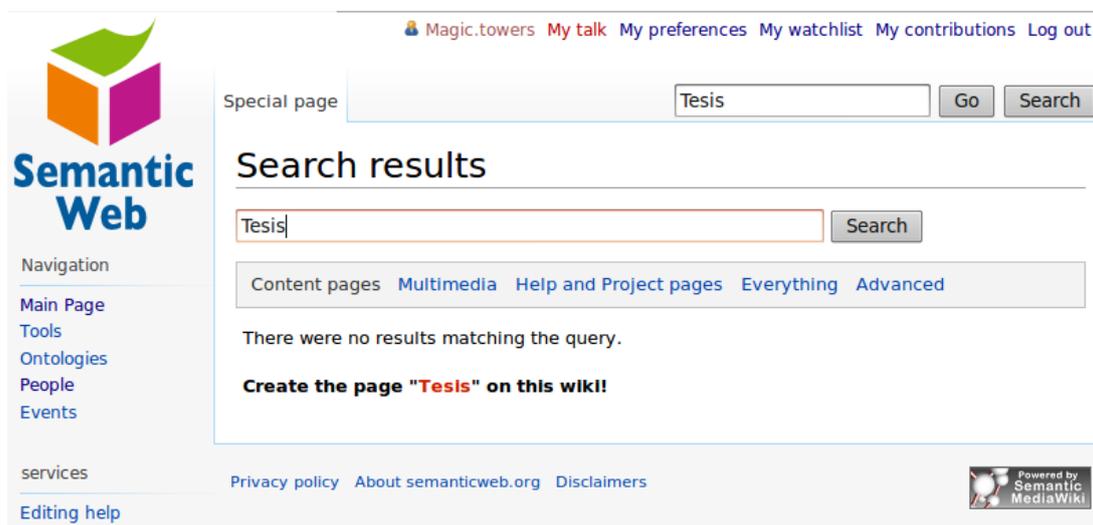


Figura 4.1: Publicación de un nuevo concepto en SMW

El caso para el cual el concepto ya existe en la SMW no es un caso menor. En esta oportunidad se presenta, en cierta medida, un conflicto entre el espacio conceptual que posee el usuario y el que está definido para la comunidad de la wiki.

¹<http://semanticweb.org>

El conflicto puede resolverse si alguna de las páginas ya creadas, se corresponde con la conceptualización del usuario. En el ejemplo de la figura 4.2 si el concepto “Diego Torres” se refiere a la persona que trabajo en el laboratorio de investigación LIFIA, no habrá conflicto. Sin embargo, si la conceptualización de la wiki semántica no es la misma a la que considera el usuario, estamos en presencia de una discrepancia en los conceptos. Por ejemplo, que el concepto sea el cantante argentino. En este caso, el usuario debe readecuar sus conceptualizaciones si es que desea incorporarlas al espacio compartido, caso contrario se verá imposibilitado de expresar esa pieza de conocimiento en la wiki.



Figura 4.2: Publicación de un nuevo concepto en SMW con conflicto

La otra forma que tiene un usuario para publicar, es agregando información semántica a una página existente. Siguiendo con el ejemplo de Semanticweb.org, deseamos agregar información semántica a la página que describe la “Universidad Nacional de La Plata”².

En este caso vamos a publicar dos elementos de información nuevos. Por un lado información sobre su ubicación, para ello agregamos el link tipado [[located::Argentina]]. Además, indicamos que la UNLP pertenece a la categoría *Trabajo* escribiendo [[Category:Trabajo]], ya que el usuario realiza actividades laborales en la institución. De esta forma, indicamos que la UNLP esta ubicada en *Argentina* y posee la categoría *Trabajo*.

En la figura 4.3 podemos visualizar en la parte inferior el *factbox*. Un factbox es un cuadro donde se detalla la información semántica que aparece en la página wiki que se encuentra por encima de este. En el caso del ejemplo, podemos ver que

²[http://semanticweb.org/wiki/Universidad_Nacional_de_La_Plata_\(UNLP\)](http://semanticweb.org/wiki/Universidad_Nacional_de_La_Plata_(UNLP))

la página de “Universidad Nacional de La Plata UNLP” contiene un link tipado (property) llamado *located* el cual está relacionado al concepto Argentina. Sobre el factbox se detallan las categorías a las que pertenece la página donde, además de *University*, indica que pertenece a *Trabajo*.

De estos ejemplos se desprenden algunas de las consecuencias propias de la publicación. En los primeros casos puede ocurrir que el concepto que intenta publicar el usuario ya se encuentre en el espacio compartido con un significado diferente, para lo cual, las únicas alternativas que posee son modificar el término que describe su concepto y adaptarse a los que determina el espacio compartido o resignar la publicación como medida extrema.

Por otra parte, en el segundo ejemplo, la categorización de la UNLP como elemento agrupado bajo la categoría *Trabajo* posee una semántica íntimamente ligada a la información personal del usuario. Es muy probable que la mayoría de los participantes que están alrededor de la comunidad Semanticweb.org no realicen tareas laborales en la UNLP. Por lo tanto, la información no es representativa del conocimiento grupal.

Universidad Nacional de La Plata (UNLP)

The Universidad Nacional de La Plata is located in [Argentina](#).

<http://www.unlp.edu.ar/>

Categories: [University](#) | [Trabajo](#)

Facts about Universidad Nacional de La Plata (UNLP) ⓘ

[RDF feed](#) 

[Located](#) [Argentina](#) + 

⌵

Figura 4.3: Agregado de información semántica en SMW

Reacción

Las acciones realizadas en los ejemplos de publicación (sección 4.1.1) pueden generar reacciones del resto de los integrantes de la comunidad semanticweb.org. La reacción involucra, en términos generales, el intercambio de opiniones y creencias a fin de determinar nuevas publicaciones que tornan más compleja la publicación inicial.

Como primer medida tomemos el caso del link tipado `[[located::Argentina]]`. Si accedemos a la página que define este link podemos visualizar una leyenda que indica: “Utilice la propiedad *Located In* en su lugar.” Esto se debe a que *Located In*, ha sido consensuada por la comunidad para significar el lugar donde se encuentra ubicado algún elemento o concepto. Teniendo en cuenta esta información, la página puede ser editada nuevamente por otro usuario modificando el nombre de la relación *located* por *located in*, quedando escrito `[[located in::Argentina]]` en el código de la página.

También, podrían haber considerado que la categorización *Trabajo* no era significativa para el grupo. En este caso, la discusión se vio reflejada en un estadio anterior. SMW posee espacios para realizar discusiones en torno a un concepto (página o propiedad) y en virtud a estas discusiones ir definiendo la definición consensuada. Es claro que no todas las publicaciones son modificadas, en muchos casos las publicaciones son aumentativas, es decir, incorporan nuevo conocimiento al concepto descrito y lo hacen más rico. En estos casos, la reacción de la comunidad se sitúa en un rol de avalar la publicación.

Las discusiones y las reacciones en SMW se pueden desarrollar en el ámbito de discusión de cada página, editando o mejorando publicaciones anteriores, modificando el contenido en función del historial de publicaciones o realizando combinaciones de todas estas alternativas. El resultado que se presenta en cada página de la wiki es producto de un consenso gradual que se produce como resultado de la autoría de la comunidad en si misma. Este proceso es un proceso iterativo constante que busca definir conceptualizaciones y una definición ontológica que describa el conocimiento de la comunidad.

En su artículo “Emergent Semantics for Ontologies” [SSN⁺02], Maedche indica que las personas requieren palabras (o símbolos en general) para hablar y comunicar sobre cosas. La relación de las palabras a las cosas es indirecta y toma desvíos intrincados entre conceptos. El significado dado a una palabra dentro de un contexto y sus referencias a cosas concretas es dado por un concepto. La relación de palabras a conceptos es el resultado de un proceso social de comunicación.

Las ontologías, también en una wiki semántica, tratan de formalizar lenguaje natural para que pueda ser procesable y comprendido por una computadora. Sin embargo, un aspecto importante es que las personas deben estar de acuerdo con la ontología específica y sus intenciones de conceptualización. El desarrollo de ontologías en forma colaborativa, indica Maedche, involucra a un proceso cooperativo de estandarización el cual posee un punto sumamente crucial: la comunicación entre los miembros. Estos deben llegar a un acuerdo final en la

definición del standard. En esta parte del proceso, la parte de reacción, debemos soportar la evolución de un conocimiento compartido, no solamente en elementos formales semánticos, sino que también en estructuras cognitivas humanas y la forma en que las personas puedan comunicarse a fin de hacer emerger, como corolario de este proceso, conocimiento de la comunidad que rodea a la wiki semántica.

Universidad Nacional de La Plata (UNLP)

The Universidad Nacional de La Plata is located in [Argentina](#).

<http://www.unlp.edu.ar/>

Category: [University](#)

Facts about Universidad Nacional de La Plata (UNLP)

RDF feed

Located in [Argentina](#) +

Figura 4.4: Agregado de información semántica en SMW

4.1.2. Navegabilidad

Las anotaciones semánticas permiten categorizar una página wiki o definir relaciones semánticas. Estos dos tipos de anotaciones semánticas proponen dos alternativas para expresar navegación.

- Representar un link tipado entre páginas. Estos links se embeben dentro del código wiki. Al tener los links un tipo que establece el significado de la relación entre las páginas representan una mejora en la navegación utilizando los links tradicionales.
- La otra alternativa navegacional es utilizando categorías, como vimos en los ejemplos anteriores. Cada categoría posee su propia página wiki. Las páginas de las categorías son útiles para navegar la wiki semántica, ya que cada una de estas páginas es un punto de acceso a las páginas wikis que han sido categorizadas por el concepto subyacente de la categoría actual. Las páginas de categorías implementan lo que Brusilovsky denomina navegación basada en conceptos[Bru07].
- Los Facbox son otra forma de navegabilidad propia de las wikis semánticas. En los factbox se detalla en un recuadro un resumen de toda la información

semántica que se refleja en la página wiki, pudiendo de esta forma navegar a las páginas de las propiedades, atributos y conceptos que son parte de las definiciones semánticas de la página actual.

En la mayoría de las wikis semánticas, la navegación es compartida por los usuarios. El espacio navegacional es producto de la actividad colaborativa que se realiza cada vez que un usuario categoriza una página o agrega una relación semántica. Sin embargo, no proveen espacios de navegación personal.

4.1.3. Espacio privado

Como analizamos en la sección anterior, el usuario no posee un espacio privado donde pueda realizar el mismo tipo de conceptualizaciones que se realizan en el espacio público. En otras palabras, el usuario no tiene la posibilidad de utilizar conceptos o relaciones que no estén consensuadas por la comunidad o si bien fueron consensuadas que no sean visibles por los demás usuarios.

SMW, al igual que el común de las wikis semánticas, no ofrece un espacio en el cual los usuarios puedan depositar conceptualizaciones sin que entren en contacto con las demás en el espacio público.

Es importante que las wikis semánticas incluyan como parte de la interacción que les brindan a sus usuarios, elementos del espacio personal propios del ciclo de construcción colaborativa de conocimiento que se desarrolla continuamente en una wiki semántica. Este espacio debe brindar capacidades para expresar conceptualizaciones, diferenciar elementos privados de los compartidos y poseer privacidad en el acceso, entre otras cosas.

4.2. Espacio Privado y Espacio Compartido

Una de las formas de determinar la diferencia entre un espacio privado y un espacio compartido, está ligado al acceso a la información. En el caso de los espacios compartidos, suponemos que la información es accedida por los miembros de la comunidad que nuclea la wiki semántica, mientras que en el caso de los espacios privados son aquellos que, dentro del espacio general de la wiki semántica, son gerenciados y accedidos únicamente por sus dueños.

Sin embargo este enfoque no es completo si tenemos en cuenta que los espacios con los que estamos tratando están ligados a la creación de conocimiento. Considerar a la wiki semántica simplemente como un repositorio deja de lado elementos que

facilitan acceso a la información. La información no solamente está presente y es accesible, sino que también esta organizada y posee relaciones. Cada concepto posee una serie de relaciones con otros conceptos, estas relaciones además de brindar una descripción generan un lazo navegacional. En las wikis semánticas también existen links entre páginas, que aunque no posean una carga de significado tan fuerte como los links tipados hacen que el acceso a los conceptos sea inmediato.

Asimismo, el espacio de conocimiento personal no debe estar aislado del espacio compartido en el que está inmerso cada individuo. Dentro de la red de conceptualizaciones y credos de cada persona, se encuentran muchos elementos que fueron tomados y son descriptos en el espacio compartido y que se encuentran, posiblemente, relacionados con elementos propios de cada espacio de significancia personal. Un ejemplo de esto es la categorización de la página “Universidad Nacional de La Plata (UNLP)” como elemento de la categoría *LugarDeTrabajo*. Esta categorización relaciona la conceptualización del lugar de trabajo personal con el concepto compartido UNLP.

En lo siguiente se brindará un modelo de personalización en wikis semánticas, el cual permita definir conceptualizaciones personales en un espacio diferente al de las conceptualizaciones compartidas, que cada usuario posea control y acceso a sus elementos personales y por último, que se permita definir un espacio de articulación entre el universo compartido y el personal.

4.3. Información Personal en una Wiki Semántica

Esta sección presenta un enfoque para agregar anotaciones personales para el caso de una wiki semántica P2P. La idea central es extender la funcionalidad de los sistemas wikis semánticos para que consideren en forma completa el proceso de creación colaborativa de conocimiento.

El análisis para definir este espacio comprende la capacidad de expresión, el reuso del lenguaje compartido, la capacidad navegacional y la forma de enriquecer el conocimiento compartido.

Como punto de partida al enfoque de personalización se propone un modelo de anotaciones semánticas personales utilizando Categorías e Individuos.

4.3.1. El espacio personal

El espacio personal debe permitir al usuario dueño de ese espacio realizar anotaciones personales semánticas de manera que solamente sean accedidas por él mismo. Es importante que este espacio garantice privacidad de la información y control de acceso.

Sin embargo, se debe proveer de un sistema de comunicación por el cual la información del espacio compartido pueda ser accedida desde el espacio personal.

4.3.2. Anotaciones Personales

La descripción del espacio de información personal se basa en la escritura de anotaciones semánticas personales. Las anotaciones personales semánticas son el medio de expresión que posee el usuario para realizar el proceso de externalización.

Como fue indicado anteriormente, el proceso de externalización se realiza dentro del plano personal y está basado en la articulación de las conceptualizaciones utilizando el lenguaje. La capacidad expresiva de las anotaciones personales, deben estar al mismo nivel que las utilizadas en el espacio compartido. De esta forma, los procesos de publicación y de externalización pueden adecuarse más fácilmente y no requieren de posteriores procesos de readecuación del lenguaje y las conceptualizaciones.

4.3.3. Anotaciones Semánticas Personales: Categorías e Individuos

La idea de este sistema de anotaciones semánticas personales es poder categorizar cada una de las páginas en base a etiquetas con información semántica. Utilizando el enfoque de folksonomies, cada uno de los usuarios puede categorizar las páginas de la wiki semántica.

Además de las anotaciones compartidas embebidas en el texto de una página wiki, los usuarios pueden además asociar anotaciones semánticas personales a las páginas wikis. Estas anotaciones personales expresan la comprensión personal de los usuarios.

Por ejemplo, si un usuario está navegando la página wiki “Semantic Wiki” como se muestra en la figura 4.5, eventualmente podría anotar esta página como “Collaborative Tool”, “Web” y “Semantic Wiki”. Si estas anotaciones expresan solamente conocimiento personal, entonces deberían ser privadas. En cambio, otras

anotaciones como “Semantic Web” o “Wiki” podrían ser compartidas. Estas últimas podrían ser definidas por el mismo usuario o por cualquier otro del sistema. En este ejemplo, podemos visualizar que el mismo usuario interactúa simultáneamente con anotaciones personales y compartidas.

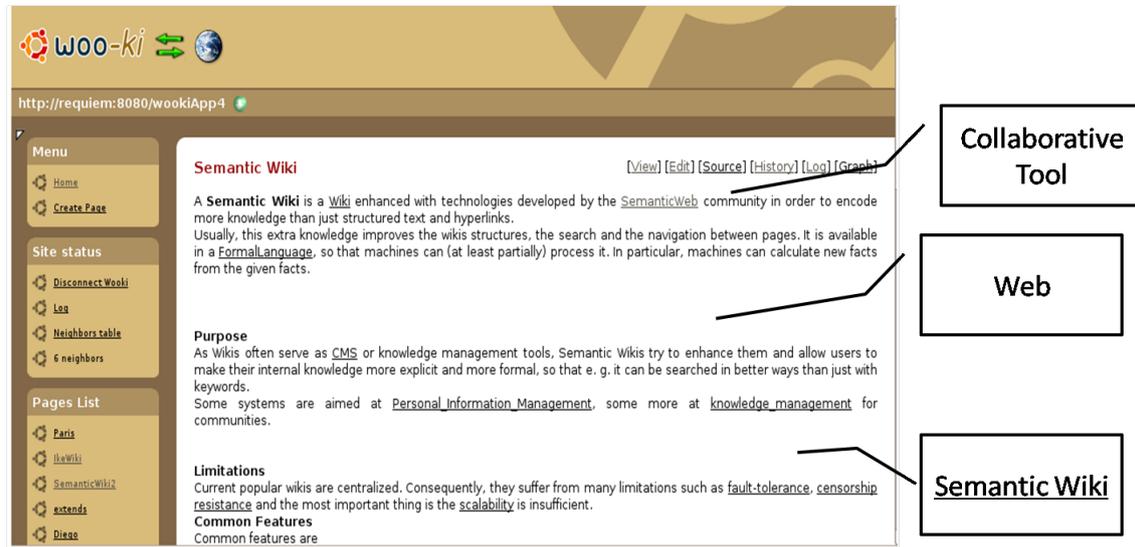


Figura 4.5: Agregado de Anotaciones Personales a una Wiki Semántica

Cada página semántica de la wiki puede ser categorizada con varias anotaciones semánticas personales. Una anotación semántica personal puede ser una Categoría o un Individuo.

Las Categorías definen una familia de elementos. Por ejemplo, en el ejemplo anterior la anotación Semantic Wiki se utilizó para indicar que la página es en si misma la definición de la categoría que involucra a las Wikis Semánticas.

Los Individuos denotan elementos que pertenecen al menos en una categoría. Por ejemplo, Semantic Media Wiki es un individuo que pertenece a la categoría Semantic Wiki. Además, un Individuo puede pertenecer a varias categorías.

Cada página de una wiki semántica puede ser anotada utilizando varias anotaciones personales. Por ejemplo, un usuario desea anotar la página “Swooki” como Individuo de las categorías “Semantic Wiki” y “P2P application”, ya que considera que Swooki es una wiki semántica y además es una aplicación P2P.

Actualmente el modelo es simple y permite definir solamente Categorías e Individuos, sin embargo se espera que en un futuro se puedan incorporar al espacio personal la capacidad de definir relaciones y atributos. Permitir expresar conocimiento personal con relaciones y atributos achicaría la brecha expresiva en el momento de realizar la externalización. Sin embargo tomamos las categorizaciones como punto de partida.

4.3.4. Reuso de conceptualizaciones compartidas

En el ejemplo sobre la wiki semántica Semanticweb.org, se intentaba incorporar información de carácter personal a la página referida a la “Universidad Nacional de La Plata”. En ese caso, la información personal estaba referida al lugar donde trabaja el usuario. Esta información, de carácter netamente personal incluye información definida en el espacio compartido, como es el caso de la Universidad Nacional de La Plata.

Categorizar a la página de la UNLP puede quedar definido en el plano personal. En esta caso, la anotación que indica que UNLP pertenece a una categoría llamada Trabajo es parte del espacio personal del usuario. Eso significa que si otros usuarios intentan acceder a la página de la UNLP, no van a percibir la categorización que realizó el usuario anterior.

Por su parte, el usuario que realiza la anotación personal, realiza un proceso de internalización tomando elementos directos de la conceptualización compartida y los incorpora a su espacio personal, complementándolos con elementos de su fuero privado. Esta combinación permite incorporar elementos del plano compartido a las conceptualizaciones personales. De esta forma, se reutiliza lenguaje y definiciones que la persona incorpora a su espacio personal. Además la relación es directa, permitiendo que sea aun más fácil realizar publicaciones sobre elementos del espacio compartido que aún no fueron considerados por la comunidad y que son parte del espacio privado de conceptualizaciones.

4.3.5. Navegabilidad

Como vimos anteriormente, el espacio navegacional en una wiki semántica tradicional se ve reflejado por las anotaciones semánticas que realizan los usuarios. Al estar estas anotaciones en el espacio compartido, la navegabilidad esta definida en términos del espacio compartido por todos los usuarios. Por lo tanto, no existe la posibilidad de definir navegabilidad personalizada para los usuarios.

La incorporación de anotaciones personales semánticas permiten definir un nuevo espacio navegacional que complementa al compartido. Las anotaciones personales están definidas solamente para el usuario que las define y permiten combinar elementos del espacio privado, con elementos del espacio compartido. De esta forma es posible navegar desde una categorización personal a una categorización compartida y viceversa. Podemos determinar una serie de ventajas con relacionadas con la navegabilidad:

Navegación personalizada Como las anotaciones personales son manipuladas en un espacio privado, estas son solamente accesibles por el usuario que las creó. Por lo tanto, los usuarios pueden adaptar la navegación de acuerdo a sus necesidades.

Navegación basada en conceptos Los usuarios pueden definir nuevas categorías utilizando anotaciones semánticas personales, por lo tanto pueden definir una navegación personalizada basada en conceptos.

Complemento del espacio navegacional compartido El espacio navegacional privado se suma al espacio compartido. No se alteran uno al otro sino que simplemente se complementan. De esta forma el hiperespacio navegacional compartido se ve aumentado por el personal.

4.4. Conclusiones del Capítulo

En este capítulo analizamos las características que brindan las wikis semánticas tradicionales. Lamentablemente no existen wikis semánticas que articulen los espacios de conocimiento personal y compartido.

Se definió un conjunto de anotaciones personales que permitan describir elementos del espacio compartido en un espacio personal y privado. Estas anotaciones se caracterizan por actuar como tags de los conceptos que representan las páginas wikis distinguiéndose entre Categorías e Individuos.

Estas anotaciones incrementan el espacio expresivo compartido incorporando elementos personales y además establecen nuevos espacios de navegación. Al espacio compartido se le suma el espacio personal, el cual además permite que los usuarios puedan personalizarlo.

Capítulo 5

P-SWOOKI

Como resultado del análisis detallado anteriormente surge P-Swooki. P-SWooki es una wiki semántica que permite realizar anotaciones semánticas personales sobre el contenido de la wiki. Este prototipo fue basado en SWOOKI, una wiki semántica P2P implementada en Java utilizando servlets dentro de un servidor Apache Tomcat.

5.1. SWOOKI: Una wiki semántica P2P

Swooki es la primera wiki semántica con arquitectura P2P. Por su parte, SWOOKI integra las tecnologías involucradas en la web semántica siguiendo la filosofía de Semantic Media Wiki. Las anotaciones semánticas son embebidas en el cuerpo de la página wiki y se distinguen del resto del texto utilizando un lenguaje de mark ups especiales.

En SWOOKI una ontología formal emerge a través de la edición colaborativa de la misma. Los usuarios editan y agregan información a la wiki de la misma forma

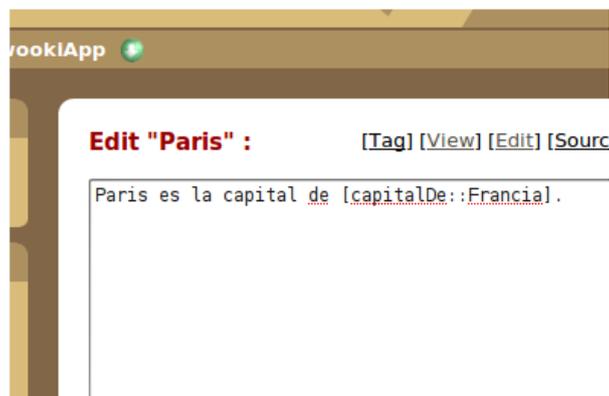


Figura 5.1: Pantalla en modo edición: agregado de una anotación semántica

que en una wiki que posee un único servidor centralizado. Los servidores pueden estar conectados par a par en una red de servidores o pueden estar desconectados. Para los servidores conectados, cada vez que se guarda un cambio en una pagina particular, el mismo se propaga a las demás servidores; en el caso de los servidores desconectados de la red, cuando se realiza la primer conexión con otros pares se replica la información con el resto de la red.

Al ser una arquitectura P2P la cantidad de usuarios que pueden colaborar alrededor de una comunidad puede ser sumamente grande si se considera el problema de los cuellos de botella que se establecen al querer acceder simultáneamente a un único servido centralizado. En esta arquitectura P2P la carga de usuarios puede balancearse por cada par (servidor) los cuales se encargan de sincronizar solamente aquella información que se no se encuentra en los otros pares. Como resultado se logra una colaboración masiva de una forma que minimiza cuellos de botellas a los recursos.

En cualquier momento cada uno de los pares puede desconectarse de la red y trabajar en modo off-line. En este modo las modificaciones se presentan solamente sobre el servidor particular al que se conectan los usuarios pero, al estar desconectado, no realiza tareas de replicación y sincronización del resto de la red. Con esta configuración, es posible tal vez que en una organización se distribuyan tareas entre diferentes grupos. Cada grupo posee a su cargo una servidor Swooki sobre el cual vuelca contenido y lo somete a verificaciones cotidianas. De esta forma, hasta que cada grupo no consolida su actividad interna y considera que el contenido producido no está terminado, no se conecta el servidor al resto de servidores de la red. Luego, cuando los diferentes servidores se conectan entre si, las tareas de sincronización y replicación generan que todos los servidores se vean enriquecidos con la participación y generación de conocimiento que se gestó en los otros pares.



Figura 5.2: Pantalla en modo normal: visualización de una anotación semántica

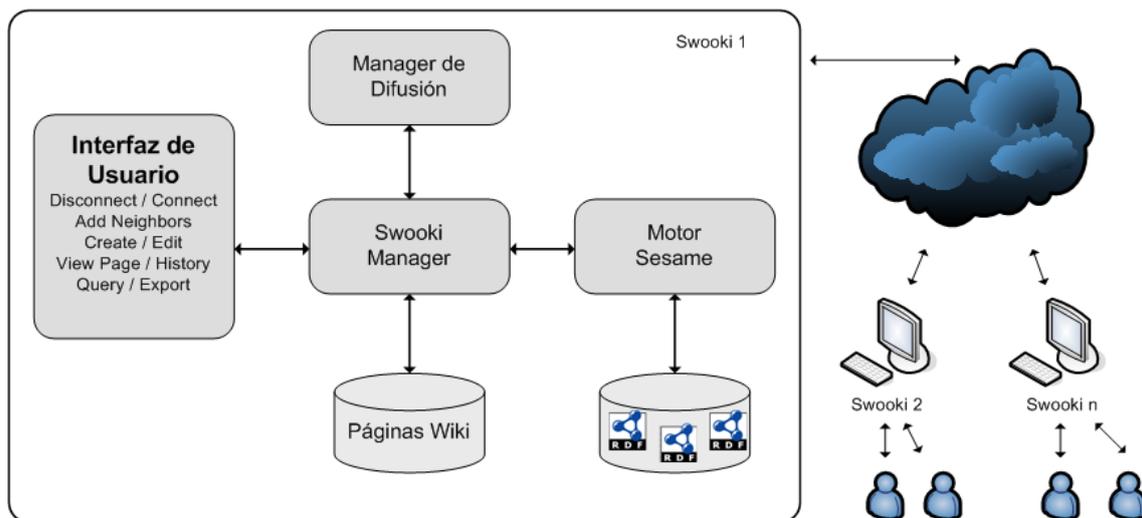


Figura 5.3: Swooki: arquitectura

En otras palabras, la edición off- line permite un trabajo local profundo sobre la generación de contenido y luego de validarse puede compartirse con el resto de la red.

5.1.1. Arquitectura

Swooki está implementada utilizando servlets Java y Tomcat como servidor de aplicaciones, además utiliza Sesame 2.0 como repositorio RDF. En Swooki, cada par es un servidor wiki. Es decir, cada servidor se encarga de gestionar la edición, creación y borrado de páginas, permite conectarse a otros servidores, gestionar la replicación y consistencia de la información y administrar el contenido semántico.

La figura 5.3 muestra la arquitectura básica dentro de una red de servidores swooki unidos a una misma red. A continuación se detallan las componentes principales de la arquitectura de un server particular.

Interfaz de usuario El componente de interfaz de usuario posee la funcionalidad de un editor wiki común. La única diferencia con un editor wiki es que realiza el envío de operaciones al mánager de SWOOKI.

Motor Sesame El motor Sesame es un repositorio RDF. Es controlado por el Manager para guardar y obtener ternas RDF. Todos los pedidos son delegados al repositorio Sesame. El motor Sesame es el responsable del parseo, búsqueda, guardado y extracción de información semántica. Los usuarios pueden utilizar SPARQL y SeRQL queris.

Manager de Difusión El mánager de difusión es el encargado de mantener la

unión del peer con la red y el de realizar el broadcasting de información hacia los otros pares. La difusión de información asegura que todos los pares conectados a la red recibirán la información difundida y no existirán particiones en la red P2P. Para los pares que están desconectados se implementa un algoritmo denominado por los autores de anti entropía el cual selecciona en forma aleatoria un vecino en la tabla de vecinos local a cada peer y le envía un mensaje con el resumen de los mensajes que ha recibido.

5.1.2. Definición Semántica

Cuando un usuario edita una pagina de SWOOKI puede agregar anotaciones semánticas (además de información textual). La forma de representación semántica está basada en links semánticos o link tipados. Un link semántico posee dos propósitos: un propósito navegacional y otro referido al significado. Estos links son tipados o categorizados para indicar el tipo de la relación que representan [Adaptive Hypermedia and Adaptive Web-Based Systems LNCS].

SWOOKI considera a cada página wiki como un concepto, los cuales pueden estar semánticamente relacionados por hechos (*facts*). Las relaciones entre los conceptos de la wiki permite que emerja la definición ontológica de la wiki.

Los facts se construyen a través de links semánticos. La forma de escribir un link semántico en Swooki es utilizando la notación [tipoDelLink::PaginaWiki]. A través de la creación de estos conceptos se va creando una definición ontológica subyacente del conocimiento volcado en la wiki. En las figuras 5.1 y 5.2 podemos ver la forma de escritura de una anotación semántica y su visualización respectivamente. La figura 5.1 muestra como se define la relación “capitalDe” entre la página que es editada - Paris - y la página Francia. Luego, en la figura 5.1 se puede apreciar que la palabra *Paris* es un link a la página con el mismo nombre y en la parte inferior la información semántica explicitada en la lista de *facts* (hechos).

5.1.3. Alcance

Cada uno de los servidores SWOOKI provee todos los servicios de un servidor de wiki semántica. Los autores analizan las diferentes prestaciones de acuerdo a los siguientes criterios:

Disponibilidad, tolerancia a fallas y balanceo de carga Los datos están disponibles en cada peer por lo tanto es accesible aun cuando alguno de los peers no está disponible. Los nombres globales de cada concepto (página

wiki), sus relaciones y los algoritmos de sincronización y propagación aseguran que la misma información se encuentra en cualquier nodo de la red.

Escalabilidad SWOOKI provee escalabilidad con respecto al número de pares.

Cada par conoce solamente a un número fijo de pares obtenidos en forma aleatoria. Estos manejan una tabla de vecinos que posee un tamaño fijo y contiene una lista parcial de los nodos en la red P2P que continuamente evoluciona pero jamás excede el tamaño fijo. El algoritmo de diseminación actualiza esta tabla durante mensajes de propagación donde se utilizan suscripciones y de suscripciones adosadas a cada mensaje. Esto no provee una solución de escalabilidad referido al tamaño de la información ya que la capacidad para guardar información está limitada en la capacidad de almacenamiento de cada nodo. SWOOKI provee escalabilidad con respecto al tamaño de la comunidad.

Trabajo off-line y cambios transaccionales Los usuarios pueden trabajar desconectados si es que no poseen conexión a internet o si ellos deciden trabajar desconectados desde la interfaz de usuario. Mientras se encuentra el nodo desconectado, los usuarios pueden realizar cambios en las páginas obteniendo consistencia en los mismos. De este modo, el usuario genera una serie de transacciones. Todos los cambios realizados en el modo desconectado son guardados en la componente de difusión. Como el algoritmo de replicación fuerza a todas las operaciones a conmutar entonces la ejecución concurrente de las diferentes transacciones produce un estado consistente en todos los casos.

Al ser una wiki semántica con las mismas prestaciones que presenta un servidor tradicional del estilo de Semantic Media Wiki, SWOOKI permite que los usuarios realicen aportes sobre un repositorio general y compartido de conocimiento (el cual está distribuido en diferentes nodos).

A medida que un usuario realiza un aporte en la wiki, por ejemplo agregar una nueva página, esta información es replicada y pasa a ser parte de la comunidad que está formada alrededor de la red de servidores SWOOKI. Bajo estas restricciones, el contenido presente en la wiki es siempre contenido que fue creado por la comunidad. En otras palabras, no es posible diferenciar información propia de un usuario e información que sea de todo el grupo.

Para representar información personal, podríamos especular con tener un nodo para cada usuario, de esta forma la información volcada en estos nodos no sería

compartida con los demás miembros de la red siempre y cuando cada uno de los nodos siempre se encuentre desconectado de otros. Al momento de conectarse, la información de cada nodo se replicaría en el resto de los nodos y toda la información originada en cada nodo pasaría a ser información común a todos los nodos de la red. El elemento personal se habría perdido.

En este tipo de wikis, se pierde en cierta medida la noción de autor personal de conocimiento. El contenido de información que posee la wiki es contenido creado por la comunidad en su totalidad. Detrás de este concepto de autoría grupal se describen dos tipos de participación: la de la persona que realiza el aporte escribiendo o editando parte del contenido de una página o grupos de páginas y la actividad del resto de las personas que valida la modificación. En caso de no estar de acuerdo, realiza una edición la cual es validada. Por lo tanto, la información presente en la wiki es aprobada por la comunidad ya sea por estar de acuerdo y no editarla o estando en desacuerdo y realizando el cambio necesario para que el contenido se alinee a su punto de vista.

Por lo tanto, así como en las wikis tradicionales, en SWOOKI no existe un espacio de creación y definición de anotaciones personales que posean los usuarios y que se diferencie del concepto de autoría compartida.

5.2. P-SWOOKI

En este trabajo de tesis, se realizó el desarrollo de P-SWOOKI, un sistema prototípico de construcción colaborativa de conocimiento el cual extiende la definición de SWOOKI añadiendo anotaciones semánticas personales.

Las anotaciones semánticas compartidas ya son provistas por SWOOKI como se detalló anteriormente, por lo tanto el trabajo sobre la extensión P-SWOOKI se centró en la forma de proveer anotaciones personales dentro del contexto de esta wiki semántica P2P, de modo tal que garantice las condiciones de privacidad del usuario, relación con los conceptos compartidos y navegabilidad del espacio de conocimiento.

Esta sección se describirán los detalles de implementación de P-SWOOKI, destacando la forma en que se manipularon los repositorios ontológicos, el guardado y la recuperación de información personal semántica y la diferenciación entre la información semántica personal y la información semántica compartida.

5.2.1. Funcionalidad

Como fue indicado, la funcionalidad principal de P-SWOOKI es permitir a los usuarios de una wiki semántica un espacio de creación y gestión de conocimiento personal. Este espacio debe permitirle acceso solamente al dueño de esas anotaciones y que las mismas puedan realizarse en función del conocimiento que fue gestado en el espacio de la comunidad, es decir en el espacio compartido.

La especificación del prototipo con estas características involucra distintos elementos funcionales: Una interfaz gráfica que integre la información gestada por la comunidad con la información personal, permitir navegar entre los espacios de información semántica personales y los espacios de información semántica compartidos. Por otro lado, analizar la configuración de la red P2P y por último, la forma en que se realiza la inyección de la definición ontológica y las diferentes formas de agregar y recuperar información personal semántica en la base de conocimiento (KB).

Extensión en la interfaz gráfica: del sistema compartido a la gestión de conocimiento personal

Uno de los elementos fundamentales para la realización de la extensión de P-SWOOKI fue la forma en que, visualmente, se pudieran complementar los dos espacios de información semántica en la misma wiki: el espacio personal y el espacio compartido.

Como inicio para la definición de la interfaz gráfica, el prototipo extiende la definida por SWOOKI que, en forma esquemática, podemos analizarla siguiendo la figura 5.4.

Visualmente, SWOOKI organiza la información del contenido en la parte central de la pantalla y en la parte superior y margen derecho se encuentra la información sobre la configuración de la wiki y las alternativas de navegación generales. Como indicamos, en la parte central de la pantalla podemos visualizar la información de contenido. En esta información nos basaremos en un primer análisis.

En primer lugar, cada página posee un título (en la figura 5.4 el número 1). El título indica el nombre del concepto para el cual se está desarrollando en el contenido de la página (3). El título es único para toda la wiki, esto significa que los conceptos no poseen diferentes páginas que los desarrollan y en caso que diferentes nodos inicien la edición en forma separada de un mismo concepto, en el momento en que se sincronicen el desarrollo volcado en cada uno de los nodos se

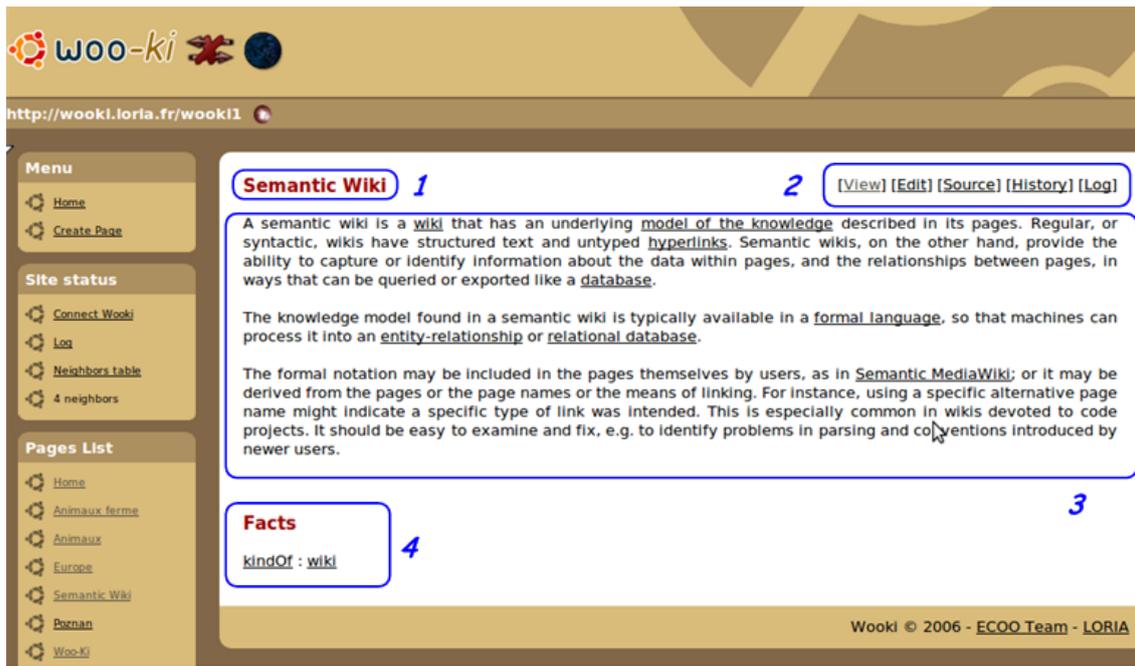


Figura 5.4: Estructura visual SWOOKI

complementará con la del otro nodo. Es importante tener presente el concepto sobre el cual se está desarrollando información, ya que la información semántica que se agregue a esa página estará expresada en función del concepto indicado por el título de la página wiki.

Siguiendo con el análisis, ubicada en la parte superior derecha se encuentra la barra de herramientas para la página que se visualiza. Las dos alternativas más utilizadas son *View* (Visualizar) y *Edit* (Editar). Visualizar es la opción predeterminada y permite visualizar la información de la página incluyendo texto con formato y links entre las páginas. Por otra parte, la opción de edición permite cambiar a una vista en la cual mediante un editor de texto y utilizando el lenguaje de markups se pueda editar la información de la página. En el modo edición es posible incluir la información semántica para la página wiki.

Ubicada en la parte central de la página (indicado con el número 3 en la figura 5.4) se visualiza el cuerpo de la página. En el modo Visualizar, esta parte muestra el texto del contenido con formato, links a páginas relacionadas y en la parte inferior un detalle de los hechos (*Facts*) que se cumplen para el concepto representado por la página. Los *facts* son las descripciones semánticas que fueron detalladas como typed links en el cuerpo de la wiki y fueron escritas en el modo edición. La particularidad de esta notación se encuentra en el hecho que en el cuerpo principal de la página no se visualiza el tipo del link, sino simplemente el link a otra página, en el ejemplo de la figura el *typed link* a la página “wiki” no se distingue de otros links .

El detalle de los *facts* permite al usuario distinguir inmediatamente la información semántica que se encuentra expresada en la descripción del concepto. Esta información semántica aparece dentro del espacio de la comunidad, es decir, que todos los usuarios pueden acceder para poder visualizarla o editarla.

El esquema visual presentado en SWOOKI define básicamente dos estadios para el gestionamiento de la información. En primera instancia la visualización de la información semántica (y sin características semánticas también) están presentes en la opción *visualize*. Por otro lado, cualquier modificación que se desea realizar a la información, aun cuando la información nueva sea estrictamente semántica como puede ser agregar tipo a los links, se realiza dentro del modo edición. Consumir, visualizar y navegar son netamente definidos en el espacio de visualización, mientras que agregar, eliminar o editar información queda destinada al espacio de edición.

Analizando los diferentes modos de visualización para cada una de las actividades que presenta SWOOKI, podemos destacar que la separación genera en el usuario un nivel de reelección y análisis diferente en cada una de las etapas de navegación o edición de la wiki. Cuando el usuario intenta modificar información, visualmente el editor presentado fuerza el cambio de contexto. En esta etapa el usuario es consciente que los cambios que realizará impactarán en el resto de la comunidad. Se puede decir que este es un cambio fuerte, ya que indirectamente necesita de la aprobación del resto de la comunidad.

Agregado de anotaciones personales: Categorías e Individuos

De la misma manera en que se gerencia la información semántica compartida, distinguiendo dos espacios de trabajo bien diferenciados, la gestión de información semántica personal posee dos espacios diferenciados. Por un lado, el contexto de visualización y por otro el contexto de agregado o edición del contenido semántico personal.

P-SWOOKI permite a cada usuario, además de interactuar con la información compartida provista por SWOOKI, incorporar información semántica dentro de un espacio personal a cualquiera de los conceptos que son parte del espacio compartido. En este prototipo, la forma en que un usuario agrega información semántica personal es categorizando los conceptos de la wiki. De cada concepto el usuario puede definirlo como un elemento que es parte de una categoría o que el concepto en si mismo define y describe una categoría.

Cada página de la wiki semántica puede ser anotada con varias anotaciones semánticas personales. Cada anotación semántica personal puede ser, en este caso,

una categoría o un individuo.

Las *Categorías* definen una familia de elementos. Por ejemplo, en la figura 5.5, la anotación “Ciudad” es utilizada para indicar que esa página wiki es una definición de la categoría *Ciudad*.

Por su parte, *Individuos* denotan elementos que pertenecen a lo sumo en una categoría. *Semantic Media Wiki* es un individuo que pertenece a la categoría *Semantic Wiki*. Un individuo puede pertenecer a varias categorías.

Una página wiki puede ser anotada con varias anotaciones personales. Por ejemplo, un usuario puede anotar la página wiki *SWOOKI* como individuo de *Semantic Wiki* y *Aplicación P2P* simultáneamente.

Actualmente el modelo de anotaciones de P-SWOOKI, solamente considera categorías e individuos. Sin embargo, el modelo de anotaciones personales está pensado para soportar relaciones y atributos en un futuro próximo.

La incorporación de esta funcionalidad en el prototipo continúa la filosofía presente en SWOOKI, es decir, el usuario posee dos etapas para la interacción con la información personal: la visualización y el agregado de las misma. Esta alternativa generó la necesidad de incluir un menú de edición de anotaciones personales y la extensión de la componente visual de SWOOKI para exhibir las anotaciones personales que se realizaron para ese concepto. La implementación del modo vista y agregado de anotaciones semánticas personales para la extensión básica puede visualizarse en las figuras 5.5 y 5.6 respectivamente.



Figura 5.5: Visualización en P-SWOOKI incluyendo la información semántica personal: el ejemplo muestra que para este usuario particular la página Paris es considerada un individuo (en inglés Individual) de las categorías Ciudad y Visitadas.

Para visualizar las anotaciones semánticas personales, se extendió el modo View de SWOOKI. La extensión incluye un recuadro el cual denominamos *Category Box*. El *Category Box* contiene todas las anotaciones personales que realizó el usuario para esa página distinguiendo las anotaciones referidas a individuos y las anotaciones referidas a categorías. Siguiendo el ejemplo de la figura 5.5, el usuario categorizó a

la página Paris como individuo (en la imagen Individuals) de las categorías Ciudad y Visitadas. Es decir, que para este usuario el concepto Paris es un elemento del conjunto representado por la categoría Ciudad y también es un elemento del conjunto representado por la categoría Visitadas. Podríamos suponer que Paris es considerada (para este usuario) en su espacio personal una Ciudad y también es un lugar que ha visitado en algún momento. En este ejemplo, el usuario ha considerado que Paris no describe a una Categoría.

Además de permitirle al usuario visualizar todas aquellas anotaciones semánticas personales que realizó, P-Swooki agrega una nueva utilidad en la barra de herramientas de edición para la página wiki actual. En el extremo superior derecho, a la izquierda de la opción View, ahora se encuentra la nueva opción Tag, la cual permite acceder al modo Edición del espacio personal. Cuando el usuario desea incorporar nueva información semántica personal a la página que esta visualizando, la opción Tag le permite acceder a la funcionalidad indicar que la página actual representa un individuo o una categoría como muestra la figura 5.6.



Tag Page : Paris

Tag as Individual which belongs to:

Tag as Category:

Figura 5.6: Agregado de información personal en P-SWOOKI

Dentro de la opción de categorizar la página el sistema le brinda al usuario la opción de etiquetar la página individuo perteneciente a una Categoría o como la definición de una Categoría. En el ejemplo de la figura 5.6, el usuario intenta etiquetar la página Paris como un individuo que pertenece a la categoría Ciudad.

Es importante destacar como es el proceso de agregado de categorías. En primer lugar si la categoría a la que se le intenta agregar un individuo no existe, primero se crea la categoría y luego se agrega el individuo. Caso contrario, se agrega directamente el individuo como miembro de esa Categoría.

El caso es similar al anotar una página como descripción de una categoría particular. En caso que la categoría no exista, primero se crea y luego se realiza la anotación de la página con la categoría indicada.

Para cada página, los usuarios pueden realizar todas las categorizaciones que deseen. Cada una de estas categorizaciones serán accesibles y visibles solamente por la persona que las creo. Y las mismas no serán propagadas al resto de los usuarios.

5.3. Navegabilidad en el espacio personal

El cuadro de anotaciones personales, además de indicarle al usuario aquellas categorizaciones que realizó con anterioridad, le permite navegar por las mismas.

Cada uno de los elementos de este cuadro es un link que lleva a una página donde se visualizan todos los conceptos relacionados con esta la categoría. Si se realiza un clic sobre un Individuo, se navega a la página de la categoría de ese individuo. Si en cambio, se hace click sobre una Categoría, se navega a la pagina de esa categoría.

De esta forma, se pueden navegar el espacio compartido utilizando navegación basada en conceptos. En la figura 5.7 se visualiza a la categoría Ciudad con dos links: uno a la página Paris y otro a la página Córdoba.



Figura 5.7: Navegación Personalizada

5.4. Implementación

En esta sección, se describirán las principales características referidas a la implementación del prototipo. En primer lugar una descripción de la ontología básica para realizar los repositorios de conocimiento utilizados, luego la definición de la ontología básica para realizar las anotaciones personales y la forma en que se agregan y recupera la información personal semántica diferenciándose de la compartida.

5.4.1. Modelo de Datos

El modelo de datos que utiliza P-Swooki es una extensión del modelo de datos de Swooki. Por lo tanto, cada wiki semántica (nodo) posee un identificador global único. Este identificador, hace que todos los repositorios pueden distinguirse.

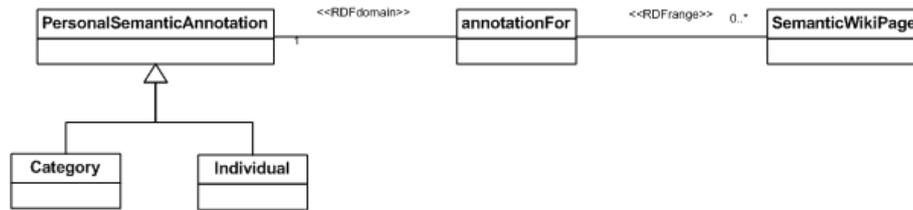


Figura 5.8: Anotaciones Semánticas Personales: Modelo de Datos

Como en cualquier sistema wiki, el elemento básico es una página wiki, por lo tanto cada página wiki posee un identificador único llamado PageID, el cual es el nombre de la página. El nombre de la página es asignado cuando la página se crea. Si ocurre que diferentes servidores crean, concurrentemente, páginas con el mismo nombre el contenido de las mismas se sincronizará utilizando el algoritmo de sincronización.

En la figura 5.8 se describe el modelo de datos de las anotaciones semánticas personales. Este modelo de datos se encuentra descripto utilizando ODM (Ontology Definition Meta-model) [GDDS06]. Con un nivel de detalle mayor, se incluye en la sección 7.2 del Anexo la definición del modelo de datos en formato OWL/RDF.

5.4.2. Definición de repositorios

Como se mencionó en la sección 5.1.1, SWOOKI utiliza Sesame como repositorio RDF embebido en cada uno de los nodos de la red P2P. Veamos un poco más en detalle las alternativas que presenta Sesame.

Sesame

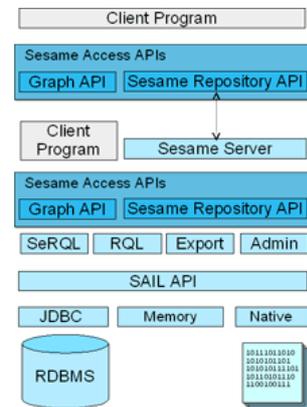
Sesame es un framework Java de código abierto que permite almacenar, consultar y realizar inferencias utilizando RDF y RDF Schema. Puede ser utilizado como una base de datos para RDF o RDF Schema, o simplemente como una biblioteca Java para aplicaciones que necesitan trabajar con RDF internamente. Este último caso es el que presenta SWOOKI y P-SWOOKI. Además Sesame provee herramientas para parsear, interpretar, consultar y almacenar información RDF.

Un concepto central en Sesame es el repositorio. Un repositorio es un contenedor de almacenamiento para RDF. Esto puede significar un objeto o conjuntos de objetos Java en memoria o una base de datos relacional. Cualquiera sea el tipo de almacenamiento elegido, es importante destacar que todas las operaciones en Sesame se producen con respecto a un repositorio. Esto significa que cada vez que se agrega información RDF a Sesame, esta información se agrega en un repositorio.

Cuando se realiza una consulta, la consulta se realiza sobre un repositorio particular.

Por otro lado, Sesame soporta inferencia a nivel de RDF Schema. Esto significa que dado un conjunto de datos RDF y/o un RDF Schema, Sesame puede encontrar la información subyacente en los datos.

Sesame posee una arquitectura basada en capas (figura 5.9) donde el acceso básico a la funcionalidad queda dado por la API de Acceso (Sesame Access API). La API de acceso provee acceso directo a los módulos funcionales de Sesame, esto es ya sea al programa del cliente como es el caso de SWOOKI y P-SWOOKI que utilizan Sesame como biblioteca interna, o directamente al siguiente componente de la arquitectura que es el Server Sesame.



La instalación de Sesame en modo embebido en el cliente, requiere simplemente incorporar a la aplicación un conjunto de bibliotecas java. Estas bibliotecas se descomponen en:

Figura 5.9: Arquitectura de Sesame

sesame.jar Las clases básicas Sesame.

rio.jar RIO (RDF I/O) es un conjunto de parsers y writers para diferentes formatos de serialización RDF (RDF/XML, Turtle, N-Triples).

openrdf-model.jar Interfaces y clases compartidas para el modelo RDF.

openrdf-util.jar Clases utilitarias.

Luego, la interacción se realiza mediante el uso de la API básica de comunicación dividida en una API de acceso al grafo de la ontología y otra para la manipulación de repositorios. Las operaciones básicas que se pueden desarrollar incluyen creación de repositorios, agregado de información y recuperación e inferencia de información en un repositorio particular. Como fue indicado, todas las operaciones que se realizan en Sesame se hacen sobre un repositorio particular.

La API para el manejo de repositorios es el punto de acceso central a los repositorios en Sesame. Puede ser utilizada para consultar y actualizar el contenido de repositorios remotos o locales. La API para repositorios considera todos los detalles de comunicación cliente-servidor, permitiendo de esta forma la misma complejidad de uso para repositorios locales como remotos. Las funcionalidades

principales incluyen acceso al repositorio, realizar consultas sobre un repositorio en particular y agregar información RDF a un repositorio.

La API para manipular Grafos provee una representación de un grafo RDF modelado en un objeto Java. La interfaz principal de la API es la clase `org.openrdf.model.Graph`. El propósito de esta clase es ofrecer un manejo conveniente de grafos RDF desde código. Los grafos pueden construirse programáticamente agregando sentencias al mismo o por medio de una evaluación de una consulta particular (SeRQL-construct) en un repositorio Sesame.

Para finalizar, Sesame incluye un lenguaje de consultas RDF/RDFS llamado SeRQL ("Sesame RDF Query Language", se pronuncia *circle*) . SeRQL combina las funcionalidades de los lenguajes de consulta RQL, RDQL, N-Triples y N3 agregándoles algunas características que permitan consultas más flexibles. Las características mas importantes de SeRQL son: transformación de grafos, soporte de esquemas RDF (RDF Schema), y soporte para XML Schema datatype,

Repositorios en P-Swooki

SWOOKI y P-SWOOKI utilizan un repositorio Sesame embebido dentro de cada nodo. En este repositorio se almacenan ternas RDF que describen los conceptos semánticos compartidos y privados. En P-Swooki, las anotaciones personales son almacenadas localmente. Cuando un usuario actualiza sus anotaciones personales se inicia la ejecución de operaciones de almacenamiento. Estas operaciones son ejecutadas localmente en el repositorio de conocimiento local bajo un espacio de nombres diferente al de las anotaciones compartidas. Esto genera una idea virtual de dos repositorios, uno para las anotaciones personales y otro para las compartidas por la comunidad.

Bajo la idea de dos repositorios, en P-Swooki cada nodo posee dos repositorios RDF locales: Sentencias personales (Personal Statements) y sentencias compartidas (Shared Statements). Estos representan al repositorio para almacenar entendimiento personal y el repositorio para almacenar entendimiento compartido.

El repositorio con sentencias compartidas (Shared Statements) contiene un conjunto de sentencias RDF que fueron extraídas de las páginas wiki. Una sentencia queda definida como una terna (Sujeto, Predicado, Objeto) donde el sujeto es el nombre de la página wiki, el predicado el tipo del link o nombre del *fact* y el objeto el concepto que está involucrado por medio del predicado. Esto es, el nombre de una página wiki.

El repositorio de sentencias personales (Personal Statements) contiene anota-

ciones semánticas personales las cuales son representadas, también, bajo sentencias RDF. Una sentencia personal RDF esta definida por una terna (Sujeto, Predicado, Objeto) para la cual el sujeto es una página wiki y el predicado es la anotación que se realiza bajo un determinado tipo. En este caso se realiza la anotación de una pagina como un individuo o una categoría.

Conceptualmente cada repositorio define dos operaciones básicas:

insertRDF(R,t) Agrega una sentencia t al repositorio de sentencias personales o compartidas R.

deleteRDF(R,t) Elimina una sentencia t del repositorio de sentencias personales o compartidas R.

Estas operaciones no son manipuladas directamente por el usuario final, sino que son llamadas implícitamente a partir de las acciones de edición que se realizan sobre la wiki, como fue comentado anteriormente.

Operaciones para Manipular Anotaciones Semánticas Personales

Existen cuatro operaciones de edición para editar anotaciones semánticas personales: `addIndividual` que agrega un Individuo, `addCategory` que agrega una categoría, `delIndividual` para eliminar un individuo y `delCategory` para eliminar una categoría. Las operaciones de actualización se consideraron como una doble operación de eliminación del valor viejo seguido del agregado del nuevo valor.

1. **addCategory(PageID, CategoryName):** donde PageID es el identificador que posee cada página wiki y CategoryName es el nombre de la nueva categoría.

Esta operación setea a la página wiki PageID como una categoría en el repositorio personal. Esta operación invoca al `insertRDF(Personal Statements,(PageId, RDF.Type, CategoryName))` para agregar una nueva terna en repositorio de información personal.

2. **addIndividual(PageID, CategoryName):** setea la página wiki PageID como miembro de la categoría CategoryName. Si CategoryName no existe, esta es agregada automáticamente al repositorio personal y luego automáticamente anota la página PageID como miembro de la categoría CategoryName.

Durante esta operación una sentencia RDF es agregada al repositorio personal por medio del llamado a `insertRDF(Personal Statements, (PageId, belongsTo, CategoryName))` donde *belongsTo* es un predicado para asociar un individuo a una categoría.

3. **delIndividual(PageID, CategoryName)**: elimina a PageID como miembro de la categoría CategoryName del repositorio personal RDF llamando a DeleteRDF(Personal Statements, (PageId, belongsTo, CategoryName)).
4. **delCategory(PageID, CategoryName)**: primero elimina todos los individuos de esa categoría y luego elimina la categoría CategoryName del repositorio RDF personal.

Implementación: agregado de una categoría

A continuación se detallará en un breve ejemplo, la forma en que se implementó la operación addIndividual(PageID, CategoryName). Como ya dijimos antes, SWOOKI y por ende P-Swooki también, están implementadas utilizando lenguaje Java en una arquitectura basada en servlets.

Cuando un usuario decide anotar una pagina wiki como individuo de una categoría, en base a una implementación MVC, se dispara una acción desde la vista al modelo solicitando la ejecución de la operación addIndividual(PageID, CategoryName) con la semántica que se describió antes.

La clase encargada de implementar el método que interactúa directamente con la API de Sesame posee el nombre SesameEngine. Dentro de esta clase se puede ver la implementación del método que se detalla en el listado a continuación.

```

1 public void tagPageAsCategory(String pageId, String categoryName) {
2     ValueFactory f = this.getSesameOp().getSesameRepository().
3         getValueFactory();
4     URI wikiPageURI = f.createURI(ILIFIAConstants.WIKI_PAGE_CLASS);
5     URI categoryURI = f.createURI(ILIFIAConstants.CATEGORY_CLASS);
6     URI annotationForURI = f.createURI(ILIFIAConstants.
7         ANNOTATION_FOR);
8     URI pageIdDataProperty = f.createURI(ILIFIAConstants.
9         PAGE_ID_DATA_PROPERTY);
10    URI pageIdURI = f.createURI(this.serverUrl+pageId);
11    URI categoryNameURI = f.createURI(ILIFIAConstants.
12        LIFIA_ONTOLLOGY+categoryName);
13    URI pageURI = f.createURI(ILIFIAConstants.LIFIA_WIKI_PAGE_CODE+
14        pageId);
15    RepositoryConnection c;
16    try {
17        c = this.getSesameOp().getSesameRepository().
18            getConnection();
19        //page is a WikiPage
20        c.add(pageURI, RDF.TYPE, wikiPageURI);

```

```

15         c.add(pageURI, pageIdDataProperty, pageIdURI);
16         //categoryNameURI is a category
17         c.add(categoryNameURI, RDF.TYPE, categoryURI);
18         //CategoryName is the the annotation for page
19         c.add(categoryNameURI, annotationForURI, pageURI);
20         c.close();
21     } catch (RepositoryException e) {
22         // TODO Auto-generated catch block
23         e.printStackTrace();
24     } finally{
25     }
26 }

```

Listing 5.1: Implementación: taggear una página wiki como Categoría

Desde las líneas 3 a la 9 se declaran las URIs que identifican a cada uno de los elementos que se van a utilizar para formar las ternas RDF. Por ejemplo, en la línea número 4 se declara una variable `categoryURI` la cual va a tener asignada una URI que identifica a la clase `Category` en nuestro modelo de datos. Luego de obtener todos los identificadores se procede a agregar las ternas en el repositorio.

En la línea 13 se abre una conexión al repositorio local donde se agregarán las sentencias RDF. En la línea 14 se indica que la página recibida como parámetro es una página wiki, ya que las anotaciones personales se realizan sobre páginas wikis, y además se le setea un identificador interno como propiedad. Luego, en la línea siguiente 18 se crea la categoría. En caso de que la categoría exista la línea no tiene efecto. Por último, en la línea 20 se realiza la asignación de la categorización de la página wiki con la categoría de la que se recibió el nombre.

Implementación: Recuperación de categorías de una página

Otro aspecto asociado al agregado de información personal semántica a las páginas de la wiki, es la posterior recuperación de dicha información. Esta operación se ve reflejada cada vez que el usuario accede a una página de la wiki y en el modo vista puede apreciar el cuadro de información personal donde se detallan las anotaciones previas que se hicieron sobre la página wiki (Figura 5.5).

En la misma clase donde se encuentra el método para agregar la información para categorizar una página wiki como Categoría, se encuentra el método que permite recuperar la información relacionada a esa página wiki. En el listado siguiente se detalla la implementación.

```

1 public Object getCategoryTagAbout(String page, String serverURL) {

```

```

2      String query = "SELECT ?tag WHERE { ?tag <<" + RDF.TYPE.
          toString() + ">>" + ILIFIAConstants.CATEGORY_CLASS + ">." + " ?
          tag <<" + ILIFIAConstants.ANNOTATION_FOR + ">>" + page + ">." + " }";
3      String result = "";
4      try {
5          result = this.execTagQuery(query, serverURL);
6      } catch (Exception e) {
7          // TODO Auto-generated catch block
8          e.printStackTrace();
9      }
10     return result;
11 }

```

Listing 5.2: Implementación: Obtener las categorías con que fue anotada la página

En la línea 2 se define dinámicamente el texto de una consulta SPARQL, compatible con Sesame. La consulta puede comprenderse como muestra el listado siguiente.

Al momento de ejecutar la consulta, la variable *page* tendrá como valor asignado el nombre de la página wiki. Por lo tanto, la única variable en esta consulta es *?tag*. De esta forma, la consulta obtendría todos aquellos elementos en el repositorio que satisfacen ser del tipo *CATEGORY* y que además satisfacen ser anotación para la página *page*. En otras palabras, se obtienen todas las categorías de la página *page*.

El prefijo en la consulta *LIFIAConstants*, permite obtener el namespace completo para la definición de *CATEGORY* y *ANNOTATION_FOR*. De esta forma, se obtienen solamente los elementos del repositorio virtual personal.

```

1 SELECT ?tag WHERE
2 { ?tag TYPE CATEGORY .
3   ?tag ANNOTATION_FOR page. }

```

Listing 5.3: Consulta SPARQL para obtener las categorías de una página

5.4.3. Arquitectura P-Swooki

Luego de haber detallado la forma en que se realizó la extensión de *SWOOKI*, podemos definir la arquitectura resultante que resume la nueva funcionalidad y las nuevas componentes.

Las componentes que han sido modificadas fueron la componente visual donde la nueva funcionalidad para agregar anotaciones personales fue agregada y por otro lado la componente de interacción con el repositorio Sesame. En la figura 5.10 podemos

distinguir con fondo gris las componentes originales de Swooki y con fondo blanco los agregados propios de P-Swooki.

Otro cambio importante es la configuración en la red. Para utilizar P-Swooki, debe configurarse un nodo por usuario. Los nodos pueden estar ubicados en una granja de servidores o las computadoras personales de cada usuario. Luego de esto, la utilización y configuración e los nodos vecinos se realiza de la misma forma que para Swooki.

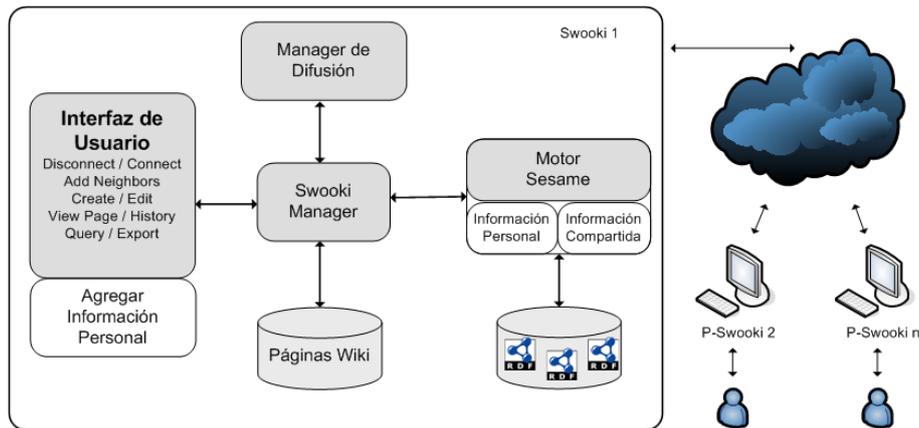


Figura 5.10: Arquitectura de P-SWOOKI

5.5. Conclusiones del Capítulo

En este capítulo se describió la implementación del prototipo P-SWOOKI. P-SWOOKI incorpora anotaciones semánticas personales a una wiki semántica.

Se describieron detalles de la implementación y las funcionalidades para realizar anotaciones semánticas, poder visualizarlas y navegar.

Se dieron detalles técnicos sobre la implementación, la forma de manipular las ontologías y como se interactuó con el repositorio RDF Sesame de forma de poder tener dos espacios de representación de conocimiento separados.

Por último, se describió la arquitectura resultante en función de los cambios producidos a la wiki original.

Capítulo 6

Evaluación

En esta etapa, presentare una evaluación acerca del enfoque propuesto en P-Swooki. La experiencia fue realizada en dos etapas separadas, la primera realizada en Francia ,durante una estadía de investigación en la cual se realizó gran parte de la extensión P-Swooki, y la segunda en Argentina¹.

El total de participantes fue de 15 personas. Las edades de los participantes estaban comprendidas entre 25 y 45 años. Tods los participantes estaban relacionados con la informática, todos estaban familiarizados con wikis y 5 de ellos estaban, además, familiarizados con wikis semánticas y poseian experiencia en construcción de ontologías.

Al momento de realizar la prueba, los participantes fueron ubicados en habitaciones diferentes y no tenian permitido comunicarse entre ellos durante la experiencia.

Iniciamos la experiencia en Francia con una breve explicación sobre wikis semánticas, conocimiento compartido y conocimiento personal. Pedimos a los participantes que construyan, en base a la breve explicación, el contenido de una wiki semántica utilizando los tipos de notaciones (compartidas y personales).

La actividad se inicio utilizando la wiki con dos páginas basicas para que actuen como disparador de la temática sobre el contenido de la wiki. Para esto, creamos dos páginas semánticas; una sobre Semantic Wiki y otra sobre Semantic Web. Además les sugerimos a los participantes que utilicen un sintáxis básica en la definición de las anotaciones para tratar de controlar la explosión de vocabulario como ocurre en folksonomies [SLR⁺06].

Para poder consolidar la primer experiencia realizada en Francia, repetimos la misma en Argentina en la cual confirmamos los resultados que habiamos

¹Esta experiencia se encuentra detallada en el artículo [TSMDM09c]

obtenido previamente. Estas experiencias mostraron una evidencia preliminar de la contribución realizada con respecto a la usabilidad de anotaciones semánticas personales y como se complementan con el conocimiento compartido. A continuación presentaré los resultados obtenidos en estas experiencias. Como ambas experiencias brindaron los mismos resultados, los detalles serán basados en la primer experiencia realizada en Francia.

Los cuadros 6.1 y 6.2 muestran el tipo (Individuo o Categoría) y la cantidad de anotaciones semánticas personales que cada participante agregó a las páginas wiki Semantic Wiki y Semantic Web respectivamente.

La página *Semantic Wiki* fue anotada por todos los participantes. Ellos anotaron esta página como Individuo 17 veces y como categoría 15 veces. El participante más activo agregó 11 anotaciones semánticas personales a esta página. El promedio de anotaciones por participante fue de 4,5. Y el porcentaje de anotaciones sin tener en cuenta al participante más activo fue de 3,5 anotaciones.

La página wiki *Semantic Web* también fue anotada por todos los participantes. Han anotado esta página como Individuo 8 veces y como Categoría 9. El participante más activo realizó 11 anotaciones personales sobre esta página. El promedio del número de anotaciones por participante fue de 2.5. El promedio, sin tener en cuenta al participante más activo, fue cercano a 1.

Usuario	Individuos	Categorías	Total
1	1	1	2
2	1	0	1
3	2	5	7
4	6	5	11
5	2	1	3
6	4	3	7
7	1	0	1

Cuadro 6.1: Anotaciones Semánticas Personales para la página *Semantic Wiki*

Los resultados confirmaron nuestra hypothesis sobre la utilidad de las anotaciones semánticas personales, ya que todos los participantes han agregado anotaciones personales en las páginas que utilizaron. El cuadro 6.3 muestra las anotaciones semánticas personales y compartidas utilizadas por los participantes en la página wiki *Semantic Wiki*.

La columna Anotaciones Compartidas agrupa las anotaciones semánticas compartidas. Al comienzo de la experiencia, esta fue vacía. Esta página podría ser an-

Usuario	Individuos	Categorías	Total
1	0	1	1
2	1	0	1
3	0	2	2
4	6	5	11
5	1	1	2
6	1	0	1
7	1	0	1

Cuadro 6.2: Anotaciones Semánticas Personales para la página *Semantic Web*

Anotaciones Compartidas	Individuos	Categorías
Category: PersonalInformationManagement	ResearchTopic	SemanticWiki (4)
Category: KnowledgeManagement	SemanticWeb (4)	Wiki(2)
Category: SemanticWeb	CollaborativeTool(2)	SemanticWeb
Category:FormalLanguage	Web	Web (2)
KindOf: Wiki	NoDelete	CollaborativeTool
has: FactBox	Semantics	WebOfData
limitation : fault-tolerance	KnowledgeWeb	NoUndoTag
limitation : scalability	Something	WWW
limitation : censorship	Wiki (3)	CSCW
	Web	Semantic

Cuadro 6.3: Anotaciones Semánticas Personales y Compartidas para la página wiki *Semantic Wiki*

otada como una categoría, por ejemplo `Category:PersonalInformationManagement` o como una propiedad como puede ser `limitation:scalability`. La segunda y tercer columna agrupan todas las anotaciones semánticas personales. Note, que en algunos casos varios usuarios usaron las mismas anotaciones semánticas. Por ejemplo, cuatro usuarios usaron la anotación *SemanticWeb* y dos utilizaron *Wiki*.

Podemos observar que el número total de anotaciones semánticas se incrementó. Por lo tanto, las anotaciones semánticas personales pueden ser útiles para aumentar el conocimiento compartido.

Además, esta evaluación nos dejó algunas lecciones:

- Cada participante utilizó ambas anotaciones semánticas: personales y compartidas.
- Varios participantes dijeron que era fácil utilizar las anotaciones semánticas personales ya que no era necesario embeberlas dentro del texto
- Algunos participantes tuvieron dificultades para distinguir entre una Categoría y un Individuo.
- Todos los participantes manifestaron la importancia de tener una mejor interfáz de usuario para facilitar navegación personal.
- Para algunos participantes las anotaciones semánticas personales fueron útiles para estructurar su propio mapa navegacional acorde a sus propias taxonomías.
- Las anotaciones personales fueron más fáciles de utilizar para aquellas personas que no estaban familiarizadas con las wikis semánticas, mientras que aquellas familiarizadas con las wikis semánticas no llegaron a ver exactamente que valor agregado les producía utilizar anotaciones semánticas personales.
- Un participante no comprendió la diferencia entre anotaciones semánticas personales y anotaciones semánticas compartidas.
- Para muchos participantes fue más sencillo agregar anotaciones personales cuando estaban browseando la wiki y agregar anotaciones compartidas cuando estaban editando las páginas wikis.
- Muchos participantes encontraron que combinar ambos tipos de anotaciones les ayudaría en tener una mejor recuperación del conocimiento.

Aunque la experiencia es pequeña, el porcentaje de anotaciones personales mostro una tendencia: las personas se sintieron a gusto al utilizar anotaciones

personales y el hecho de agregar información semántica personal lo consideraron una actividad complementaria en una wiki semántica.

Estos resultados preliminares nos dan una visión positiva para seguir investigando en esta dirección, sin embargo, es clara la necesidad de realizar experiencias a mayor escala para consolidar estos resultados.

Capítulo 7

Conclusiones y Trabajo a Futuro

En el trabajo se describió el contexto de la Web y la relación con el conocimiento. En primer lugar se dio una general de las tecnologías y desafíos que involucra la Web Semántica y luego se comenzó a tratar la los procesos involucrados a la construcción colaborativa de conocimiento.

La construcción colaborativa de conocimiento plantea un ciclo de actividades sociales y cognitivas complejo al cual se busca brindarle un sustento informático. Este sustento debe estar acorde a las nuevas prestaciones tecnológicas y de participación de los usuarios. Durante el análisis a las formas en que la informática ofrece alternativas viables a estos procesos sociales complejos se detectó una ausencia en la representación de espacios personales para la construcción de conocimiento en particular en el contexto de las wikis semánticas.

Se propuso la definición de un espacio personal en wikis semánticas mediante la incorporación de anotaciones semánticas personales. Estas anotaciones quedan definidas en término de Categorizaciones de los elementos del espacio compartido utilizando la idea de folksonomies. Las anotaciones se realizan sobre un espacio privado al que solamente puede acceder el dueño del espacio, pero permite hacer referencia a elementos del espacio compartido.

Se realizó un prototipo que extiende una wiki semántica P2P a la que se le incluyó la definición para realizar anotaciones personales. Por medio de estas anotaciones, los usuarios pueden utilizar espacios privados en los que pueden combinar conocimiento personal. La información personal es accedida únicamente por los usuarios, mientras que la información semántica compartida se propaga al resto de los usuarios. Para ello, la información semántica se dividió en dos espacios de nombres diferentes.

Además, la navegabilidad en términos de información personal mejoró y

extendió la que presentaba el modelo compartido. De esta manera se obtuvo navegación personalizada, navegación personalizada basada en conceptos y se extendió el espacio navegacional compartido.

Finalizando, las experiencias demuestran que la utilización de las anotaciones personales no generan rechazo en los usuarios y la gran mayoría las utilizó libremente. Para destacar, varios de las personas que formaron parte de la experiencia indicaron que se sentían más cómodas utilizando las anotaciones personales en la etapa de navegación y consulta, mientras que las anotaciones compartidas las utilizaron en momentos de edición estructural del concepto.

7.1. Trabajo a Futuro

Como perspectivas a futuros debemos continuar con la definición de las anotaciones semánticas personales de forma que puedan ser más expresivas, permitiendo describir relaciones y atributos. De esta forma, estaremos en condiciones de analizar la publicación y la internalización basada en información con el mismo nivel de expresividad.

Otro plano para continuar con las líneas de investigación esta ligado a analizar las anotaciones personales de diferentes usuarios a fin de encontrar conceptualizaciones comunes basadas en las categorizaciones. Existen líneas de investigación relacionadas al tema, las cuales involucran técnicas como data mining y FCA.

Para finalizar, este trabajo de tesis es producto de trabajos de investigación realizados en LIFIA en colaboración con el equipo ECOO del Laboratorio LORIA, Nancy, Francia los cuales significaron la publicación de dos artículos[TSMDM09c, TSMDM09b] y un poster[TSMDM09a].

Anexo

7.2. Definición semántica RDF/OWL

```
1
2 <?xml version=" 1.0" ?>
3 <!DOCTYPE rdf:RDF [
4     <!ENTITY owl " http://www.w3.org/2002/07/owl#" >
5     <!ENTITY xsd " http://www.w3.org/2001/XMLSchema#" >
6     <!ENTITY owl2xml " http://www.w3.org/2006/12/owl2-xml#" >
7     <!ENTITY rdfs " http://www.w3.org/2000/01/rdf-schema#" >
8     <!ENTITY rdf " http://www.w3.org/1999/02/22-rdf-syntax-ns#" >
9     <!ENTITY FolksoWiki " http://www.lifia.info.unlp.edu.ar/ontologies
10     /2008/10/FolksoWiki.owl#" >
11 ]>
12 <rdf:RDF xmlns=" http://www.lifia.info.unlp.edu.ar/ontologies/2008/10/
13 FolksoWiki.owl#"
14     xml:base=" http://www.lifia.info.unlp.edu.ar/ontologies/2008/10/
15     FolksoWiki.owl"
16     xmlns:rdfs=" http://www.w3.org/2000/01/rdf-schema#"
17     xmlns:owl2xml=" http://www.w3.org/2006/12/owl2-xml#"
18     xmlns:owl=" http://www.w3.org/2002/07/owl#"
19     xmlns:FolksoWiki=" http://www.lifia.info.unlp.edu.ar/ontologies
20     /2008/10/FolksoWiki.owl#"
21     xmlns:xsd=" http://www.w3.org/2001/XMLSchema#"
22     xmlns:rdf=" http://www.w3.org/1999/02/22-rdf-syntax-ns#">
23 <owl:Ontology rdf:about="" />
24
25 <!--
26 ////////////////
27 //
28 // Object Properties
29 //
30 ////////////////
31 -->
```

```

28     <!-- http://www.lifia.info.unlp.edu.ar/ontologies/2008/10/
        FolksoWiki.owl#annotatedBy -->
29     <owl:ObjectProperty rdf:about="#annotatedBy">
30         <owl:inverseOf rdf:resource="#annotationFor" />
31     </owl:ObjectProperty>
32     <!-- http://www.lifia.info.unlp.edu.ar/ontologies/2008/10/
        FolksoWiki.owl#annotationFor -->
33
34     <owl:ObjectProperty rdf:about="#annotationFor">
35         <rdfs:domain rdf:resource="#SemanticTag" />
36         <rdfs:range rdf:resource="#WikiElement" />
37     </owl:ObjectProperty>
38
39     <!-- http://www.lifia.info.unlp.edu.ar/ontologies/2008/10/
        FolksoWiki.owl#belongsTo -->
40     <owl:ObjectProperty rdf:about="#belongsTo">
41         <owl:inverseOf rdf:resource="#hasIndividual" />
42     </owl:ObjectProperty>
43     <!-- http://www.lifia.info.unlp.edu.ar/ontologies/2008/10/
        FolksoWiki.owl#hasIndividual -->
44     <owl:ObjectProperty rdf:about="#hasIndividual">
45         <rdfs:domain rdf:resource="#Category" />
46         <rdfs:range rdf:resource="#Individual" />
47     </owl:ObjectProperty>
48     <!--
49     ////////////////
50     //
51     // Data properties
52     //
53     ////////////////
54     -->
55     <!-- http://www.lifia.info.unlp.edu.ar/ontologies/2008/10/
        FolksoWiki.owl#pageId -->
56     <owl:DatatypeProperty rdf:about="#pageId" />
57     <!-- http://www.lifia.info.unlp.edu.ar/ontologies/2008/10/
        FolksoWiki.owl#tagName -->
58     <owl:DatatypeProperty rdf:about="#tagName">
59         <rdfs:domain>
60             <owl:Restriction>
61                 <owl:onProperty rdf:resource="#pageId" />
62                 <owl:someValuesFrom rdf:resource="&xsd:string" />
63             </owl:Restriction>
64         </rdfs:domain>
65     </owl:DatatypeProperty>
66     <!--

```

```

67 ////////////////
68 //
69 // Classes
70 //
71 ////////////////
72   →
73
74   <!-- http://www.lifia.info.unlp.edu.ar/ontologies/2008/10/
       FolksoWiki.owl#Category →
75   <owl:Class rdf:about="#Category">
76     <rdfs:subClassOf rdf:resource="#SemanticTag" />
77   </owl:Class>
78
79   <!-- http://www.lifia.info.unlp.edu.ar/ontologies/2008/10/FolksoWiki
       .owl#Individual →
80   <owl:Class rdf:about="#Individual">
81     <rdfs:subClassOf rdf:resource="#SemanticTag" />
82   </owl:Class>
83
84   <!-- http://www.lifia.info.unlp.edu.ar/ontologies/2008/10/
       FolksoWiki.owl#SemanticTag →
85   <owl:Class rdf:about="#SemanticTag">
86     <rdfs:subClassOf rdf:resource="&owl;Thing" />
87   </owl:Class>
88   <!-- http://www.lifia.info.unlp.edu.ar/ontologies/2008/10/
       FolksoWiki.owl#WikiElement →
89   <owl:Class rdf:about="#WikiElement">
90     <rdfs:subClassOf rdf:resource="&owl;Thing" />
91   </owl:Class>
92   <!-- http://www.lifia.info.unlp.edu.ar/ontologies/2008/10/
       FolksoWiki.owl#WikiPage →
93   <owl:Class rdf:about="#WikiPage">
94     <rdfs:subClassOf rdf:resource="#WikiElement" />
95   </owl:Class>
96   <!-- http://www.w3.org/2002/07/owl#Thing →
97   <owl:Class rdf:about="&owl;Thing" />
98 </rdf:RDF>
99
100 <!-- Generated by the OWL API (version 2.2.1.941) http://owlapi.
       sourceforge.net →
101 }

```

Listing 7.1: Consulta SPARQL para obtener las categorías de una página

Bibliografía

- [ADR06] Sören Auer, Sebastian Dietzold, and Thomas Riechert. Ontowiki - A tool for social, semantic collaboration. In Isabel F. Cruz, Stefan Decker, Dean Allemang, Chris Preist, Daniel Schwabe, Peter Mika, Michael Uschold, and Lora Aroyo, editors, *International Semantic Web Conference*, volume 4273 of *Lecture Notes in Computer Science*, pages 736–749. Springer, 2006.
- [Av08] Grigoris Antoniou and Frank van Harmelen. *A Semantic Web Primer, 2nd Edition*. The MIT Press, 2 edition, March 2008.
- [BLHL01] Tim Berners-Lee, James Hendler, and Ora Lassila. The semantic web (berners-lee et. al 2001). May 2001.
- [Bor97] W.N. Borst. *Construction of Engineering Ontologies*. PhD thesis, University of Twente, Enschede, The Netherlands, 1997.
- [Bru07] Peter Brusilovsky. Adaptive navigation support. In Peter Brusilovsky, Alfred Kobsa, and Wolfgang Nejdl, editors, *The Adaptive Web*, volume 4321 of *Lecture Notes in Computer Science*, chapter 8, pages 263–290. Springer-Verlag, Berlin, Germany, May 2007.
- [Car07] Jorge Cardoso. *Semantic Web Services: theory, tools and applications*. IGI Global, 2007.
- [DC04] Alicia Díaz and Gérome Canals. Divergence occurrences in knowledge sharing communities. In *CRIWG*, pages 17–24, 2004.
- [FS08] Thomas Franz and Sergej Sizov. Communication systems for semantic work environments. In Jörg Rech, Björn Decker, and Eric Ras, editors, *Emerging Technologies for Semantic Work Environments: Techniques, Methods, and Applications*, chapter 2, pages 16–32. IGI Global, USA, 2008.

- [GDDS06] Dragan Gaaevic, Dragan Djuric, Vladan Devedzic, and Bran Selic. *Model Driven Architecture and Ontology Development*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.
- [Gru93] Thomas R. Gruber. A translation approach to portable ontology specifications. *Knowledge Systems Laboratory, Technical Report, Computer Science Department, Stanford University, California, KSL 92-71*, 1993.
- [Heg67] G. W. F. Hegel. *Phenomenology of Spirit*. New York, NY: Harper & Row. (Original work published in 1807), 1967.
- [Hei] Martin Heidegger. *Being and Time*. Blackwell Publishing (original work published in 1927).
- [KVV06] Markus Krötzsch, Denny Vrandečić, and Max Völkel. *Semantic MediaWiki*. 2006.
- [mAYGS08] Ching man Au Yeung, Nicholas Gibbins, and Nigel Shadbolt. A study of user profile generation from folksonomies. In Peter Dolog, Markus Krötzsch, Sebastian Schaffert, and Denny Vrandečić, editors, *SWKM*, volume 356 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2008.
- [MOtORT07] John Musser, Tim O'Reilly, and the O'Reilly Radar Team. *Web 2.0 Principles and Best Practices*. O'Reilly Media Inc., 2007.
- [NT95] Ikujiro Nonaka and Hirotaka Takeuchi. *The Knowledge - Creating Company: How Japanese Companies Create the Dynamics of Innovation*. Oxford University Press, 1995.
- [OBD06] Eyal Oren, John G. Breslin, and Stefan Decker. How semantics make better wikis. In *Proceedings of the International World-Wide Web Conference*, pages 1071–1072, New York, NY, USA, 2006. ACM Press.
- [Ore05] Eyal Oren. Semperwiki: a semantic personal wiki. In *Proc. of 1st Workshop on The Semantic Desktop - Next Generation Personal Information Management and Collaboration Infrastructure, Galway, Ireland*, NOV 2005.
- [OVBD06] Eyal Oren, Max Völkel, John Breslin, and Stefan Decker. Semantic wikis for personal knowledge management. *Database and Expert Systems Applications*, pages 509–518, 2006.

- [OWL04] Owl web ontology language, 2004.
- [Pas04] Thomas B. Passin. *Explorer's Guide to the Semantic Web*. Manning, Greenwich, 2004.
- [RSMM08] Charbel Rahhal, Hala Skaf-Molli, and Pascal Molli. Swooki: A peer-to-peer semantic wiki. In *The 3rd Workshop: 'The Wiki Way of Semantics'-SemWiki, co-located with the 5th Annual European Semantic Web Conference (ESWC), Tenerife, Spain, June 2008*.
- [SBBK08] Sebastian Schaffert, François Bry, Joachim Baumeister, and Malte Kiesel. Semantic wikis. *IEEE Software*, 25(4):8–11, 2008.
- [Sch83] Donald A. Schön. *The reflective practitioner: how professionals think in action*. Basic Books, New York, 1983.
- [SLR⁺06] Shilad Sen, Shyong K. Lam, Al Mamunur Rashid, Dan Cosley, Dan Frankowski, Jeremy Osterhouse, F. Maxwell Harper, and John Riedl. tagging, communities, vocabulary, evolution. In *CSCW '06: Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*, pages 181–190, New York, NY, USA, 2006. ACM.
- [SSN⁺02] S. Staab, S. Santini, F. Näck, L. Steels, and A. Maedche. Emergent semantics. *Intelligent Systems, IEEE [see also IEEE Expert]*, 17(1):78–86, 2002.
- [Sta06] Gerry Stahl, editor. *Group cognition: Computer support for building collaborative knowledge*. Cambridge, MA: MIT Press, 2006.
- [ter98] Gerard Henri Hofte ter. *Working apart together: foundations for component groupware*. PhD thesis, Enschede, June 1998.
- [TSMDM09a] Diego Torres, Hala Skaf-Molli, Alicia Diaz, and Pascal Molli. Personal and shared knowledge building in p2p semantic wikis. In *Poster presented at 6th European Semantic Web Conference (ESWC 2009), Heraklion, Greece, May 2009*.
- [TSMDM09b] Diego Torres, Hala Skaf-Molli, Alicia Diaz, and Pascal Molli. Personal navigation in semantic wikis. In *International Workshop on Adaptation and Personalization for Web 2.0 in connection with UMAP'09, Trento, Italy, June 2009*.

-
- [TSMDM09c] Diego Torres, Hala Skaf-Molli, Alicia Díaz, and Pascal Molli. Supporting personal semantic annotations in p2p semantic wikis. In Sourav S. Bhowmick, Josef Küng, and Roland Wagner, editors, *DEXA*, volume 5690 of *Lecture Notes in Computer Science*, pages 317–331. Springer, 2009.
- [VH07] Gottfried Vossen and Stephan Hagemann. *Unleashing Web 2.0: From Concepts to Creativity*. Morgan Kaufmann, 2007.
- [VK06] Denny Vrandečić and Markus Krötzsch. Reusing ontological background knowledge in semantic wikis. In *SemWiki*, 2006.
- [Wal07] Thomas Vander Wal. Folksonomy, 2007. <http://vanderwal.net/folksonomy.html>.
- [WUM07] Stéphane Weiss, Pascal Urso, and Pascal Molli. Wooki: a p2p wiki-based collaborative writing tool. In *Web Information Systems Engineering*, Nancy, France, December 2007. Springer.