

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ

ESCUELA DE POSGRADO



**PUCP**

**ARQUITECTURA PARA SISTEMAS DE MEMORIA ORGANIZACIONAL EN  
INSTITUCIONES DE EDUCACIÓN SUPERIOR**

Tesis para optar el grado de Magíster en Informática con mención en  
Ingeniería de Software que presenta

ALEXANDER EDWIN QUILCA CONDORI

Dirigido por

DR. HÉCTOR ANDRÉS MELGAR SASIETA

San Miguel, 2014

# Resumen

El presente trabajo de investigación tiene como objetivo definir una arquitectura para sistemas de Memoria Organizacional en Instituciones de Educación Superior. El objetivo surge del problema identificado, en el cual la construcción de sistemas de gestión del conocimiento es tratada como construcción de sistemas típicos, aplicando métodos iterativos que si bien proporcionan versiones tempranas del sistema, no permiten predecir y gestionar adecuadamente el desarrollo del sistema. Además, de la revisión del estado del arte se identificó que se ha desarrollado pocos estudios de memoria organizacional en instituciones de educación superior pese a ser un lugar natural para la investigación y transferencia de conocimiento.

La revisión del estado del arte permitió validar que hay pocos estudios de memoria organizacional en instituciones de educación superior. Por lo cual, se tuvo que ampliar las opciones de búsqueda a otros dominios, con lo cual se identificó arquitecturas conceptuales de memoria organizacional que apoyen en la definición y validación de la arquitectura propuesta.

Para estructurar la arquitectura se usó el modelo CESM y de la metodología CommonKADS. Además, para apoyar el diseño de la arquitectura propuesta se revisaron las arquitecturas del estado del arte y se comparó la arquitectura propuesta con las arquitecturas de la revisión del estado del arte. La arquitectura propuesta presenta muchas similitudes con las arquitecturas identificadas en el estado del arte, ya que comparten muchos de sus elementos como los usuarios, herramientas para interactuar con el sistema, repositorio de documentos y estas arquitecturas hacen uso de ontologías.

La arquitectura propuesta para sistemas de memoria organizacional en instituciones de educación debe ser validada con la construcción de frameworks y prototipos en un departamento académico de una universidad. Posteriormente se podría evaluar la arquitectura propuesta en otros dominios de aplicación fuera del ámbito académico como salud, gobierno, industria, entre otros.

# Índice general

Índice de Figuras	III
Índice de Tablas	IV
<b>1 Presentación de la Investigación</b>	<b>1</b>
1.1 Introducción . . . . .	1
1.2 Definición del Problema . . . . .	2
1.3 Objetivo General . . . . .	3
1.4 Objetivos Específicos . . . . .	3
1.5 Justificación . . . . .	4
1.6 Métodos y procedimientos . . . . .	6
<b>2 Marco Conceptual</b>	<b>10</b>
2.1 Introducción . . . . .	10
2.2 Gestión del Conocimiento . . . . .	10
2.2.1 Conocimiento . . . . .	10
2.2.2 Gestión del Conocimiento . . . . .	11
2.2.3 Sistemas de Gestión del Conocimiento . . . . .	12
2.2.4 Procesos de la Gestión del Conocimiento . . . . .	12
2.3 Memoria Organizacional . . . . .	13
2.3.1 Modelos de Memoria Organizacional . . . . .	13
2.4 Arquitectura de Software . . . . .	16
2.5 Institución de Educación Superior . . . . .	18
<b>3 Revisión del Estado del Arte</b>	<b>19</b>
3.1 Introducción . . . . .	19
3.2 Fases de la Revisión Sistemática . . . . .	19
3.2.1 Primera Búsqueda . . . . .	19
3.2.2 Segunda Búsqueda . . . . .	21
3.2.3 Tercera Búsqueda . . . . .	23
3.2.4 Artículos seleccionados . . . . .	25
3.3 Resumen de la Revisión Sistemática . . . . .	33
<b>4 Resultados Esperados</b>	<b>34</b>
4.1 Describir los procesos de negocios típicos en las Instituciones de Educación Superior, que requieran tareas intensivas en conocimiento . . . . .	34
4.2 Métodos y técnicas de la Ingeniería de Conocimiento que pueden ser usados en los procesos de memoria organizacional . . . . .	36

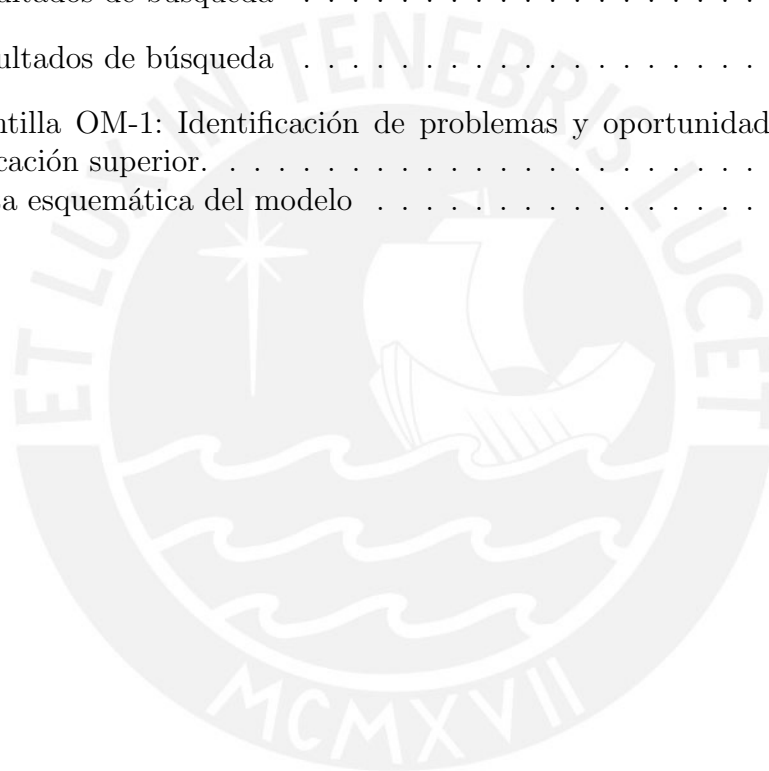
4.2.1	La Ingeniería del Conocimiento como apoyo a la adquisición de conocimiento . . . . .	37
4.2.2	La Ingeniería del Conocimiento como apoyo a la codificación del Conocimiento . . . . .	39
4.2.3	La Ingeniería del Conocimiento como apoyo al uso del conocimiento . . . . .	40
4.3	Procesos de Memoria Organizacional en el contexto de Instituciones de Educación Superior . . . . .	41
4.3.1	Adquisición . . . . .	41
4.3.2	Codificación . . . . .	41
4.3.3	Recuperación . . . . .	41
<b>5</b>	<b>Modelo Propuesto</b>	<b>43</b>
5.1	Modelo de Organización . . . . .	43
5.2	Descripción del Modelo propuesto . . . . .	45
5.2.1	Composición . . . . .	46
5.2.2	Ambiente . . . . .	48
5.2.3	Estructura . . . . .	48
5.2.4	Mecanismo . . . . .	49
5.2.5	Validación del Modelo Propuesto a través de comparación con otras arquitecturas conceptuales de la revisión de literatura . . . . .	49
<b>6</b>	<b>Conclusiones y Trabajos Futuros</b>	<b>55</b>
6.1	Trabajos Futuros . . . . .	56

# Índice de figuras

1.1	Conjunto de modelos de CommonKADS (Schreiber and Akkermans, 1999) . . . . .	8
2.1	Datos, información y conocimiento (Watson, 2003). . . . .	11
2.2	Modelado de Memoria Organizacional según (Ramesh, 1997). . . . .	14
2.3	Modelado de Memoria Organizacional según (Abecker et al., 1998). . . . .	15
2.4	Modelado de Memoria Organizacional según (Nevo and Wand, 2005). . . . .	16
2.5	Modelado de Memoria Organizacional según (Jackson, 2008). . . . .	16
3.2	Arquitectura de una memoria organizacional para seguridad vial (Boury-Brisset and Tourigny, 2000). . . . .	26
3.1	Arquitectura de referencia de Memoria Corporativa (Kuhn and Abecker, 1997). . . . .	27
3.3	Arquitectura conceptual (Zouaq et al., 2006). . . . .	29
3.4	Memoria organizacional de dos capas (Ju, 2006). . . . .	29
3.5	La arquitectura <i>Onto-Dom</i> (Ale et al., 2008). . . . .	30
3.6	Arquitectura de un sistema de gestión del conocimiento para una universidad (Adil Laoufi, 2011) (Adaptado al español). . . . .	31
3.7	Propuesta de arquitectura conceptual para sistemas de memoria organizacional (Toledo et al., 2011). . . . .	32
4.1	Infraestructura de Información como apoyo a la construcción de sistemas de memoria organizacional (Abecker et al., 1998) . . . . .	36
5.1	Arquitectura de software propuesta para instituciones de educación superior . . . . .	46
5.2	Ejemplo de mapeo semántico de un documento - Libro de (Zelle, 2002) . . . . .	48

# Índice de Tablas

3.1	Resultados de búsqueda . . . . .	21
3.2	Cadena de búsqueda . . . . .	22
3.3	Resultados de búsqueda . . . . .	23
3.4	Cadena de búsqueda . . . . .	24
3.5	Resultados de búsqueda . . . . .	24
4.1	Resultados de búsqueda . . . . .	35
5.1	Plantilla OM-1: Identificación de problemas y oportunidades en la educación superior. . . . .	44
5.2	Vista esquemática del modelo . . . . .	45



# Capítulo 1

## Presentación de la Investigación

El presente capítulo describe la definición del problema, indicando de manera detallada cual es la problemática de la construcción de sistemas de gestión del conocimiento en instituciones de educación superior, posteriormente se definen los objetivos tanto general como específicos. Asimismo, se justifica el proyecto de investigación y se detalla los métodos y procedimientos.

### 1.1. Introducción

El conocimiento es uno de los activos más importantes de las organizaciones. La Ingeniería de Conocimiento soporta los procesos de memoria organizacional, sin embargo es criticada por su esquema de rápido prototipado, en el cual la adquisición del conocimiento es tratada directamente como un desarrollo de software de computadora.

Las instituciones de educación superior son un lugar natural para la gestión del conocimiento puesto que tienen como misión primordial la transferencia de conocimientos y la investigación científica, sin embargo de la revisión del estado se ha evidenciado que se han realizado pocos estudios relacionados a memoria organizacional en este tipo de organizaciones.

La gestión del conocimiento ha tenido buenos resultados en la industria ((Grunstein and J.-P, 1996; Nonaka, 2000) apud (Adil Laoufi, 2011)), sin embargo hay pocas aplicaciones en universidades.

El objetivo general de la investigación es definir una arquitectura para sistemas de Memoria Organizacional para Instituciones de Educación Superior. Para conseguir los objetivos específicos se usará primero la técnica de revisión sistemática en los principales indexadores académicos IEEE, SCOPUS, ACM y ScienceDirect. Además se hará uso de la metodología de Ingeniería de Conocimiento CommonKADS, modelo



Césmico de (Bunge, 2006), modelo de arquitectura propuesta por (Melgar Sasieta, 2011), finalmente se realizará una comparación de la arquitectura propuesta con los modelos y/o arquitecturas de la revisión del estado del arte.

## 1.2. Definición del Problema

El conocimiento es uno de los activos más importantes de las organizaciones y afecta decisivamente su competitividad (Abecker et al., 1998). La competitividad de una organización es definida como la capacidad de conseguir una posición competitiva favorable que le permita alcanzar, sostener y mejorar determinada situación socioeconómica.

La memoria organizacional se ocupa de la reutilización y del intercambio de conocimiento (Ackerman and Hadverson, 2000) para su uso en las actividades actuales, mejorando la efectividad organizacional (Stein, 1995). Además, la memoria organizacional se puede definir como un sistema capaz de almacenar las cosas percibidas, experimentadas o vividas más allá de la situación actual, y permite recuperarlas y usarlas en un futuro (Lehner and Maier, 2000). Es el conocimiento de cómo hacer las cosas, la forma de abordar los problemas, la forma en que se deben tratar los unos a los otros (Jackson, 2008).

Asimismo, la memoria organizacional es definida comúnmente en términos de contenido de la memoria organizacional y los procesos asociados a la memoria organizacional (Stein, 1995). El contenido de la memoria organizacional esta compuesto por todo aquello que se encuentre documentado como información general de la organización, reportes de compras, inventarios, políticas de contratación de la empresa, manuales de procedimientos y archivos de computadora, etc.; asimismo también está compuesto por aquello que no está documentado como experiencias, formas de pensar, actitudes sobre la toma de decisiones, opiniones, anécdotas, etc. Los procesos asociados a una memoria organizacional son la adquisición, codificación y recuperación del conocimiento. Mediante estos procesos se traslada conocimiento al presente, permitiendo influir en las actividades actuales, lo que puede conllevar a niveles más altos o más bajos de competitividad organizacional.

La ingeniería de conocimiento soporta los procesos de memoria organizacional proporcionando una serie de técnicas para la gestión del conocimiento como codificación, organización y recuperación del conocimiento haciendolo accesible y recuperable. Sin embargo esta es siendo observada por el esquema de rápido prototipado que se viene utilizando, en el cual la adquisición del conocimiento es tratada directamente como un proceso de un sistema de información. El rápido prototipado se ha vuelto popular en los sistemas basados en conocimiento, dado que permite aprender y cambiar el curso del proyecto cuando sea necesario, pero la desventaja de este enfoque es su naturaleza, difícil de predecir y gestionar (Kingston, 2007). Una arquitectura de software por su naturaleza permite definir un conjunto de estructuras necesarias para el entendimiento de un sistema y por consiguiente apoyaría la construcción exitosa de un sistema de memoria organizacional.



Las instituciones de educación superior o universidades, son organizaciones intensivas en conocimiento dado que tienen 2 misiones principales: la transferencia de conocimiento e investigación científica (Adil Laoufi, 2011).

Pese a que las instituciones de educación superior son organizaciones intensivas en conocimientos, no se han desarrollado muchos estudios sobre memoria organizacional en estas organizaciones (ver Primera Búsqueda del Estado del Arte). Al respecto solo se encontraron 4 estudios relacionados a memoria organizacional en instituciones de educación superior, y estos abordan principalmente la construcción de sistemas de memoria organizacional y solo en uno de ellos se propone una arquitectura de sistema de memoria organizacional (Adil Laoufi, 2011), el cual es capaz de integrar diferentes ontologías, permitiendo el razonamiento y recuperación inteligente de información.

En base a este contexto surge la pregunta que motiva este proyecto. ¿Cómo representar una arquitectura para sistemas de Memoria Organizacional en Instituciones de Educación Superior?.

### 1.3. Objetivo General

Definir una arquitectura para sistemas de Memoria Organizacional en Instituciones de Educación Superior.

### 1.4. Objetivos Específicos

- OE1** Describir los procesos de negocios típicos en las Instituciones de Educación Superior, que requieran tareas intensivas en conocimiento.
- OE2** Detallar los métodos y técnicas de la Ingeniería del Conocimiento que pueden ser usados en los procesos de Memoria Organizacional.
- OE3** Explicar los procesos de Memoria Organizacional en el contexto de Instituciones de Educación Superior.
- OE4** Diseñar una propuesta de arquitectura para sistemas de Memoria Organizacional en Instituciones de Educación Superior.
- OE5** Validar la Arquitectura propuesta a través de un estudio cualitativo.

## 1.5. Justificación

¿Cómo representar una arquitectura para sistemas de Memoria Organizacional en Instituciones de Educación Superior? es la pregunta que motiva el planteamiento de este proyecto de investigación.

Las Instituciones de Educación Superior por lo general poseen sistemas de información o plataformas informáticas que se encargan de la administración y control de los datos de la organización, siendo común que estos sistemas almacenen tanto la información académica (proyectos de investigación de profesores, alumnos, artículos publicados, proyectos de fin de carrera, disertaciones, tesis, entre otros) así como la información administrativa (planillas, presupuesto, entre otros). Estos sistemas de información por lo general son construidos sobre bases de datos relacionales. Es por esta razón que estos sistemas solamente pueden almacenar datos. Al procesar estos datos dentro de los sistemas informáticos, estos se convierten en información.

El problema con las bases de datos, tal como su nombre lo indica, es que estas solamente almacenan datos. Los datos por si solos son elementos en bruto (números, textos, fechas, etc.) que no permiten realizar inferencia de ningún tipo. Para ilustrar un ejemplo vamos a considerar un sistema de información en donde se almacene la información académica de los profesores como por ejemplo artículos publicados en revistas y eventos (congresos, simposios, workshops, etc). Por lo general estos sistemas solicitan que los investigadores registren su producción académica ingresando sus títulos, co-autores, nombre de evento o revista, entre otros datos relevantes de la producción. Las bases de datos relacionales permiten recuperar estos datos de una forma fácil, simple y rápida siendo posible la elaboración de reportes del tipo: ¿cuántos artículos publicó determinado profesor en determinado año?, ¿cuántos artículos publicaron los profesores de determinado departamento? y ¿cuántos artículos se han publicado por determinada línea de investigación?

Los sistemas de memoria organizacional no solo necesitan datos en bruto sino conocimiento que pueda ser recuperado fácilmente por los individuos de una organización y les permita aprender de los errores y potencializar las lecciones aprendidas. Los sistemas creados a partir de bases de datos buscan texto pero sin tener en consideración el contexto, estos sistemas no permiten recuperar los datos cuando el término de búsqueda no coincide con el término almacenado en la base de datos. Por ejemplo si una publicación tuviese el nombre de “Una Arquitectura para Sistemas de Memoria Organizacional en Instituciones de Educación Superior”, es muy probable que este título este almacenado en algún campo de la base de datos, probablemente también estén almacenadas las palabras claves seleccionadas por el autor como “memoria organizacional, sistemas de memoria organizacional, ingeniería del conocimiento, modelo, sistemas basados en conocimiento”. Al ingresar como término de búsqueda algunas palabras que coincidan con el título o las palabras claves, un sistema basado en base de datos lo podrá retornar. Pero si alguien ingresa como término de búsqueda “universidad”, ciertamente este título no será retornado a pesar que una institución de educación superior podría ser una universidad. El título no será retornado pues los sistemas construidos son sintácticos, es necesario incluirles semántica para que se puedan realizar inferencias, es en este punto en donde la

Ingeniería del Conocimiento es aplicada.

La Ingeniería del Conocimiento utiliza formalismos que permiten representar el conocimiento haciendo posible que la semántica sea entendida por las personas y por las computadoras facilitando búsquedas semánticas, además al trabajar a nivel de conceptos la búsqueda podría inclusive recuperar artículos escritos en otros idiomas debido a que la recuperación semántica no trabaja con términos sino con conceptos y los conceptos son independientes del lenguaje. La Ingeniería del Conocimiento se basa en la aplicación de Inteligencia Artificial para poder realizar las inferencias.

Además, la Ingeniería del Conocimiento que soporta los procesos de memoria organizacional es observada principalmente por su enfoque de prototipado (Kingston, 2007), en el cual la adquisición de conocimiento es manejada directamente como un desarrollo iterativo de software (técnica reconocida de la ingeniería de software). Si bien el enfoque de prototipado puede brindar rápidamente un sistema de Memoria Organizacional, es difícil de validar y realizar el mantenimiento del sistema, además es complicado de predecir y gestionar el desarrollo del sistema; puesto que se cuenta con muy poca documentación. Una solución para los problemas descritos se puede encontrar haciendo uso de una metodología de Ingeniería del Conocimiento a través de diferentes modelos con diferentes niveles de detalle y del modelo césmico de Bunge que permite describir cualquier tipo de sistema. La adquisición del conocimiento resulta particularmente crítico puesto que el conocimiento según (Walsh and Ungson, 1991) se retiene en seis componentes de la organización: persona, cultura organizacional, cambio organizacional, estructuras organizativas, la ecología organizacional y archivos externos, por lo que la adquisición no puede ser tratada por métodos convencionales de iteración.

Según ((Grunstein and J.-P, 1996; Nonaka, 2000) apud (Adil Laoufi, 2011)) la gestión del conocimiento ha tenido buenos resultados en la industria, sin embargo hay muy pocas aplicaciones en universidades, pese a ser un lugar natural de creación y transferencia de conocimiento. Además, existen diversos estudios sobre memoria organizacional, sin embargo la mayoría de ellos están aplicados a instituciones tradicionales y no a instituciones de educación superior.

Además, de la revisión de literatura realizada el 17 de Octubre del 2013 en los principales indexadores académicos como: ACM, IEEE, Scopus y ScienceDirect, se encontró solo 04 artículos relacionados a Sistemas de Memoria Organizacional en Instituciones de Educación Superior, sin embargo solo en uno de ellos se propone una arquitectura o framework para construcción de sistemas de memoria organizacional (ver Primera Búsqueda del Estado del Arte), lo cual confirma el ineditismo del presente proyecto de investigación.

Una arquitectura para sistemas de Memoria Organizacional en Instituciones de Educación Superior, servirá como base para la construcción de un *framework* que soporte la construcción de Sistemas Basados en Conocimiento.

## 1.6. Métodos y procedimientos

**OE1** Describir los procesos de negocio típicos en las Instituciones de Educación Superior, que requieran tareas intensivas en conocimiento.

**R1.1** Un reporte técnico donde se describan los procesos de negocio típico de las Instituciones de Educación Superior que requieran tareas intensivas en conocimiento.

**Método para obtener R1.1** Se realizará una revisión sistemática de la literatura en los principales indexadores de artículos académicos: ACM, IEEE, Scopus y ScienceDirect. La revisión sistemática seguirá la guía de propuesta por (Kitchenham, 2004), que según dicho autor es la más adecuada para investigaciones en ingeniería de software.

La revisión sistemática de la literatura permite identificar, evaluar e interpretar los trabajos de investigación relevantes y disponibles para una pregunta de investigación, área temática o fenómeno de interés. Los estudios individuales que contribuyen a una revisión sistemática son llamados estudios primarios y la revisión sistemática es una forma de estudio secundario. Se realizarán diversas búsquedas en los principales indexadores académicos, cada fase de búsqueda está compuesta por el objetivo, criterios de inclusión/exclusión y parámetros de búsqueda. Finalmente, se describirán brevemente los artículos académicos seleccionados.

**OE2** Detallar los métodos y técnicas de la Ingeniería del Conocimiento que pueden ser usados en los procesos de Memoria Organizacional.

**R2.1** Un reporte técnico que describa los métodos y técnicas de la Ingeniería del Conocimiento que sean los más adecuados para la Memoria Organizacional en el contexto de Instituciones de Educación Superior.

**Método para obtener R3.1** Se realizará una revisión sistemática de la literatura en los principales indexadores académicos: ACM, IEEE, Scopus y ScienceDirect.

**OE3** Explicar los procesos de Memoria Organizacional en el contexto de Instituciones de Educación Superior (codificación, organización y recuperación del conocimiento).

**R3.1** Un reporte técnico que describa los procesos de Memoria Organizacional en el contexto de Instituciones de Educación Superior.

**Método para obtener R2.1** Se realizará una revisión sistemática de la literatura en los principales indexadores académicos: ACM, IEEE, Scopus y ScienceDirect.

**OE4** Diseñar una propuesta de Arquitectura de Memoria Organizacional para Instituciones de Educación Superior.

**R4.1** Una propuesta de Arquitectura de software de Memoria Organizacional adecuada para las Instituciones de Educación Superior.



**Método para obtener R4.1** Para la definición de la arquitectura propuesta se usará el modelo CESM de (Bunge, 2006) y la metodología CommonKADS (Schreiber and Akkermans, 1999). Asimismo, para enriquecer el contenido del repositorio de documentos se usará ontologías, con las cuales se realizarán anotaciones semánticas en los documentos de investigación. Además, se tomará como referencia el modelo propuesto de (Melgar Sasieta, 2011), donde se desarrolla un modelo para la visualización de conocimiento basado en imágenes semánticas.

El Modelo CESM describe a los objetos complejos como un sistema cuyas partes o componentes se mantienen unidos por algún tipo de relaciones (Bunge, 2006). Estas relaciones son lógicas en el caso de sistemas conceptuales, como una teoría; y son físicos en el caso de sistemas concretos, como un átomo, células, sistema inmunológico, familia, hospital, etc. Cada sistema " $\sigma$ " está compuesto por su: Composición ( $C$ ), Ambiente ( $E$ ), Estructura ( $S$ ) y Mecanismo ( $M$ ) (Bunge, 2006):

$$\mu(\sigma) = \langle C(\sigma), E(\sigma), S(\sigma), M(\sigma) \rangle$$

**$C(\sigma)$**  : Describe al conjunto de partes de " $\sigma$ ".

**$E(\sigma)$**  : La colección de elementos de entorno que influyen en " $\sigma$ " o los elementos que son influenciados por " $\sigma$ ".

**$S(\sigma)$**  : La estructura es el conjunto de relaciones o enlaces que mantienen unidos los componentes de " $\sigma$ " y los que enlazan a su entorno.

**$M(\sigma)$**  : Es la colección de procesos que se dan dentro de un sistema y que lo hacen cambiar de alguna forma ( la investigación es un proceso que se realiza en una institución de educación superior).

La metodología CommonKADS (Acrónimo de "*Knowledge Acquisition and Design System*") muestra a la Ingeniería del Conocimiento como una actividad de modelado, donde cada modelo es una propuesta de abstracción de una porción del mundo real (representación intermedia entre la realidad y alguna aplicación del modelo, como la implementación de un sistema de base de conocimientos) (Schreiber and Akkermans, 1999). CommonKADS propone un conjunto de 6 modelos de conocimiento (Ver figura 1.1), si se desarrolla todos los modelos, el conocimiento debería transformarse gradualmente desde un conjunto de requerimientos y necesidades de la organización a una especificación de diseño de alto nivel de un sistema basado en conocimiento. De los 06 modelos descritos por la metodología CommonKADS, en el presente trabajo de investigación se usará el modelo de organización, el cual permite representar los procesos, estructuras y recursos de una organización, con el objetivo de identificar problemas y oportunidades para sistemas de conocimiento. Además, la metodología CommonKADS está compuesta por los modelos de tareas, agentes, conocimiento, comunicación y diseño.

Por otro lado, una ontología es una definición explícita de una conceptualización (Gruber, 1993). La conceptualización se refiere a una interpretación estructurada de una parte del mundo, las cuales usan las personas para pensar y comunicar acerca del mundo (Borst, 1997). Por ejemplo en el presente

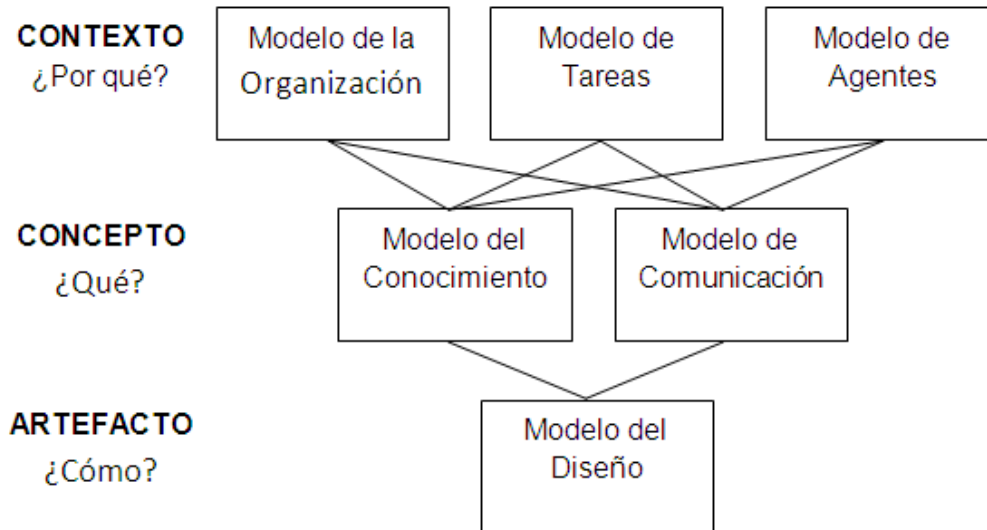


Figura 1.1: Conjunto de modelos de CommonKADS (Schreiber and Akkermans, 1999)

proyecto la conceptualización puede incluir a los docentes que pueden ser clasificados en diversos tipos como Docente Principal, Asociado o Asistente, y los docentes por lo general pertenecen a una facultad de la universidad. Otro ejemplo de conceptualización en biología, las especies en base a sus hábitos de alimentación pueden sub-categorizarse en herbívoros, carnívoros y omnívoros.

Para (Borst, 1997) la definición de ontología es muy similar a la de (Gruber, 1993), una ontología es una especificación formal de una conceptualización, derivándose dos ideas claves. La primera que una ontología debe ser especificada usando un lenguaje formal, que pueda ser procesado por ordenadores y no solo por personas, esta perspectiva es particularmente útil para la implementación de un sistema de gestión del conocimiento. La segunda, que una ontología debe ser un conocimiento compartido por un grupo de personas sobre un dominio de aplicación en particular, utilizando un vocabulario común en el contexto de un dominio de aplicación en particular.

El conocimiento de un dominio particular puede ser representando a través de los 5 componentes de las ontologías: conceptos, relaciones, funciones, instancias y axiomas. Los conceptos son las ideas básicas que se intentan formalizar, los conceptos pueden ser clases de objetos, métodos, planes, estrategias, procesos de razonamiento entre otros, por ejemplo, en el presente trabajo se definen usuarios de la memoria organizacional, estos usuarios son personas del mundo académico (docentes y estudiantes). Las relaciones representan la interacción y enlace entre los conceptos del dominio suelen formar la taxonomía del dominio, por ejemplo subclase-de, parte-de, parte-exhaustiva-de, conectado-a, entre otros, por ejemplo, en el presente trabajo de investigación, una tesis pertenece a un área de investigación como ingeniería de software, computación gráfica,

compiladores, estructura de datos, entre otros. Las funciones son un tipo concreto de relación donde se identifica un elemento mediante el cálculo de una función que considera varios elementos de la ontología, por ejemplo si se brinda un proyecto de investigación a una función y/o relación que obtenga el área de investigación, siempre devolverá el mismo resultado. Las instancias se utilizan para representar objetos determinados de un concepto, por ejemplo para el concepto Docente, se puede tener muchas instancias como un Docente llamado Juan, con un número de DNI, con un área de investigación, entre otros. Los axiomas son teoremas que se declaran sobre relaciones que deben cumplir los elementos de la ontología, por ejemplo, un docente no podría pertenecer a más de una facultad a la vez.

Las ontologías se pueden clasificar de acuerdo al nivel de generalidad de la ontología en ontologías de alto nivel, dominio, tareas y de aplicación (Guarino, 1998). Las ontologías de alto nivel describen conceptos muy generales como: tiempo, espacio, objeto, evento, acción, etc., los cuales son independientes de un problema particular o dominio de aplicación, por lo cual es razonable tener ontologías de alto nivel para grandes comunidades de usuarios. Las ontologías de dominio describen el vocabulario relacionado a un dominio genérico (como medicina, ciencias de la computación, agricultura, etc.). Las ontologías de tareas describen el vocabulario relacionado a una tarea o actividad genérica (como diagnóstico, vender, estudiar, investigar, etc.), especializando los términos definidos en las ontologías de alto nivel. Las ontologías de aplicación describen conceptos dependientes de un dominio y tarea en particular, las cuales son a menudo especializaciones de ambas ontologías, estos conceptos corresponden a roles realizados por entidades de dominio mientras realizan cierta actividad.

**OE5** Validar la Arquitectura propuesta a través de un estudio cualitativo.

**R5.1** Un reporte técnico que muestre la validez de la arquitectura propuesta.

**Método para obtener R5.1** Se validará la arquitectura a través de un estudio cualitativo, comparando la arquitectura propuesta con otras arquitecturas de la revisión del estado del arte.

**R5.2** Un reporte técnico con los resultados de la comparación de la arquitectura propuesta con otras arquitecturas de Memoria Organizacional.



# Capítulo 2

## Marco Conceptual

### 2.1. Introducción

En esta sección se definen los conceptos que son necesarios e importantes para la comprensión del problema, dichos conceptos están estrechamente relacionados entre sí. Se definirá los conceptos de Gestión del Conocimiento, Memoria Organizacional, Arquitectura de Software e Institución de Educación Superior. Asimismo, se profundizará en aquellos conceptos que son relevantes para el presente trabajo de Investigación.

### 2.2. Gestión del Conocimiento

#### 2.2.1. Conocimiento

Para entender la Gestión del Conocimiento, primeramente se debe entender claramente el término conocimiento. Los términos datos, información y conocimiento están estrechamente relacionados (Watson, 2003).

**Datos:** Los datos son un conjunto de símbolos discretos sobre una característica o medición de algo, un dato no proporciona información sobre el contexto, por ejemplo el número 9 es un dato, el cual representa una cantidad, pero no tiene ningún contexto en el que se pueda usar dicho dato.

**Información:** Hace falta información para proporcionar al dato un significado relevante en un contexto determinado. La información es un conjunto de datos procesados con sentido en un contexto determinado, por ejemplo al número 9, le agregamos el contexto de “*la temperatura en la ciudad de Lima - Perú es 9 grados centígrados*”, se entiende que hay una relación entre la temperatura de la ciudad de Lima y el número 9.

**Conocimiento:** El conocimiento es una mezcla de experiencias, valores, información contextual, reconocimiento de patrones, entendimiento de experto que provee un marco para evaluar e incorporar nuevas experiencias e información. En el ejemplo de la temperatura en la ciudad de Lima, la mayoría de nosotros entiende que incluso para invierno, 9 grados centígrados representa un día muy frío, por lo cual es conveniente abrigarse adecuadamente.

El conocimiento se deriva de la información y la información se deriva de los datos (Ver figura 2.1). Datos, información y conocimiento, se constituyen en tres conceptos los cuales no deben confundirse, ya que de ella depende el buen manejo de la toma de decisiones en una organización.

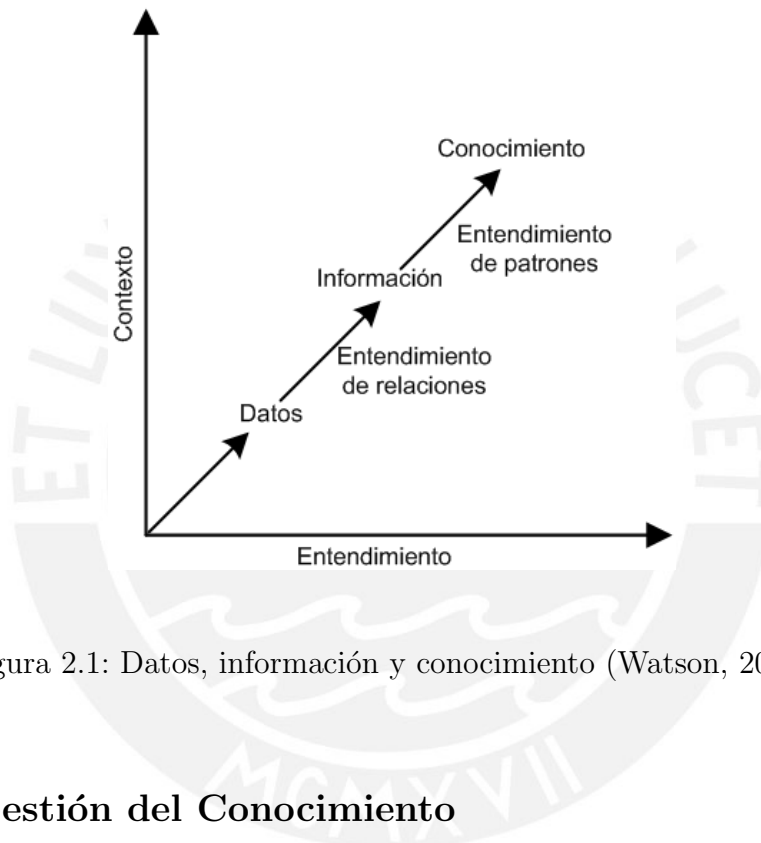


Figura 2.1: Datos, información y conocimiento (Watson, 2003).

### 2.2.2. Gestión del Conocimiento

Se tiene diferentes definiciones de Gestión del Conocimiento, sin embargo todas estas definiciones presentan descripciones similares, la gestión del conocimiento se refiere a la forma de capturar, organizar y comunicar el conocimiento tácito y explícito de los empleados, para que otros empleados usen dicho conocimiento para ser más efectivos y productivos en su trabajo (Jung et al., 2007). La gestión del conocimiento administra las diferentes fases del ciclo de vida del conocimiento: identificación, adquisición, desarrollo, diseminación, uso y preservación del conocimiento (Abecker et al., 1998).

De acuerdo a (Vas, 2001), se tiene 02 formas de clasificar el conocimiento: la primera forma clasifica el conocimiento en tácito o explícito, la segunda forma clasifica al conocimiento en: declarativo, procedimental y heurístico, asimismo de la revisión de la literatura, se observa que algunos autores clasifican al conocimiento en: individual y grupal (colectivo).

El conocimiento explícito es el saber que puede ser transmitido o compartido entre las personas o el seno de una organización con relativa facilidad, es un conocimiento formal (estructurado), almacenado y distribuido. Es todo aquel conocimiento capturado y codificado en manuales, normas y procedimientos, por lo que resulta sencilla su transmisión. El conocimiento tácito es el saber que presenta dificultad y complejidad para ser transmitido o comunicado a otras personas. Es aquel que reside en las experiencias de aprendizajes personales que reside en cada persona, por tanto resulta difícil capturar, estructurar, almacenar y distribuir el conocimiento. Son modelos mentales, intuiciones, creencias, perspectivas relacionadas con la concepción personal del mundo y/o habilidades (Nonaka, 1994).

### 2.2.3. Sistemas de Gestión del Conocimiento

Los sistemas de gestión del conocimiento, son una clase de sistemas de información aplicado a la gestión del conocimiento organizacional. Los sistemas de gestión del conocimiento son herramientas de tecnologías de la información que se desarrollan para apoyar y mejorar los procesos organizacionales de creación, almacenamiento/recuperación, transferencia y aplicación del conocimiento (Alavi and Leidner, 2001). Cabe señalar que no todos los proyectos de gestión del conocimiento conllevan a la implementación de tecnologías de la información, pero es muy común enfocarse en las tecnologías a expensas de los aspectos culturales y sociales de la gestión del conocimiento.

### 2.2.4. Procesos de la Gestión del Conocimiento

De acuerdo a (Alavi and Leidner, 2001), los procesos de la gestión del conocimiento son: creación, almacenamiento/recuperación, transferencia y aplicación del conocimiento. Las tecnologías de la información apoyan los diversos procesos de la gestión del conocimiento a través de diversas herramientas. El uso de herramientas tecnológicas mejora la creación de conocimiento por ejemplo a través del uso de Datawarehousing, datamining, repositorio de documentos, entre otros. Además, se puede hacer uso de diversos sistemas de información para apoyar los procesos de trabajo colaborativo, coordinación y comunicación; las intranets permiten acceder a gran cantidad de información institucional que permiten crear conocimiento tácito en las personas.

Para que el conocimiento pueda ser almacenado y recuperado es necesario estructurarlo con ontologías, anotaciones semánticas y taxonomías, con lo cual el conocimiento se formaliza y se convierte en explícito y articulado. Las modernas tecnologías de almacenamiento de computadoras y técnicas sofisticadas de recuperación de información como lenguajes de consulta, bases de datos multimedia y sistemas de gestión de base de datos, son herramientas efectivas que mejoran la memoria organizacional. Estas herramientas tecnológicas incrementan la velocidad de acceso a la memoria organizacional (Alavi and Leidner, 2001).

Las tecnologías de la información mejoran el proceso de transferencia de conocimiento brindando una serie de herramientas como: boletines electrónicos, foros de discusión, disponibilidad de nuevos canales de comunicación, directorios de conocimiento, redes internas y rápido acceso al conocimiento.

El conocimiento por si mismo no genera valor, la aplicación del conocimiento genera valor y crea las ventajas competitivas de una organización (Alavi and Leidner, 2001). Las herramientas de tecnologías de la información facilitan la aplicación del conocimiento. Los sistemas expertos y los sistemas de apoyo a la toma de decisiones se diseñan con el propósito de aplicar el conocimiento adecuado para la ejecución de tareas organizacionales.

## 2.3. Memoria Organizacional

La Memoria Organizacional (MO), se puede definir como un sistema capaz de almacenar las cosas percibidas, experimentadas o vividas más allá de la situación actual, y permite recuperarlas y usarlas en un futuro (Lehner and Maier, 2000). Es el conocimiento de cómo hacer las cosas, la forma de abordar los problemas y cuestiones, la forma en que se deben tratar los unos a los otros (Jackson, 2008). La memoria de la organización se ocupa de la reutilización e intercambio del conocimiento (Ackerman and Hadverson, 2000) para su uso en las actividades actuales, mejorando así la eficacia de la organización (Stein, 1995). El concepto de memoria organizacional está íntimamente ligada al concepto de aprendizaje organizacional (Lehner and Maier, 2000; Atwood, 2002; Ackerman and Hadverson, 2004). La memoria y el aprendizaje organizacional son procesos sociales profundamente enraizados y la eficacia organizacional se logra a través de una integración sinérgica de una cultura de intercambio de conocimientos y recursos tecnológicos. De acuerdo con (Jackson, 2008) el aprendizaje organizacional hace uso de procesos dentro de la organización que mantiene y usa la memoria organizacional, estos procesos garantizan que la memoria sea almacenada a lo largo del tiempo y pueda ser accedida cuando sea necesario. Pero para que una organización pueda aprender, los resultados de las experiencias del pasado (memoria organizacional) deben estar disponibles para cada individuo de la organización (Ackerman, 1998).

El conocimiento es uno de los activos más importantes de las organizaciones competitivas, la memoria organizacional extiende y amplifica este activo con la captura, organización, diseminación y reuso del conocimiento creado por sus empleados (Conklin, 2001). La memoria organizacional no es solo una forma para acumular y retener conocimiento, sino también permite compartirlo con los demás miembros de la organización.

### 2.3.1. Modelos de Memoria Organizacional

A continuación, se detalla los diferentes modelos de Memoria Organizacional que se han propuesto:

### 2.3.1.1. Modelo para representar una MO en diferentes niveles de abstracción

El meta modelo de (Ramesh, 1997) representa una memoria organizacional en diferentes niveles de abstracción (Ver figura 2.2). Este modelo tiene primitivas que permiten representar tanto a los agentes implicados en los procesos de organización, así como las entradas y salidas de estos procesos, además de las relaciones entre ellos. En este meta-modelo los *stakeholders* representan a los actores implicados en los procesos de toma de decisiones. El objeto representa las entradas y salidas de los procesos (requisitos, supuestos, juicios, justificaciones, alternativas, factores críticos de éxito, entre otros). Los objetos se crean por diferentes tareas de la organización, como tareas de análisis y diseño. Todos los objetos están documentados por fuentes que pueden ser objetos físicos, tales como los documentos o referencias. Las fuentes son los documentos de planificación, notas, memorandos, llamadas telefónicas, así como referencias a los distintos *stakeholders*. Los *stakeholders* crean las fuentes y desempeñan diferentes roles para la creación, uso y mantenimiento de varios objetos. La trazabilidad a través de diversos objetos está representado por las trazas a los enlaces.

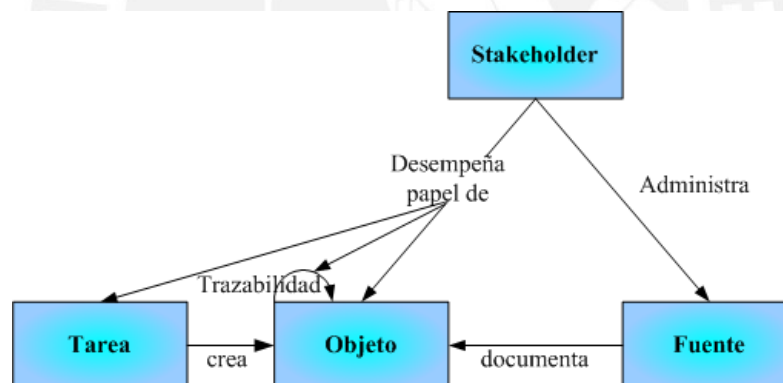


Figura 2.2: Modelado de Memoria Organizacional según (Ramesh, 1997).

### 2.3.1.2. Meta-modelo basado en información

El meta-modelo basado en información de (Abecker et al., 1998) está basado en tres ontologías para describir los elementos de información y conocimiento. En este enfoque, cada elemento es descrito por una serie de conceptos representados en las ontologías de información, organización y de dominio (Ver figura 2.3).

El meta-modelo describe diferentes fuentes de información con sus respectivas estructuras, formatos y propiedades de acceso. El vocabulario es provisto por la ontología de información el cual contiene atributos y conceptos genéricos que se



pueden aplicar a todo tipo de información, asimismo provee enlaces y especializa algunos conceptos que se aplican a fuentes de información específicas. Básicamente la ontología de información cubre todos los aspectos de las fuentes de información y conocimiento que no sea contenido específico. También ofrece enlaces a la ontología de dominio la cual es usada para la descripción de contenido, y enlaces a la ontología de organización que es usado para describir contextos de los elementos de conocimiento. Para modelar el contenido de las fuentes de información se utiliza la ontología de dominio. Además para modelar el contexto de la información se utiliza la ontología de organización. El modelado de contexto se ocupa principalmente de describir el contexto previsto de utilización de un elemento y el contexto en que se creó el elemento. El modelado de contexto se ocupa principalmente de describir el contexto previsto de utilización de un elemento y el contexto con el cual se creó el elemento.

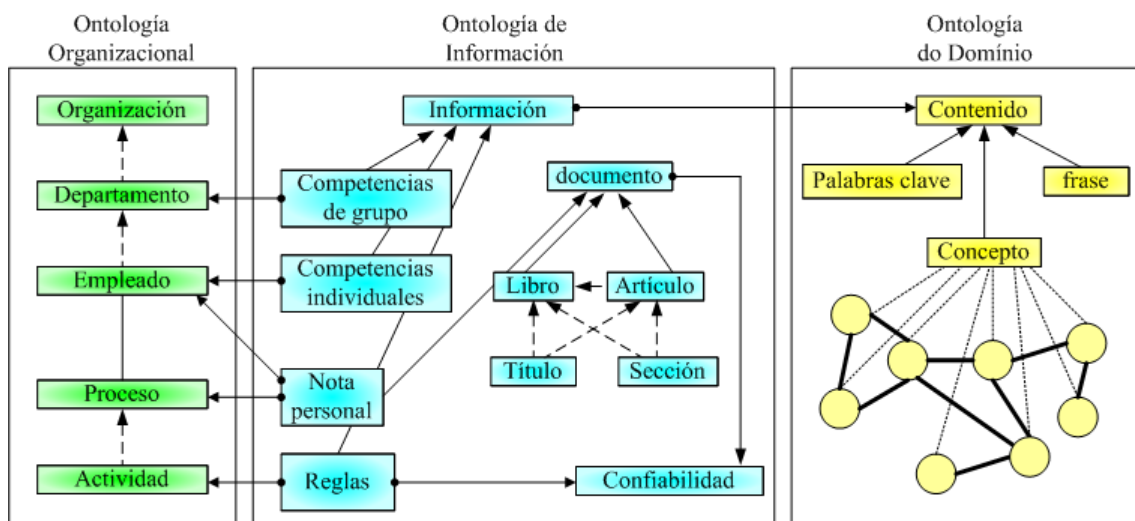


Figura 2.3: Modelado de Memoria Organizacional según (Abecker et al., 1998).

### 2.3.1.3. Modelo basado en tres dimensiones de memoria organizacional

(Nevo and Wand, 2005) propuso un modelo basado en tres dimensiones de memoria organizacional: retenedores de conocimiento, el conocimiento en sí y el meta conocimiento (Ver figura 2.4). En este modelo, los retenedores de conocimiento se caracterizan por el meta conocimiento cognitivo y descriptivo (por ejemplo, años de educación y experiencia). Estos retenedores tienen conocimiento de los conceptos y/o instancias de la organización. El conocimiento que tienen los retenedores se encuentra en forma de predicados (afirmaciones verdaderas) sobre el estado de las cosas en un momento dado, los cambios pasados y los posibles cambios en un dominio de interés.

### 2.3.1.4. Modelo ampliado de (Nevo and Wand, 2005)

(Jackson, 2008) extiende el modelo propuesto por (Nevo and Wand, 2005) mediante la adición de nuevas entidades en el modelo (Ver figura 2.5). En este nuevo

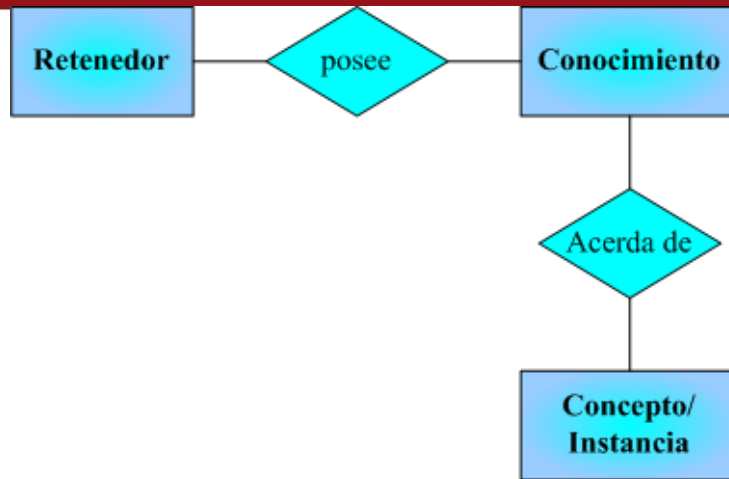


Figura 2.4: Modelado de Memoria Organizacional según (Nevo and Wand, 2005).

modelo los retenedores de conocimiento están relacionados a algún tipo de almacenamiento el cual puede ser de tipo tecnología, procedimientos, personas, entre otros. Se creó la entidad dominio del conocimiento el cual permite representar los conceptos, asimismo se han añadido sinónimos y definiciones de los conceptos. La entidad Conocimiento se ha vinculado a las entidades Modo de uso y Modo de Adquisición, para incluir información de contexto en el modelo.

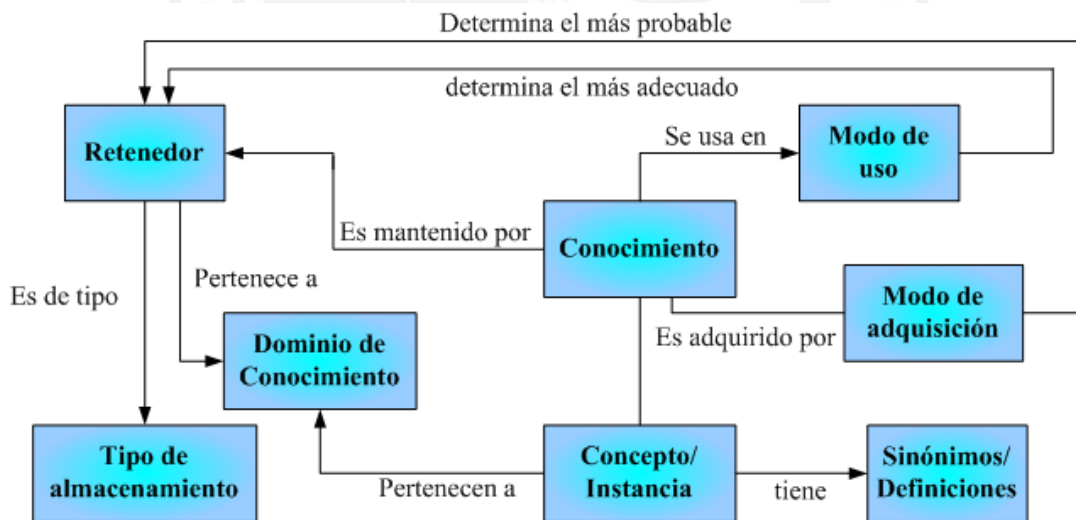


Figura 2.5: Modelado de Memoria Organizacional según (Jackson, 2008).

## 2.4. Arquitectura de Software

De acuerdo a (Garlan et al., 2010) la Arquitectura de Software define el conjunto de estructuras necesarios para el entendimiento de un sistema, el cual comprende elementos del software, relaciones entre estos y las propiedades de estos. Estas es-



estructuras cuidadosamente seleccionadas y diseñadas, son las claves para comprender los objetivos de diseño del sistema. Además estas estructuras son la clave para comprender la estructura. Para (Bass et al., 2003) la Arquitectura de Software de un programa o sistema de computadora es la estructura o estructuras del sistema, que comprende elementos del software, las propiedades visibles externamente de los elementos y la relación entre estos.

**La arquitectura define elementos:** La arquitectura expresa la información acerca de como los elementos se relacionan unos a otros. Esto significa que la arquitectura suprime cierta información acerca de elementos que no conciernen a su interacción. Así una arquitectura es principalmente una abstracción de un sistema que suprime detalles de elementos que no afectan como se usa, por quienes será usado, relacionado o interacción con otros elementos. En casi todos los sistemas modernos, los elementos interactúan con otros a través de interfaces que dividen los detalles acerca de un elemento en partes públicas y privadas. Arquitectura tiene que ver con el lado público de esta división, detalles privados de los elementos tiene que ver únicamente con la implementación interna y no con la arquitectura.

**Estructuras de un sistema:** Los sistemas comprenden normalmente más de una estructura y una sola estructura puede ser considerada una arquitectura, por ejemplo los proyectos no triviales se dividen en unidades de implementación, a estas unidades se les asigna ciertas responsabilidades y son la base de asignación de trabajo para equipos de programación. Este tipo de elementos constan de programas y datos del software, en otras implementaciones las unidades pueden llamar o acceder a programas y datos que son privados. En proyectos grandes, los elementos se dividirán en sub-equipos. Este tipo de estructuras usualmente usados para describir un sistema. Es una estructura muy estática, en la cual se enfoca en la funcionalidad del sistema, la cual se divide y asigna a unidades de implementación.

**Los sistema software tienen una arquitectura:** Todo sistema software tiene una arquitectura puesto que todo sistema puede ser visto como la composición de sus elementos y relaciones entre estos. Incluso en los casos más triviales, un sistema tiene una arquitectura de un solo elemento.

**Comportamiento de un elemento:** El comportamiento de todo elemento es parte de la arquitectura en la medida que dicho comportamiento puede ser observado y distinguido desde el punto de vista de otro elemento. Este comportamiento es el que permite a los elementos interactuar con otros, que son claramente parte de la arquitectura.

**Evaluación de la Arquitectura:** Es importante evaluar una arquitectura para un sistema en particular, puesto que no se debe de usar la técnica de prueba y error en la definición de una arquitectura, y además no se debe definir al azar la arquitectura de software.

Además, de acuerdo a (Group, 2000), una arquitectura está definida por las prácticas recomendadas como: organización fundamental de un sistema mediante sus componentes, relaciones entre ellos, el entorno y los principios que regulan su diseño y

evolución. Se debe comprender y controlar los elementos de un sistema para que se capture la utilidad, costo y riesgos del sistema, en la mayoría de los casos estos elementos son componentes físicos del sistema y sus relaciones, en otros casos serán componentes lógicos, y en algunos casos los elementos son principios duraderos o patrones que crean estructuras de organización duraderos. Esta definición de arquitectura busca abarcar los distintos usos, pero relacionados entre sí, fomentando así una definición más rigurosa que constituya la organización fundamental de un sistema en un dominio particular.

## 2.5. Institución de Educación Superior

Una institución de educación superior o universidad es básicamente una organización tradicional ya que está estructurada, ordenada y maneja recursos (materiales, humanos y financieros) para lograr un objetivo común (Adil Laoufi, 2011). Sin embargo se diferencia a través de 3 elementos principales:

**La diversidad de sus actores** Una institución de educación superior tiene diversos actores como estudiantes, profesores, administrativos, equipos técnicos y socios; los cuales tienen diferentes funciones, en donde los profesores y estudiantes crean e investigan nuevo conocimiento, para lo cual deben obtener documentación básica (libros, artículos, tesis, etc.).

**La heterogeneidad de recursos** Además de los recursos tradicionales, las instituciones de educación ofrecen una gran variedad de recursos a sus diferentes actores como documentos administrativos, cursos, manuales, portal académico, libros, catálogo de libros, libros electrónicos, patentes, reportes, etc.

**Misión de una institución de educación superior** La universidad desde su concepción tiene dos misiones principales la transferencia de conocimiento e investigación.

## Capítulo 3

# Revisión del Estado del Arte

### 3.1. Introducción

Existen pocas investigaciones sobre sistemas de memoria organizacional en instituciones de educación superior, pese a ser un lugar natural para la transferencia del conocimiento e investigación (Adil Laoufi, 2011). Por lo cual se ha realizado una revisión sistemática sobre investigaciones de sistemas de memoria organizacional en instituciones sin importar el rubro de negocio, lo cual ha permitido encontrar en total 7 investigaciones donde se propongan arquitecturas conceptuales de memoria organizacional.

La revisión sistemática se realizó sobre los indexadores académicos Scopus, ScienceDirect, ACM DL e IEEE, en total se seleccionaron 07 artículos del indexador SCOPUS, cabe señalar que los artículos seleccionados de los otros indexadores ya se encontraban contenidos en el resultado del indexador SCOPUS. Se realizaron 3 búsquedas sobre los indexadores antes mencionados con diferentes cadenas de búsqueda, posteriormente se muestran los resultados y se presenta una breve descripción de los artículos seleccionados.

### 3.2. Fases de la Revisión Sistemática

#### 3.2.1. Primera Búsqueda

Se realizó para determinar si ya se han desarrollado investigaciones similares en los que se hayan propuesto arquitecturas o frameworks conceptuales para instituciones de educación superior y/o universidades. Esta búsqueda permitió justificar el problema de investigación, en las siguientes búsquedas se amplió las opciones de

búsqueda.

### Objetivo

Encontrar investigaciones similares en instituciones de educación superior que permitan determinar la originalidad de la investigación. Asimismo, contribuyó en la justificación del problema de investigación.

### Criterios de inclusión

- Trabajos de investigación o artículos, los cuales deben estar relacionados a memoria organizacional en universidades.
- Trabajos de investigación o artículos, de los indexadores académicos Scopus, ScienceDirect, ACM DL e IEEE.

### Criterios de Exclusión

- Trabajos de investigación o artículos, que no se enfoquen en propuestas de arquitecturas/frameworks conceptuales para construcción de sistemas de memoria organizacional.
- Trabajos de investigación o artículos, donde se propone arquitecturas de sistemas de memoria organizacional enfocadas en el diseño de software y/o sistemas de información. Esta exclusión, obedece a que este tipo de arquitecturas se centran en aspectos tecnológicos de construcción e implementación de sistemas.

### Parámetros

A fin de ampliar las opciones de búsqueda, se usó las siguientes cadenas de texto para buscar en los indexadores:

- *(organizational memory and university) or.*
- *(organizational memory and education).*

Para ampliar las opciones de encontrar investigaciones similares, en los parámetros de búsqueda se han usado otras denominaciones de memoria organizacional como: *corporate memory* e *institutional memory*, por lo cual la cadena de búsqueda final se ha definido por la combinación de estas denominaciones con *university* and *education*. Por ejemplo se ha usado la siguiente cadena de búsqueda para el indexador ScienceDirect (para los indexadores Scopus, IEEE y ACM DL se realizaron modificaciones específicas de cada indexador):

- *(title-abstr-key("organizational memory") and title-abstr-key("university")) or.*
- *(title-abstr-key("organizational memory") and title-abstr-key("education")) or.*
- *(title-abstr-key("corporate memory") and title-abstr-key("university")) or.*
- *(title-abstr-key("corporate memory") and title-abstr-key("education")) or.*

- *(title-abstr-key(“institutional memory”) and title-abstr-key(“university”)) or.*
- *(title-abstr-key(“institutional memory”) and title-abstr-key(“education”)).*

La búsqueda fue realizada el 17/10/2013, retornando los resultados de la tabla 3.1.

Tabla 3.1: Resultados de búsqueda

Búsqueda	Memoria Organizacional		
	Fecha de la búsqueda	Cantidad de resultados	Artículos relacionados a la investigación
Scopus	17/10/2013	55	4
Science Direct	17/10/2013	6	0
ACM DL	17/10/2013	0	0
IEEE Xplore	17/10/2013	2	0

De la búsqueda realizada, solo se encontraron 4 artículos en SCOPUS, de los cuales solamente el artículo de (Adil Laoufi, 2011) propone una arquitectura de sistema de memoria organizacional.

### 3.2.2. Segunda Búsqueda

En la primera búsqueda solo se encontró un artículo en el cual se propone una arquitectura o framework para construcción de sistemas de memoria organizacional en instituciones de educación superior y/o universidades, por lo cual se ha ampliado la cadena de búsqueda a fin de encontrar más investigaciones en las que proponga arquitecturas o frameworks en cualquier tipo de organización. Asimismo, las búsquedas se realizaron en los títulos, resúmenes y palabras claves de los artículos.

#### Objetivo

Encontrar investigaciones en las que se proponga una arquitectura o framework para construcción de sistemas de Memoria Organizacional en diversos tipos de instituciones (no solo en instituciones de educación superior o universidades).

## Criterios de inclusión y exclusión

Se tomará en cuenta los mismos criterios de inclusión y exclusión de la "Primera Búsqueda".

## Parámetros

La cadena de búsqueda es: "Organizational memory and architecture", la misma que ha sido personalizada de acuerdo a la sintaxis de búsqueda de los indexadores: Scopus, Science Direct, ACM DL e IEEE Xplore:

Tabla 3.2: Cadena de búsqueda

Scopus	TITLE-ABS-KEY(organizational memory) and TITLE-ABS-KEY(architecture)
Science Direct	TITLE-ABSTR-KEY(organizational memory) AND TITLE-ABSTR-KEY(architecture)
ACM DL	((Title:organizational memory) and (Title:architecture)) or ((Abstract:organizational memory) and (Abstract:architecture)) or ((Keywords:organizational memory) and (Keywords:architecture))
IEEE Xplore	(Document Title : organizational memory AND Document Title : architecture ) OR (Abstract: organizational memory AND Abstract : architecture ) OR (Author Keywords: organizational memory AND Author Keywords : architecture ) OR (Index Terms: organizational memory AND Index Terms : architecture )

La búsqueda fue realizada el 10/10/2013, retornando los siguientes resultados:



Tabla 3.3: Resultados de búsqueda

Búsqueda	Organizational memory and architecture		
	Fecha de la búsqueda	Cantidad de resultados	Artículos seleccionados
Scopus	10/10/2013	49	6
Science Direct	10/10/2013	10	2
ACM DL	10/10/2013	13	1
IEEE Xplore	10/10/2013	9	2

En total se seleccionaron 06 artículos del indexador SCOPUS, además los artículos seleccionados de los indexadores ScienceDirect, IEEE y ACM DL ya estaban incluidos en los artículos seleccionados del indexador Scopus.

### 3.2.3. Tercera Búsqueda

Para ampliar las opciones de encontrar arquitecturas de software para construcción de memoria organizacional, se ha modificado la cadena de búsqueda de la **"Segunda Búsqueda"**, con lo cual la cadena de búsqueda es: *"Organizational memory and framework"*.

#### Objetivo

Encontrar investigaciones en las que se proponga una arquitectura o framework para construcción de sistemas de Memoria Orgazacional en diversos tipos de instituciones (no solo en instituciones de educación superior o universidades).

#### Criterios de inclusión y exclusión

Se tomará en cuenta los mismos criterios de inclusión y exclusión de la *"Segunda Búsqueda"*.

#### Parámetros

La cadena de búsqueda es: *"Organizational memory and framework"*, la misma que ha sido personalizada de acuerdo a la sintaxis de búsqueda de los indexadores: Scopus, Science Direct, ACM DL e IEEE Xplore:



Tabla 3.4: Cadena de búsqueda

Scopus	TITLE-ABS-KEY(organizational memory) and TITLE-ABS-KEY(framework)
Science Direct	TITLE-ABSTR-KEY(organizational memory) AND TITLE-ABSTR-KEY(framework)
ACM DL	((Title:organizational memory) and (Title:framework)) or ((Abstract:organizational memory) and (Abstract:framework)) or ((Keywords:organizational memory) and (Keywords:framework))
IEEE Xplore	(Document Title : organizational memory AND Document Title : framework ) OR (Abstract: organizational memory AND Abstract : framework ) OR (Author Keywords: organizational memory AND Author Keywords : framework ) OR (Index Terms: organizational memory AND Index Terms : framework )

La búsqueda fue realizada el 17/10/2013, retornando los siguientes resultados:

Tabla 3.5: Resultados de búsqueda

Búsqueda	Organizational memory and framework		
	Fecha de la búsqueda	Cantidad de resultados	Artículos seleccionados
Scopus	17/10/2013	87	1
Science Direct	17/10/2013	7	0
ACM DL	17/10/2013	32	0
IEEE Xplore	17/10/2013	14	0

Solamente se encontró un artículo relacionado a propuestas de arquitecturas/-frameworks conceptuales para construcción de memorias organizacionales.

### 3.2.4. Artículos seleccionados

A continuación se describe brevemente los trabajos de investigación o artículos seleccionados de la segunda búsqueda (06 artículos) y tercera búsqueda (01 artículo) de la revisión sistemática. Cabe señalar que el único artículo seleccionado de la primera búsqueda se encuentra contenido en el resultado de la segunda búsqueda, por lo cual en total se describirán 7 artículos seleccionados.

#### 3.2.4.1. Arquitectura general para memorias corporativas de (Kuhn and Abecker, 1997)

La arquitectura general para memorias corporativas propuesta por (Kuhn and Abecker, 1997) (Ver figura 3.1), fue derivada de tres (03) casos de estudio presentados en el artículo (Kuhn and Abecker, 1997), la arquitectura general se desarrolló a lo largo de muchos años. La arquitectura se usó como esquema o plantilla general para proponer otros proyectos de memoria corporativa en el dominio de mesa de ayuda y soporte al cliente.

La arquitectura de (Kuhn and Abecker, 1997), clasifica a la información de acuerdo a su complejidad y a la estructura de los datos, conocimiento y documentos, los cuales conforman el núcleo del sistema denominado repositorio de información. Los datos y los documentos usualmente ya se encuentran almacenados en bases de datos y herramientas de gestión de documentos de la organización, los cuales pueden ser recuperados por mecanismos tradicionales de recuperación de información. El conocimiento caracterizado por una representación formal, puede ser procesado con inferencias complejas (inducción, deducción, etc) y no solo por mecanismos tradicionales de recuperación de información.

Asimismo, de acuerdo al nivel de abstracción, el rol en el proceso de solución de problemas, y la estabilidad temporal de la información, se distinguen tres capas de información: información específica de casos, información general e información ontológica o metainformación. La capa de información específica de casos, describe los procesos y control de información usada/generada durante la ejecución de tareas operacionales específicas. La capa de información general, describe como usar e interpretar datos específicos, esta información puede ser usada para sugerir, explicar y criticar. La capa de información ontológica o metainformación, describe el marco conceptual para las demás partes del sistema, el cual es importante para controlar la evolución de la base de conocimientos.

Finalmente, los servicios de gestión y procesamiento de información, permiten acceder al repositorio de información a través de componentes de evolución y utilización de conocimiento. Los componentes de evolución de conocimiento, permiten

al administrador de la memoria corporativa ingresar nuevas reglas de inducción, deducción, entre otros. Los componentes de utilización de conocimiento permiten a usuarios hacer uso del conocimiento.

### 3.2.4.2. Arquitectura de memoria organizacional de (Boury-Brisset and Tourigny, 2000)

(Boury-Brisset and Tourigny, 2000) propuso una arquitectura de una memoria organizacional (*OM*) para seguridad vial (Ver figura 3.2). La arquitectura propuesta está compuesta por un sistema basado en conocimiento que explota todo el conocimiento disponible en la organización; captura y formaliza la teoría de dominio, el conocimiento del *know how* y las experiencias. La teoría de dominio está representada por ontologías y modelos de tareas, el conocimiento del *know how* está representado por la base de conocimientos, y las experiencias están basadas en casos.

El sistema está compuesto por el entorno el cual integra todo el conocimiento existente en la organización en una memoria corporativa, además formaliza la teoría de dominio y las experiencias, ofrece herramientas para la construcción y mantenimiento de la memoria corporativa e implementa tareas múltiples que explotan el contenido de la memoria corporativa. La memoria corporativa esta compuesta por diferentes tipos de conocimiento: base de datos de computadora, documentos electrónicos, ontologías de dominio, modelos de tareas declarativo, bases de conocimiento y bases de casos conteniendo soluciones de problemas.

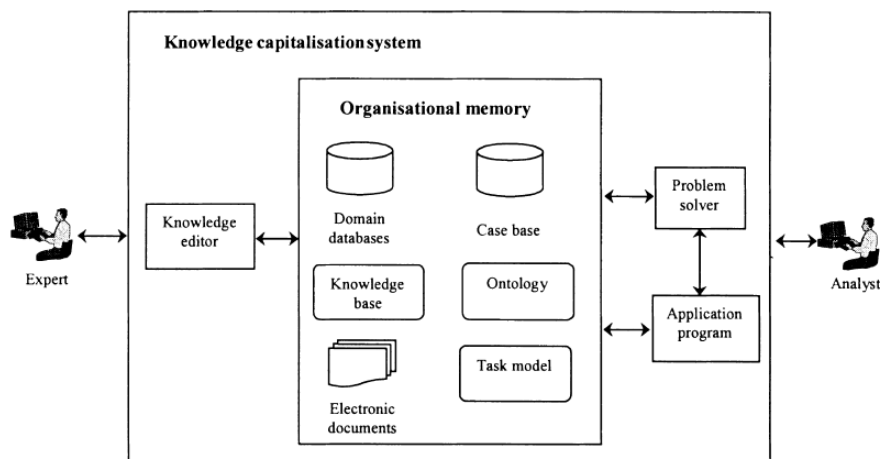


Figura 3.2: Arquitectura de una memoria organizacional para seguridad vial (Boury-Brisset and Tourigny, 2000).

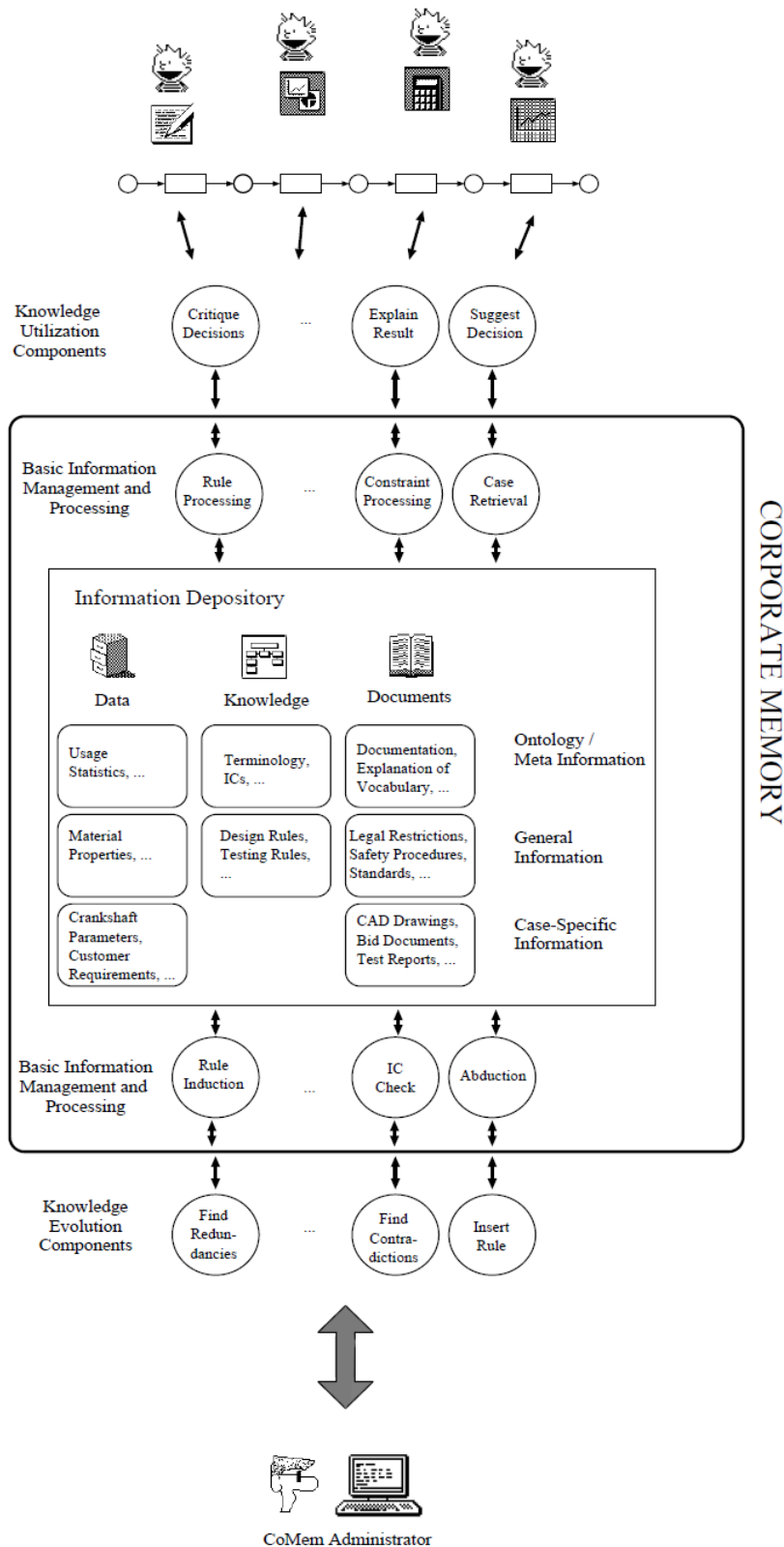


Figura 3.1: Arquitectura de referencia de Memoria Corporativa (Kuhn and Abecker, 1997).

### 3.2.4.3. Memoria organizacional basada en ontologías

(Zouaq et al., 2006) propuso una arquitectura conceptual que se enfoca en la creación y mantenimiento de una memoria organizacional basada en ontología (Ver figura 3.3). La ontología es usada para la creación y organización de metadatos del contenido, asimismo proporciona un entendimiento común de toda la organización. La ontología es usada también para describir la organización, el dominio en el cual trabaja la organización, sus competencias, los documentos de la organización, etc. Cada objeto en el sistema (Empleado, competencia, material, rol, etc.) es un objeto OWL el cual puede ser identificado únicamente por un URI. Todos los objetos son descritos por un número de conceptos de la ontología de dominio y por lo tanto pueden recuperarse mediante técnicas basadas en similitud utilizando el algoritmo de clasificación coseno vector.

La arquitectura toma como entrada datos estructurados (bases de datos, casos, empleados, competencias, etc) y datos no estructurados (documentos) de los empleados que quieren compartir su conocimiento (producción de conocimiento). El sistema identifica conocimiento y conceptos, y los enlaza a las ontologías existentes o crea nuevos conceptos en las ontologías. La identificación de conocimiento y conceptos se realiza de forma manual (con la ayuda de empleados) o semi-automática (con técnicas de recuperación de información). El administrador de metadatos (*metadata manager*) crea los metadatos en archivos OWL y referencia a los objetos OWL. Luego, un objeto de conocimiento es añadido a la memoria organizacional. Como cada objeto está definido de acuerdo a las ontologías, el sistema es capaz de ofrecer herramientas eficientes de recuperación de conocimiento.

### 3.2.4.4. Memoria organizacional de dos capas

(Ju, 2006) propuso una memoria organizacional de dos capas (Ver figura 3.4). En la primera capa solo se puede lograr la gestión del conocimiento básico, cada documento es marcado con meta-información que apoye la búsqueda y recuperación eficiente a través de un sistema de gestión de documentos. En la segunda capa, se toman medidas para utilizar las computadoras para los recursos de información, cuyo contenido es altamente estructurado o de naturaleza procesal, entre estas medidas se tiene un esquema de representación del conocimiento (*KR*) el cual es el núcleo, como parte del esquema de representación del conocimiento, se adoptaron los constructos *KR* de lógica, ontología y computación, y se eligió *Topic Maps in XML* como la tecnología de representación.

Los constructos de representación del conocimiento se operacionalizan en una ontología de información (describe la estructura y el acceso a la memoria organizacional), una ontología empresarial (describe el contexto de una organización para utilizar correctamente el conocimiento) y una ontología de dominio (describe el contenido del conocimiento).

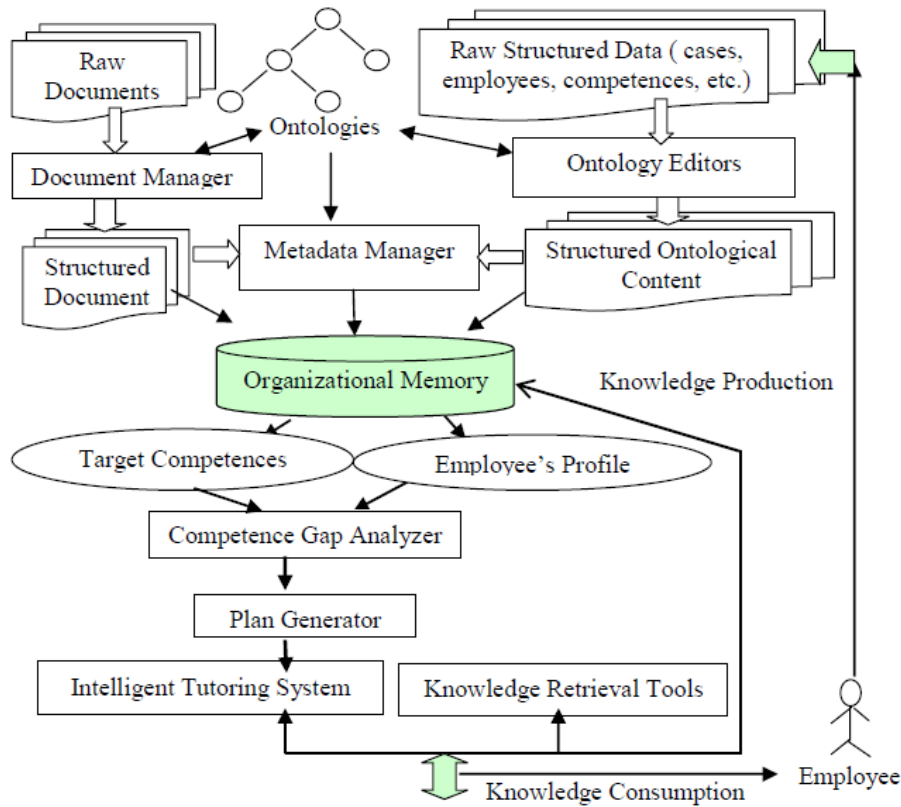


Figura 3.3: Arquitectura conceptual (Zouaq et al., 2006).

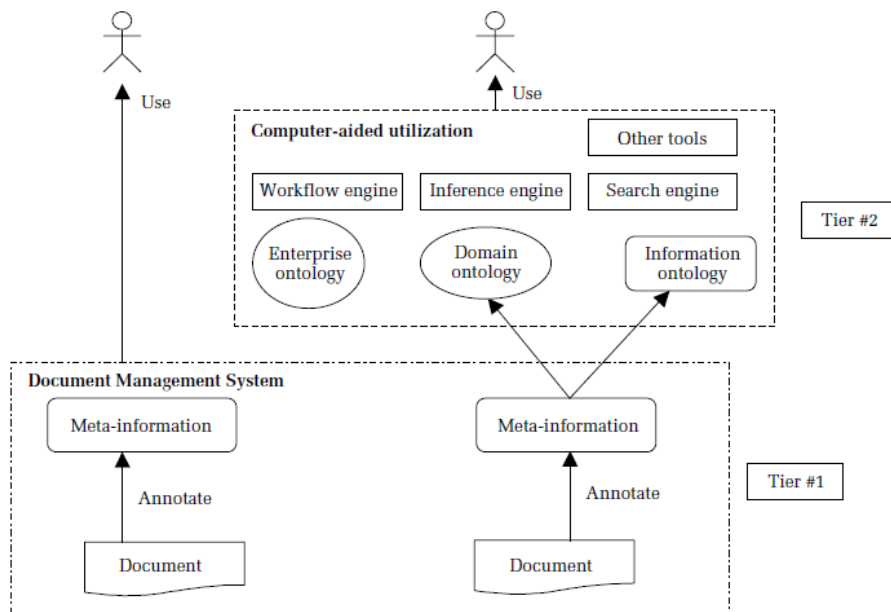


Figura 3.4: Memoria organizacional de dos capas (Ju, 2006).



### 3.2.4.5. Arquitectura *Onto-DOM*

(Ale et al., 2008) define la arquitectura *Onto-DOM* para representar de forma homogénea los conocimientos que son de naturaleza heterogénea, dicha arquitectura tiene 3 componentes principales:

**Capa de recuperación y procesamiento de información:** Es responsable del análisis de consultas del usuario, transformación de las consultas en un formato adecuado y de la recuperación de la información.

**Capa de representación del conocimiento:** Este componente es responsable de la extracción y representación de conocimiento procedente de diversas fuentes heterogéneas. (Ale et al., 2008) propuso una estrategia para la representación de documentos semánticos donde las ontologías son usadas como la estructura principal para el proceso de clasificación, las ontologías de dominio contienen todos los conceptos relevantes y relaciones en un dominio específico.

**Interfaz de otros dominios:** Es responsable de propagar las consultas de usuario a memorias organizacionales en dominios diferentes que puedan proveer una respuesta.

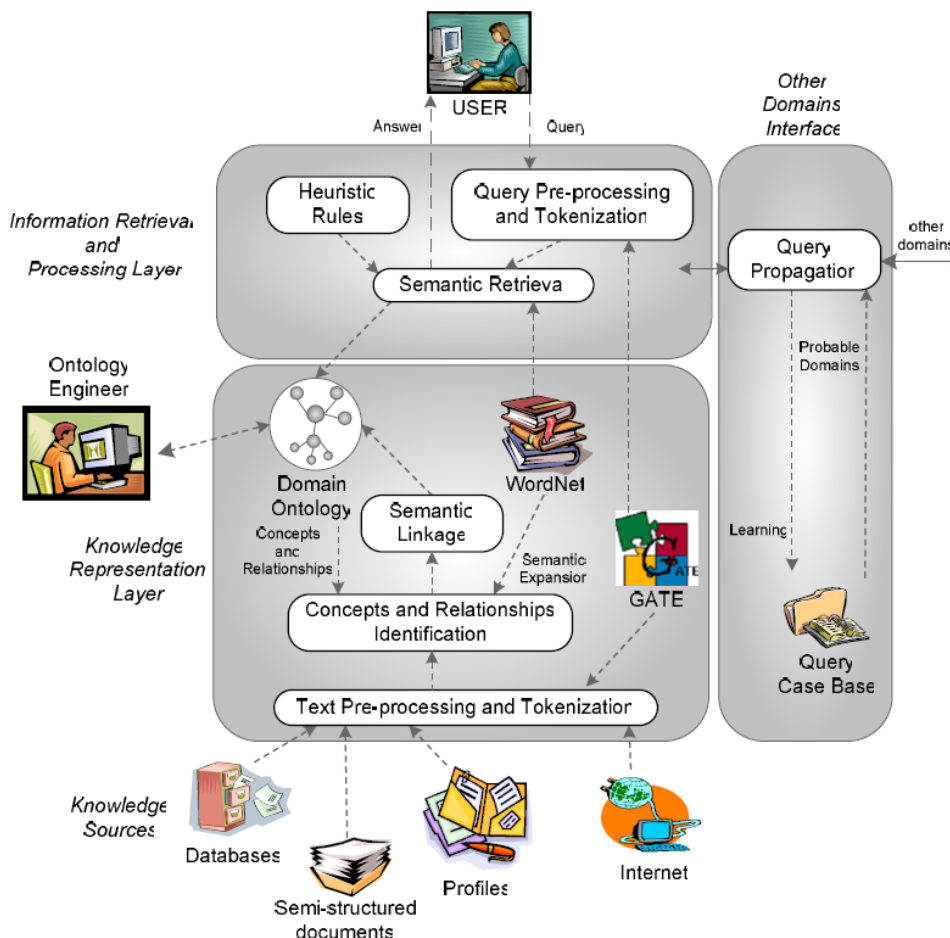


Figura 3.5: La arquitectura *Onto-Dom* (Ale et al., 2008).



### 3.2.4.6. Arquitectura de un sistema de gestión del conocimiento para una universidad

(Adil Laoufi, 2011) presenta una arquitectura de un sistema de gestión del conocimiento para una universidad (*Université IBN ZOHR Agadir - Morocco*) (Ver figura 3.6), el cual es capaz de integrar diferentes ontologías, y permite el razonamiento y recuperación de información inteligente. La arquitectura del sistema está compuesta por 5 niveles: exploración, capitalización, gestión, recuperación y base de conocimientos. La capa de base de conocimiento está construida en base a 04 ontologías: la ontología de perfil de usuario, la ontología de dominio, la ontología de contexto y la ontología de información.

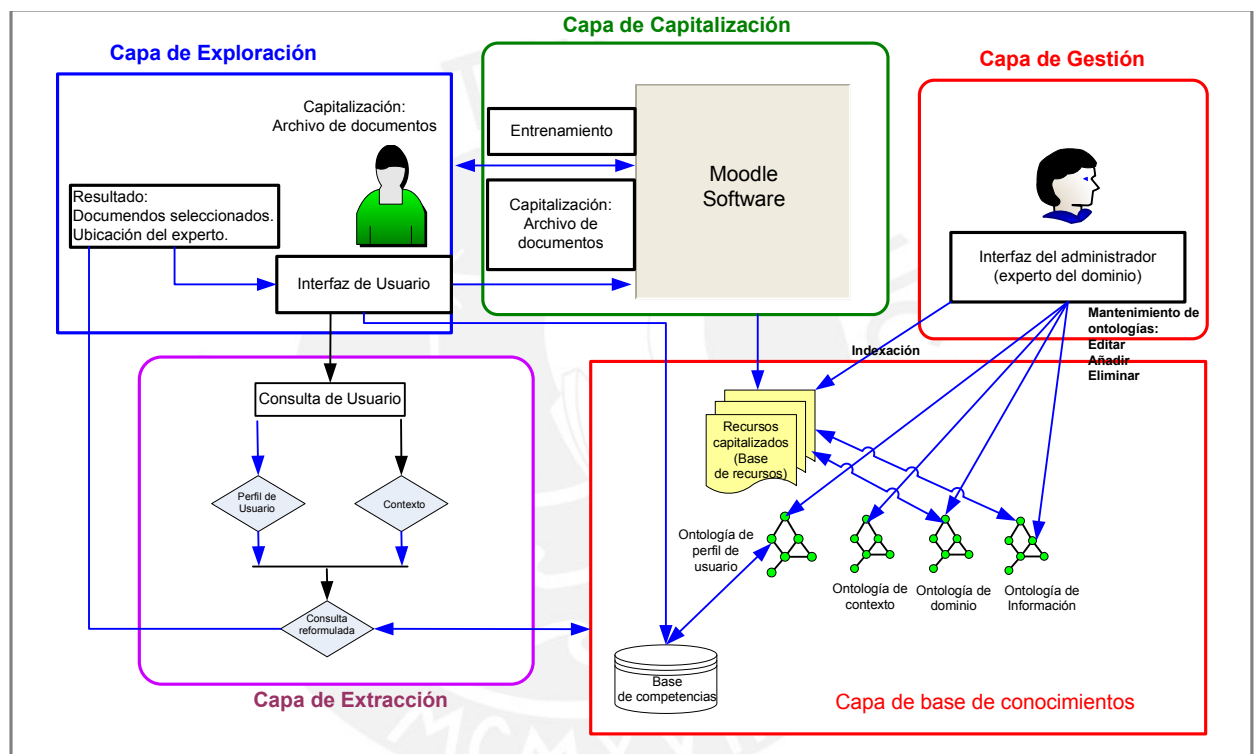


Figura 3.6: Arquitectura de un sistema de gestión del conocimiento para una universidad (Adil Laoufi, 2011) (Adaptado al español).

### 3.2.4.7. Arquitectura de memoria organizacional de (Toledo et al., 2011)

(Toledo et al., 2011) propuso una arquitectura de memoria organizacional, en esta arquitectura las organizaciones están compuestas por varias unidades funcionales (áreas, departamentos, grupos, etc.). Cada unidad funcional es un dominio de conocimiento que cumple con los principios de autonomía y coordinación permitiendo un enfoque de gestión del conocimiento basado en una arquitectura distribuida, cada

dominio de conocimiento implementa su propio sistema de gestión de conocimiento guiado por ontologías, este sistema de gestión del conocimiento está compuesto por 02 módulos principales: representación y recuperación del conocimiento, se tiene los módulos de comunicación entre dominios y la evolución de la ontología. El módulo de representación del conocimiento recupera el conocimiento de diversas fuentes como: documentos, base de datos, hojas de cálculo, memorandos, etc., y los almacena con un formato adecuado; el módulo de recuperación del conocimiento del sistema transforma la consulta del usuario eliminando ambigüedades y obtiene los conceptos más representativos dentro del dominio; el módulo de comunicación entre dominios permite propagar una consulta de usuario a otros dominios, en el caso de no obtener respuestas adecuadas en el dominio local; el módulo de evolución de la ontología es responsable por mantener actualizada la ontología de dominio.

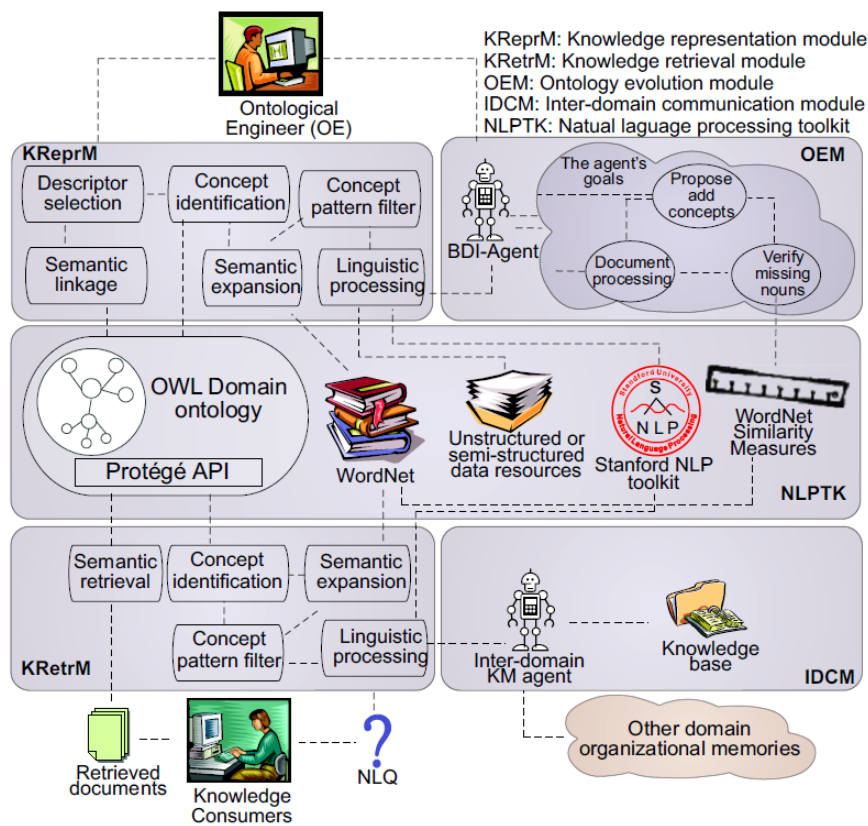


Figura 3.7: Propuesta de arquitectura conceptual para sistemas de memoria organizacional (Toledo et al., 2011).

### 3.3. Resumen de la Revisión Sistemática

En la presente sección se revisó el estudio del arte sobre arquitecturas de memoria organizacional para construcción de sistemas de memoria organizacional. La primera búsqueda enfocada en investigaciones en instituciones de educación superior, solo brindó un artículo relevante pese a tener hasta 55 resultados en el indexador SCOPUS. Por lo cual se realizaron 2 búsquedas adicionales ampliando las opciones de búsqueda de arquitecturas de memoria organizacional sin importar el rubro de negocio de la organización, obteniendo en total 7 artículos seleccionados.

Todas las arquitecturas de sistemas de memoria organizacional de los artículos seleccionados están basadas en ontologías, dado que permite agregar contenido semántico a la información y/o documentos, facilitando la construcción de bases de conocimientos. Cabe señalar que algunos artículos centran su arquitectura en el uso de ontologías para la construcción de una base de conocimientos. Mientras otros autores consideran a las ontologías como otro componente de su arquitectura, las cuales están basadas en base de casos, base de conocimientos, el propio sistema de gestión de conocimiento, entre otros.

Asimismo, los artículos seleccionados, precisan que es posible agregar contenido semántico a los documentos de las organizaciones, permitiendo la recuperación de información basada en inferencias y no solo por el contenido almacenado en los campos de una base de datos relacional. Dicho contenido semántico se logra con anotaciones semánticas a los documentos y enlazándolos a las ontologías que se hayan definido.

Por otro lado, en las arquitecturas propuestas se tiene como componente importante a los usuarios que harán uso del sistema de memoria organizacional, los cuales interactúan con diversos componentes del sistema que retornan el conocimiento almacenado. Otros componentes importantes, son el administrador de ontologías (base de conocimientos) y el experto del dominio (base de casos).

## Capítulo 4

### Resultados Esperados

#### 4.1. Describir los procesos de negocios típicos en las Instituciones de Educación Superior, que requieran tareas intensivas en conocimiento

De la primera búsqueda, solo 4 artículos están relacionados a memoria organizacional en instituciones de educación superior, en estos artículos no se describen las tareas intensivas en conocimiento de una institución de educación superior. Por lo que se realizó una búsqueda en los indexadores académicos IEEE, SCOPUS, ACM DL y ScienceDirect, se usó la siguiente cadena de búsqueda, con ciertas particularidades para cada indexador:

- (title-abs-key(“knowledge intensive tasks”) and title-abs-key(“university”)) or.
- (title-abs-key(“knowledge intensive tasks”) and title-abs-key(“education”)).

La búsqueda se realizó el 17/10/2013, no se encontraron investigaciones donde se describan tareas intensivas en conocimiento en universidades (Ver tabla 4.1), por lo cual las actividades intensivas en conocimiento se determinaron en base a su definición y a las actividades típicas que se realizan en una universidad.

Las principales actividades que realizan los docentes y estudiantes en una universidad son la búsqueda y creación de conocimiento (investigación), esta búsqueda les permite escribir artículos científicos, crear nuevo conocimiento, entre otros (Adil Laoufi, 2011). Asimismo los docentes y estudiantes realizan otras actividades rutinarias o administrativas en una universidad. Un estudiante puede encontrar los horarios de los cursos, reservar un libro en la biblioteca, matricularse en un curso,

Tabla 4.1: Resultados de búsqueda

Búsqueda	Tareas intensivas en conocimiento		
	Fecha de la búsqueda	Cantidad de resultados	Artículos relacionados a tareas intensivas en conocimiento
Scopus	17/10/2013	5	0
Science Direct	17/10/2013	1	0
ACM DL	17/10/2013	1	0
IEEE Xplore	17/10/2013	1	0

preparar una presentación, entre otros, mientras un docente revisa exámenes, registra asistencia a los cursos, prepara presentaciones sobre lo investigado, entre otros (Andreas Rath and Lindstaedt, 2013). Para que una actividad sea considerada como intensiva en conocimiento, las personas que desarrollan las actividades están sometidas a cierto grado de incertidumbre y toman decisiones en base a su experiencia, formación, competencias y juicios personales (Heravizadeh and Edmond, 2008). De todas las actividades que se realizan en una universidad, la investigación y creación de nuevo conocimiento tienen mayor grado de incertidumbre, no son actividades rutinarias y requieren de mucho tiempo para su realización, por ejemplo muchos de los trabajos de investigación de doctorados demoran hasta 4 años para su culminación, y los autores de dichos trabajos deben leer bastante información sobre la rama de investigación del proyecto que realizan.

Asimismo, se revisó 05 páginas web de las 10 mejores universidades según <http://www.topuniversities.com> dichas universidades tienen como misión la investigación y la enseñanza de calidad (transferencia de conocimiento)<sup>1 2 3 4 5</sup>. Por lo cual la investigación, es un actividad muy importante en una universidad.

Por lo antes indicado, la investigación es un proceso que requiere de actividades intensivas en conocimiento.

<sup>1</sup><http://web.mit.edu/aboutmit/>

<sup>2</sup><http://www.harvard.edu/about-harvard>

<sup>3</sup><http://www.stanford.edu/about/>

<sup>4</sup><http://www.cam.ac.uk/about-the-university/how-the-university-and-colleges-work/the-universitys-mission-and-core-values>

<sup>5</sup><http://www.caltech.edu/content/mission-statement>

## 4.2. Métodos y técnicas de la Ingeniería de Conocimiento que pueden ser usados en los procesos de memoria organizacional

Proporcionar una memoria organizacional efectiva no es un problema que la informática puede resolver solo (Atwood, 2002), para ello es necesario una solución híbrida que implica tanto a personas, procesos y tecnología (Abecker et al., 2001). En este sentido la ingeniería del conocimiento apoya la adquisición, organización y recuperación del conocimiento haciéndolo accesible (ver figura 4.1). La eficacia de la memoria organizacional no sólo depende de la cantidad de conocimiento almacenado, sino también de la facilidad con la que se puede acceder (Sinkula, 1994).

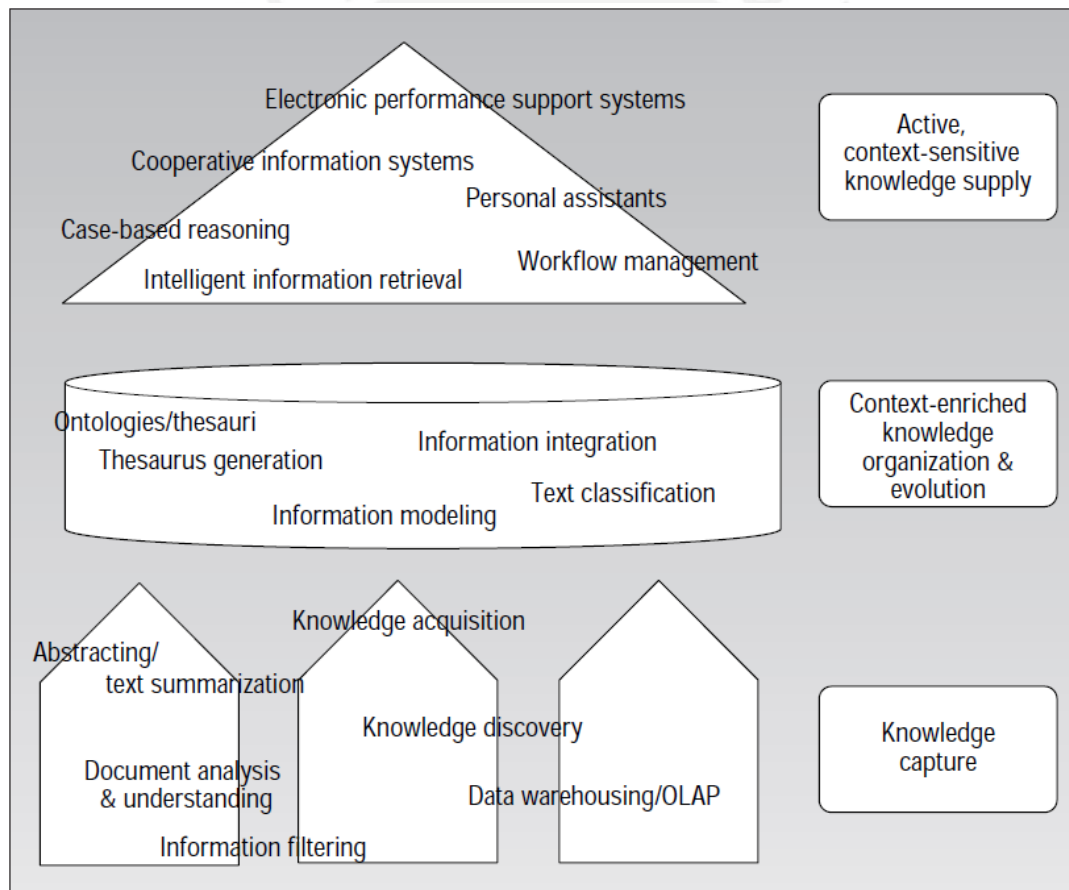


Figura 4.1: Infraestructura de Información como apoyo a la construcción de sistemas de memoria organizacional (Abecker et al., 1998)



### 4.2.1. La Ingeniería del Conocimiento como apoyo a la adquisición de conocimiento

La ingeniería del conocimiento trabaja con una variedad de métodos y técnicas que se pueden utilizar para apoyar el proceso de adquisición del conocimiento. Entre ellas se tiene la extracción de información, resumen de texto, datawarehouse y procesos de descubrimiento de conocimiento.

#### 4.2.1.1. Extracción de Información

Toma texto como entrada y produce datos en formato fijo y sin ambigüedades (Cunningham, 1997). El enfoque desde la ingeniería del conocimiento puede ser visto como la actividad de llenar información estructurada a partir de información no estructurada escrita en formato libre (Grishman, 1997). En este enfoque, los conceptos clave (es decir, hechos o acontecimientos relativos a la entidad o las relaciones entre entidades) se definen antes de que el sistema los busque (Gaizauskas and Wilks, 1998) ejecutando básicamente cinco tipos de tareas: reconocimiento de entidades, resolución de la correferencia, construcción de plantilla de elementos, construcción de plantillas de relaciones y producción de plantillas de escenarios (Cunningham, 1997).

El reconocimiento de entidades es la tarea más simple, se encarga de identificar todas las entidades en el texto (es decir, los nombres de personas, lugares, organizaciones, fechas, cantidades de dinero). La resolución de la correferencia identifica las relaciones de identidad existentes entre las entidades. La construcción de plantilla de elementos clasifica la entidades asociando información descriptiva. La construcción de plantillas de relaciones identifica las posibles relaciones entre las entidades. La producción de plantillas de escenarios es el producto de los sistemas de gestión de conocimiento, las entidades se unen describiendo sus elementos y relaciones dentro de un determinado evento. Actualmente los sistemas de ingeniería de conocimiento no son precisos para extraer información y construir base de datos de forma automática, pero los resultados pueden ser útiles para apoyar a los seres humanos en sus tareas, recolectando evidencias de los documentos (Ramshaw and Weischedel, 2005). El desempeño de cada tarea y la facilidad con que se puede desarrollar varía en función del tipo de texto (por ejemplo, artículos de revistas, HTML, mensajes electrónicos), dominio (las noticias financieras, información turística, asistencia técnica) y el escenario (es decir, misiles lanzados, fusiones de organizaciones) (Cunningham, 1997). Aunque los sistemas de ingeniería de conocimiento ofrecen un rendimiento modesto son de gran valor (con pérdida de algunos eventos y la inclusión de algunos errores). Pueden ser utilizados para extraer información y hacer una verificación manual posterior (Grishman, 1997).

#### 4.2.1.2. Resumen de textos

Consiste en centrarse en una versión más pequeña del texto preservando su contenido (Kruengkrai and Jaruskulchai, 2003), permitiendo el ahorro de tiempo de las personas que no pueden leer todos los documentos relacionados con un tema en

particular.

Hay dos formas de resumen de textos: abstracción y la extracción o resumen (Kruengkrai and Jaruskulchai, 2003). La abstracción es el proceso de comprender, interpretar y parafrasear a un nuevo extracto del texto original. Esto sigue siendo una tarea difícil de realizar usando las computadoras, y requiere el uso extensivo de las técnicas y recursos del área de procesamiento del lenguaje natural. El resumen o extracción es el proceso de clasificación y presentación de extractos más originales y relevantes dentro de un documento (oraciones, párrafos , etc).

#### 4.2.1.3. Datawarehouse

Según (Inmon, 1992), el datawarehouse (DW) se puede definir como un conjunto de datos integrados, variables en el tiempo no volátil que se utiliza principalmente en las organizaciones para apoyar la toma de decisiones. El datawarehouse puede ser visto con un enfoque de integración de la información, donde la información de interés se extrae de diferentes fuentes y se almacena en un repositorio centralizado donde se ejecutan las consultas del usuario (Widom, 1995). El datawarehouse apoya la toma de decisiones, la información almacenada es histórica, resumida y consolidada y no hay registros de las transacciones individuales (Chaudhuri and Dayal, 1997).

#### 4.2.1.4. Descubrimiento de información: Descubrimiento de conocimiento en bases de datos y en bases de datos textuales

El descubrimiento de conocimiento en bases de datos se puede definir como el proceso no trivial de identificación de datos, patrones que son válidos, útiles y comprensibles (Fayyad et al., 1996). Entre las fases que componen el proceso de descubrimiento de Conocimiento en Bases de Datos se puede mencionar la selección, tratamiento previo, tratamiento, la minería y la interpretación de los datos (Fayyad et al., 1996). El almacén de datos se puede utilizar como una aproximación a la selección, pre-procesamiento y proceso de transformación de datos en el descubrimiento de conocimiento en bases de datos.

El descubrimiento de conocimiento en texto (KDT) sigue la misma filosofía de descubrimiento de conocimiento en bases de datos. El elemento que distingue a los dos casos es la fuente de los datos, en el Descubrimiento de Conocimiento en Bases de datos textuales la fuentes de información es de tipo no estructurada , escrita en lenguaje natural de texto (Kodratoff, 1999). El primer paso del Descubrimiento de Conocimiento en Bases de datos textuales es convertir la información textual en información estructurado para implementar las fases restantes del descubrimiento de conocimiento en Bases de datos de información. En este punto, se puede utilizar el enfoque de extracción de información (Feldman and Dagan, 1995).

## 4.2.2. La Ingeniería del Conocimiento como apoyo a la co- dificación del Conocimiento

El modelado del conocimiento permite formalizar conocimiento independiente de su aplicación, facilitando su reutilización (Studer et al., 1999). Para representar el conocimiento se puede utilizar vocabularios controlados, taxonomías, tesauros y ontologías, a partir de repositorios de documentos a las ontologías. La principal diferencia entre estas formas de representación es el nivel semántico que permiten formalizar.

### 4.2.2.1. Vocabularios controlados

Un vocabulario controlado puede ser entendido como una lista cerrada de nombres. Los componentes de un vocabulario controlado se conocen generalmente como términos, donde un término es un nombre específico para un concepto particular. Es común distinguir entre término y concepto, el término básicamente se puede definir como el nombre de un concepto, en el otro lado, es común que el mismo concepto tenga varios nombres y también que el mismo término puede nombrar varios conceptos. Un vocabulario controlado consta de términos y no de conceptos directamente (Garshol, 2004).

### 4.2.2.2. Taxonomías

Una taxonomía puede ser visto como una organización de términos jerarquizados de un vocabulario controlado. La ventaja de este enfoque es que permite que los términos relacionados se agrupen y se clasifican haciendo más simple el proceso de búsqueda (Garshol, 2004).

### 4.2.2.3. Tesauros

Un tesauro se puede definir como una lista ordenada de los términos, sobre todo palabras claves en un dominio específico, los tesauros tratan con las palabras, las alternativas para estas palabras, sinónimos, traducciones y otros que se enriquecen con algunos tipos de relaciones definidas (Rees, 2003) como: "Término amplio" (BT) y "término específico" (NT), que pueden ser leídos en orden jerárquico, y "término relacionado" (RT) que se utiliza para capturar la proximidad conceptual (Hepp and Bruijn, 2007).

### 4.2.2.4. Ontologías

Las ontologías son esquemas de metadatos que proporcionan un vocabulario controlado de términos, todos definidos explícita y semánticamente que puedan ser procesados por las computadoras (Bechhofer, 2002). Una ontología puede tomar una variedad de formas, necesariamente un vocabulario de términos y una especificación

de su significado. Esto incluye definiciones y una indicación de cómo formalizar los conceptos que están interrelacionados colectivamente en un dominio y las posibles restricciones a las interpretaciones de los términos (Uschold et al., 1995).

### 4.2.3. La Ingeniería del Conocimiento como apoyo al uso del conocimiento

#### 4.2.3.1. Recuperación de Información

Un sistema de recuperación de información puede ser visto como una herramienta que interpreta el contenido de los elementos de información (documentos) de una colección y clasifica de acuerdo con el grado de relevancia para la consulta del usuario (Baeza-Yates and Ribeiro-Neto, 1999). El objetivo principal es recuperar todos los documentos relevantes y al mismo tiempo recuperar la menor cantidad de documentos no relevantes (Baeza-Yates and Ribeiro-Neto, 1999), a fin de satisfacer el estado de incertidumbre del usuario (Ingwersen, 1992).

#### 4.2.3.2. Pregunta Respuesta - *Question Answering*

El objetivo del *Question Answering* (QA) es localizar, extraer y representar una respuesta específica para una pregunta realizada por un usuario usando lenguaje natural (Roussinov and Robles-Flores, 2004). Los sistemas elaborados usando las técnicas de QA son especialmente útiles en situaciones donde el usuario necesita conocer información específica y no tiene tiempo para leer toda la documentación disponible sobre el asunto (Molla and Vicedo, 2007). Para responder, el sistema debe analizar la pregunta, procurar una o más respuestas consultando las fuentes de información y presentar la respuesta al usuario de manera apropiada, añadiendo la respectiva justificación en caso de ser necesario (Hirschman and Gaizauskas, 2001).

En general los sistemas QA están basados en texto y retornan a los usuarios fragmentos de textos que contienen las respuestas a una consulta determinada. Algunos sistemas de Búsqueda de Respuestas toman como entrada el contexto en el que se hace la pregunta y el conocimiento que tiene el usuario en el campo, así como sus objetivos (Hirschman and Gaizauskas, 2001).

#### 4.2.3.3. Razonamiento basado en casos

El razonamiento basado en casos (CBR) se puede definir como un enfoque que ayuda en la solución de problemas. Un sistema de razonamiento basado en casos es capaz de utilizar el conocimiento específico de las experiencias del pasado en una situación específica. Un nuevo problema se resuelve a partir del descubrimiento de casos similares y de la utilización de estos en situaciones nuevas (Aamodt and Plaza, 1994). Básicamente, los sistemas de razonamiento basado en casos utilizan procesos para recuperar el caso más similar, reutilizando la información y el conocimiento

de ese caso para resolver el problema, revisando las partes de esta experiencia que pueden ser útiles para la futura resolución de problemas.

### 4.3. Procesos de Memoria Organizacional en el contexto de Instituciones de Educación Superior

A continuación se describen los procesos de memoria organizacional en el contexto de una institución de educación superior. Cabe señalar que la revisión del estado del arte, no permite definir claramente los procesos de memoria organizacional en el contexto de instituciones de educación superior.

#### 4.3.1. Adquisición

De la primera búsqueda del estado del arte, la adquisición de conocimiento en una institución de educación superior se realiza teniendo como origen conocimiento explícito. El presente trabajo de investigación se centra en los documentos (conocimiento explícito) que pueden ser tesis, artículos, revistas, entre otros. Para recuperar conocimiento de los documentos, los cuales en su mayoría se encuentran en formato no estructurado se podrá hacer uso de las técnicas de extracción de información y en los casos que se encuentre información en bases de datos relacionales se podrá usar el descubrimiento de información en bases de datos y bases de datos textuales. Es probable que se requiera de un experto de dominio para formalizar adecuadamente el conocimiento de estos documentos, puesto que un proceso automatizado podrá encontrar muchas ambigüedades.

#### 4.3.2. Codificación

Para codificar o estructurar el conocimiento en una institución de educación se puede hacer uso de las diferentes ontologías que se han desarrollado para diferentes dominios tal como se describe en (Meenachi and Baba, 2012). En un universidad se tiene que usar ontologías del ámbito académico, como las desarrolladas para el área de ciencias de la computación (Cassel et al., 2007).

#### 4.3.3. Recuperación

Por las características del contenido de conocimiento en documentos que se utilizan en una institución de educación superior, la técnica de recuperación del cono-



cimiento que más se adecúa es la recuperación de información, puesto que permite recuperar los documentos de investigación relevantes para los docentes y estudiantes.



# Capítulo 5

## Modelo Propuesto

### 5.1. Modelo de Organización

Esta sección describe los problemas y oportunidades desde una perspectiva amplia de negocio de las instituciones de educación superior. Las oportunidades y problemas de la organización son analizados de acuerdo a la plantilla OM-1 de la metodología CommonKADS (Schreiber and Akkermans, 1999) (Tabla 5.1).

Se encontraron dos (02) problemas principales: la memoria organizacional en una universidad reside en diferentes retenedores (estudiantes, docentes, artículos, revistas, tesis, recursos de aprendizaje entre otros), y es difícil encontrar la información adecuada cuando se realiza investigaciones; y el otro problema es que la mayoría de los recursos de investigación de una universidad no están estructurados, dificultando las tareas de recuperación de información. Adicionalmente se identificaron dos (02) oportunidades: El uso de ontologías para estructurar los recursos de investigación y luego inferir nuevo conocimiento; la otra oportunidad es que las universidades tienen como misión principal la transferencia de conocimiento y la investigación (Adil Laoufi, 2011).

Tabla 5.1: Plantilla OM-1: Identificación de problemas y oportunidades en la educación superior.

Problemas	<p>La memoria organizacional se encuentra en diferentes fuentes (retenedores) de la organización, y es difícil encontrar la información correcta.</p> <p>Recursos de investigación no estructurados.</p>
Oportunidades	<p>El uso de ontologías para representar los recursos de investigación, y luego inferir nuevo conocimiento desde las ontologías.</p> <p>Las universidades tienen 2 objetivos primordiales: transferencia del conocimiento y la investigación.</p>
Contexto de la Organización	<p>Pocos frameworks o arquitecturas conceptuales de memoria organizacional.</p> <p>Solo se encontró un artículo que propone una arquitectura de memoria organizacional en la educación superior.</p>
Soluciones	<p>Propuesta de una arquitectura conceptual de memoria organizacional en educación superior.</p> <p>La propuesta de la arquitectura conceptual de memoria organizacional permitirá estructurar recursos de investigación y por tanto facilitará las búsquedas de recursos de investigación.</p>

## 5.2. Descripción del Modelo propuesto

En el cuadro 5.2 - Vista esquemática del modelo, se presenta la propuesta de modelo para la construcción de sistemas de memoria organizacional en instituciones de educación superior, el cual se centra en la búsqueda de recursos de investigación representados por diferentes documentos como tesis, artículos, revistas, libros, entre otros (retenedores de conocimiento) ((Bunge, 2006)). El modelo propuesto ha sido basado en el modelo CESM, el cual permite describir un sistema mediante los componentes Composición (*C*), Ambiente (*E*), Estructura (*S*) y Mecanismo (*M*) (Bunge, 2006). Cabe señalar, que en la figura 5.1 se representa gráficamente la arquitectura propuesta.

Tabla 5.2: Vista esquemática del modelo

Composición	Ambiente	Estructura	Mecanismo
Usuarios.	Necesidad de recursos de investigación.	Recursos de investigación con anotaciones semánticas.	Vista de búsqueda.
Repositorio de Documentos.			Vista de conocimiento.
Repositorio de Conocimiento.			Vista de Documentos
Visualización.			

El modelo está compuesto por usuarios (en este estudio: estudiantes y docentes), repositorio de documentos, repositorio de conocimiento y componentes de visualización. Los usuarios interactúan con el modelo a través de interacciones con las vistas de búsqueda, conocimiento y de documentos. Por tanto los usuarios podrán acceder a investigaciones previas en determinada área de investigación, las investigaciones previas representadas están compuestas por tesis, artículos, libros, etc., los cuales desde el punto de memoria organizacional son denominados retenedores de conocimiento.

El entorno se caracteriza por la necesidad que tienen los usuarios de recuperar los recursos de investigación previos y además estos deben ser los adecuados para apoyar sus investigaciones del presente. Esta necesidad es cada vez más importante en instituciones de educación superior que buscan destacar por la cantidad y calidad de las investigaciones que se realizan en sus claustros. Para recuperar las investiga-

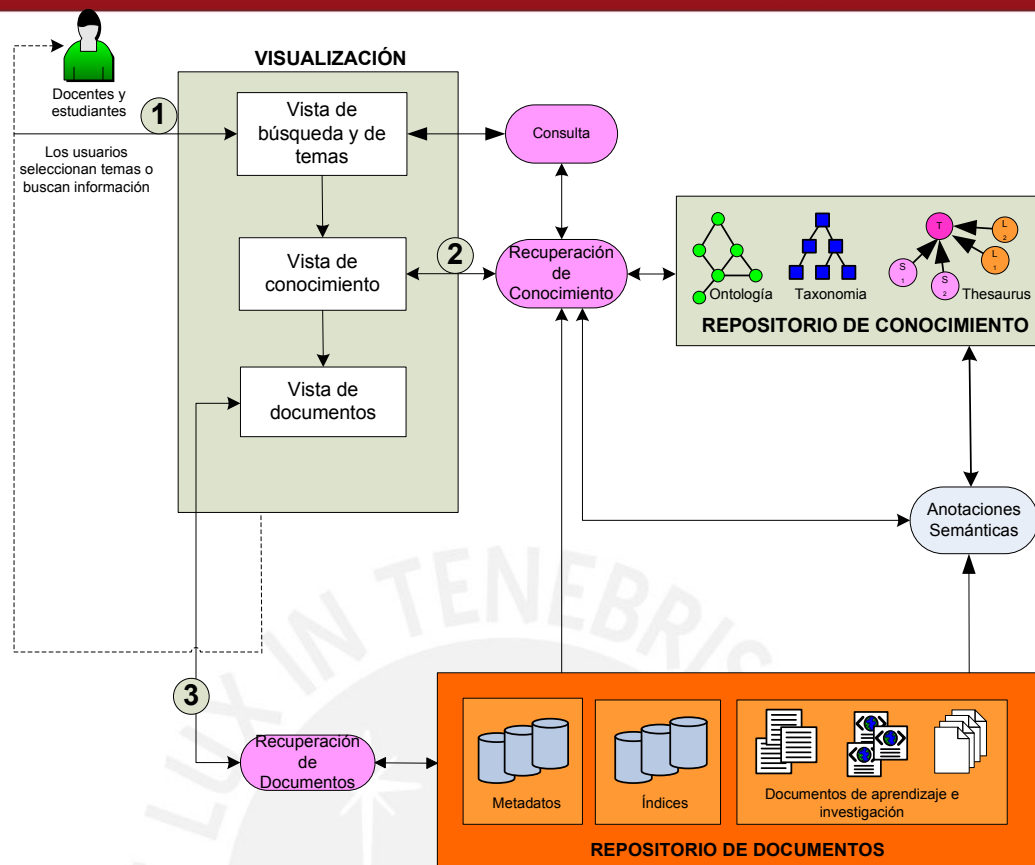


Figura 5.1: Arquitectura de software propuesta para instituciones de educación superior

ciones previas y que a su vez sean adecuadas, el repositorio de documentos debe ser enriquecido a través de anotaciones semánticas que se mapean a las ontologías.

La vista búsqueda es un mecanismo que permite buscar un recurso de investigación en base a un texto de entrada o seleccionar el recurso a través de diversos temas o áreas de investigación. La vista de conocimiento y vista de documentos son mecanismos que permiten a los usuarios recuperar un documento específico.

## 5.2.1. Composición

### 5.2.1.1. Usuarios

Los usuarios del modelo representan principalmente a los estudiantes y docentes de una institución de educación superior, quienes interactúan con el componente de visualización por la necesidad de acceder a recursos de investigación.



### 5.2.1.2. Repositorio de Documentos

Este repositorio representa a los recursos de investigación a partir del cual el conocimiento es recuperado. En el modelo propuesto se ha agregado una capa semántica, el cual permite formalizar el conocimiento contenido en los documentos (por ejemplo tesis, artículos, revistas, etc.).

El repositorio de documentos está definido por cuatro niveles: i) Nivel descriptivo, ii) Nivel de metadatos, iii) Nivel de contenido y iv) Nivel Semántico. En el nivel descriptivo tiene por objetivo identificar el documento en el repositorio, usando atributos de identificación (p.e., código, nombre, descripción, formatos de documentos soportados, etc.), atributos físicos (p.e., ubicación del repositorio) y atributos bibliográficos (p.e., nombre, autor, fecha de creación, versión, etc.). En el nivel de metadatos se busca describir la estructura de la información almacenada, por ejemplo en un repositorio de artículos académicos de ciencias de la computación los metadatos permiten describir el título del artículo, autores, palabras claves, área de investigación; en un repositorio de tesis permite describir el título, autor, universidad, facultad, asesores, entre otros. En el nivel de contenido proporciona las estructuras necesarias para almacenar los documentos y los metadatos asociados. El nivel semántico permite asociar contenido semántico a los documentos (p.e. mapeamiento a las clases e instancias definidas en una ontología).

### 5.2.1.3. Repositorio de Conocimiento

El repositorio de conocimiento es el componente que representa los artefactos usados para representar el conocimiento de dominio usado para i) realizar el mapeo semántico de los documentos, ii) realizar las inferencias cuando sea posible sobre los conceptos usados en los procesos de visualización de conocimiento.

Los artefactos para representar el conocimiento pueden ser ontologías, taxonomías o thesaurus. Por ejemplo en investigaciones en el área de ciencias de computación se puede usar la Ontología de Computación de (Cassel et al., 2007), además existen ontologías para otras ramas del conocimiento como agricultura, aviación, biología, química, ingeniería civil, ciencias de la computación entre otros (Meenachi and Baba, 2012). Las inferencias se pueden realizar a través de relaciones que estén definidas en la ontología, por ejemplo los algoritmos paralelos son parte de la teoría de diseño y análisis de algoritmos, la teoría de diseño y análisis de algoritmos son de parte de la teoría de computación, por tanto los algoritmos paralelos son parte de la teoría de computación.

### 5.2.1.4. Visualización

El componente de visualización es el encargado de presentar visualmente los recursos de investigación hacia los usuarios que requieren dicha información.

### 5.2.2. Ambiente

Los usuarios (docentes y estudiantes) requieren documentos de recursos de investigación para satisfacer los requerimientos de información. Estos requerimientos de información son importantes en las instituciones de educación superior dado que la transferencia de conocimiento y la investigación son sus objetivos primordiales, asimismo las universidades requieren destacar por la cantidad y calidad de sus investigaciones realizadas en sus claustros.

### 5.2.3. Estructura

Para recuperar adecuadamente los documentos (recursos de investigación), se debe enriquecer el repositorio de documentos con anotaciones semánticas hacia ontologías del repositorio de conocimientos. Por ejemplo, en el caso de un programa académico de ciencias de la computación se puede usar la Ontología de Computación de (Cassel et al., 2007) para enriquecer el libro de (Zelle, 2002). En la figura 5.2, se puede apreciar que el documento contiene conocimiento relacionado a algoritmos de búsquedas lineal y binaria, estas búsquedas son ejemplos de algoritmos importantes, finalmente los algoritmos importantes forman parte de la teoría de algoritmos y complejidad, por lo que se puede inferir que las búsquedas lineales y binarias forman parte de la teoría de la teoría de algoritmos y complejidad.

#### 13.1.4 Comparing Algorithms

So far, we have developed two solutions to our **simple searching** problem. Which one is better? Well, that depends on what exactly we mean by better. The **linear search algorithm** is much easier to understand and implement. On the other hand, we expect that the **binary search** is more efficient, because it doesn't have to look at every value in the list. Intuitively, then, we might expect the **linear search** to be a better choice for small lists and **binary search** a better choice for larger lists. How could we actually confirm such intuitions?

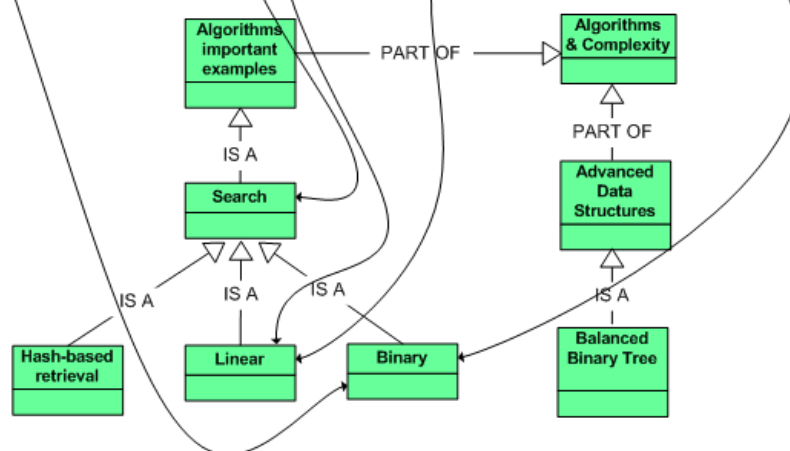


Figura 5.2: Ejemplo de mapeo semántico de un documento - Libro de (Zelle, 2002)

## 5.2.4. Mecanismo

### 5.2.4.1. Vista de búsqueda

El proceso de vista de Búsqueda inicia cuando un usuario (docentes y estudiantes) requieren recursos de investigación, el usuario ingresa una cadena de texto, a través del cual se debe obtener información relevante para el usuario. Asimismo, en el caso de artículos académicos, la vista de búsqueda permitirá buscar documentos por diversos criterios de clasificación como área de investigación, autor, fecha de publicación, entre otros.

### 5.2.4.2. Vista de conocimiento

Una vez ingresada la cadena de texto para recuperación de información relacionada, se debe presentar al usuario un listado de documentos junto con el conocimiento asociado (conceptos de la ontología). Por ejemplo en el caso de artículos académicos de ciencias de la computación, si el usuario ingresa como cadena de texto “Análisis semántico”, el sistema retornará artículos académicos con conocimiento de conceptos o clases de compiladores, web semántica, procesamiento de lenguaje natural, luego el usuario podrá seleccionar el área de investigación de su interés. Si el usuario está interesado en la construcción de compiladores, deberá seleccionar los documentos con conocimiento de conceptos o clases asociados a la construcción de compiladores.

### 5.2.4.3. Vista de documentos

En este proceso, el usuario recupera un documento específico seleccionado en la vista de conocimiento. Por ejemplo si el usuario seleccionó un documento con conocimiento de creación de compiladores, esta vista presentará al usuario el artículo académico.

## 5.2.5. Validación del Modelo Propuesto a través de comparación con otras arquitecturas conceptuales de la revisión de literatura

### 5.2.5.1. Modelo Propuesto - Arquitectura de referencia de Memoria Corporativa (Kuhn and Abecker, 1997)

En la Arquitectura de referencia de Memoria Corporativa de (Kuhn and Abecker, 1997), se distingue como componentes de información a Datos, Conocimiento y Documentos, no distinguiendo claramente la diferencia entre Datos, Información y Conocimiento, en el modelo propuesto se incluye un repositorio de Conocimiento y

de Documentos los cuales se complementan y no conforman una única de capa de información, se encuentran claramente separados.

Si bien en el modelo (Kuhn and Abecker, 1997), se precisa que se hace uso de ontologías/meta información, información general e información específica de casos; no se precisa como se formaliza el conocimiento. En el modelo propuesto claramente se define que se debe hacer uso de ontologías para formalizar el conocimiento y se debe realizar mapeos semánticos para relacionar el conocimiento del repositorio de conocimiento al repositorio de documentos. Cabe precisar que en el modelo (Kuhn and Abecker, 1997) y el modelo propuesto, definen un usuario que requiere información o conocimiento sobre un tema en particular. Asimismo en el modelo propuesto se define claramente que está aplicado a Instituciones de Educación Superior y el modelo (Kuhn and Abecker, 1997).

En el modelo propuesto (Kuhn and Abecker, 1997) propone una serie de interfaces para que los usuarios puedan interactuar con la memoria corporativa, al igual que el usuario del modelo propuesto dispone de Vistas de Búsqueda, de conocimiento y de documentos.

En el modelo (Kuhn and Abecker, 1997), definen un Administrador de la Memoria Corporativa el cual debe actualizar el conocimiento actualizado. En el modelo propuesto no se define un actor que pueda actualizar el conocimiento, ontologías, documentos, etc., puesto que el modelo propuesto tiene como objetivo representar el conocimiento, y no tiene como finalidad definir como se actualizará el conocimiento.

#### **5.2.5.2. Modelo Propuesto - Arquitectura de una Memoria Organizacional para Seguridad Vial (Boury-Brisset and Tourigny, 2000)**

En la propuesta de (Boury-Brisset and Tourigny, 2000), la arquitectura integra todos los activos de conocimiento como bases de datos, documentos electrónicos, ontologías de dominio, modelos de tareas declarativas, bases de conocimientos, bases de casos que contienen soluciones a problemas específicos. Por otro lado el presente trabajo de investigación se centra en un repositorio de documentos con contenido semántico que se mapean a ontologías.

Tanto en el modelo propuesto en el presente trabajo de investigación como en el modelo de (Boury-Brisset and Tourigny, 2000), se tiene a usuarios que usan los sistemas y/o repositorios de conocimientos, en este trabajo los usuarios son docentes y estudiantes, y en la propuesta de (Boury-Brisset and Tourigny, 2000) los usuarios son los analistas que consultan soluciones a problemas y también usan el repositorio como entorno de aprendizaje. En el presente trabajo de investigación los usuarios (docentes y estudiantes) tienen necesidad de recuperar recursos de investigación relevantes, y en el modelo propuesto de (Boury-Brisset and Tourigny, 2000) se cubrirá las necesidades de entrenamiento y la solución a problemas específicos de los analistas de seguridad vial.

En la propuesta de (Boury-Brisset and Tourigny, 2000), se ofrece diversas herramientas para construir y actualizar la memoria organizacional, con lo cual los expertos del dominio pueden actualizar y enriquecer el repositorio con nuevas teorías y experiencias. Mientras en la propuesta del presente trabajo de investigación no se ha considerado a los actores ni la forma en que se actualiza el conocimiento, tal como se indicó en la comparación anterior, el presente trabajo se centra en diseñar un modelo para representación de conocimiento.

Tanto en el modelo propuesto como en el modelo de (Boury-Brisset and Tourigny, 2000), se ofrecen interfaces para los usuarios finales, que en nuestro trabajo son diversas vistas, y en la propuesta de (Boury-Brisset and Tourigny, 2000) se presenta un programa de aplicación (solucionador de problemas).

En la propuesta de (Boury-Brisset and Tourigny, 2000), indica que la arquitectura propuesta integra todos los activos de conocimiento de la organización, sin embargo no se precisa como están integrados estos activos de conocimiento. Si los activos de conocimientos no están integrados, en realidad se tendría varios módulos de memoria organizacional. En el modelo propuesto, los documentos se relacionan con el repositorio de conocimiento a través de mapeos semánticos.

### **5.2.5.3. Modelo Propuesto - Arquitectura conceptual gestión de memoria organizacional basada en una ontología (Zouaq et al., 2006)**

En la arquitectura de (Zouaq et al., 2006), los objetos son representados por ontologías de dominio. Tanto en la propuesta de (Zouaq et al., 2006) y la arquitectura del presente trabajo de investigación, se define claramente un repositorio de documentos, dicho repositorio se mapea a ontologías existentes. En la arquitectura del presente trabajo, se define un repositorio de conocimiento, el cual está compuesto por ontologías, thesaurus y taxonomías, y en la arquitectura de (Zouaq et al., 2006) dicho repositorio es el componente Memoria Organizacional (está compuesto por objetos de conocimientos).

La arquitectura de (Zouaq et al., 2006), presente una serie de herramientas de recuperación de conocimiento para el uso de los empleados y en la arquitectura propuesta se tiene una serie de vistas para buscar los recursos de investigación. La arquitectura de (Zouaq et al., 2006), presenta varias similitudes con la arquitectura propuesta, dado que se están enfocando en la representación del conocimiento. Sin embargo la arquitectura de (Zouaq et al., 2006) se distingue por el énfasis en la captura de conocimiento y como esta se agrega a la memoria organizacional, mientras que en la arquitectura propuesta, tal como se indicó previamente, no tiene como finalidad definir como se actualizará el conocimiento.



#### 5.2.5.4. Modelo Propuesto - Arquitectura de memoria organizacional de dos capas (Ju, 2006)

La arquitectura de (Ju, 2006) dos capas, define claramente un repositorio de documentos (capa 1) y un repositorio de conocimiento (capa 2). Asimismo los documentos están mapeados a las ontologías de información y de dominio, al igual que la arquitectura propuesta, donde los documentos (recursos de investigación) están mapeados al repositorio de conocimiento a través de mapeos semánticos. Por lo que se puede observar muchas similitudes entre la arquitectura propuesta y la arquitectura de (Ju, 2006), principalmente el repositorio de documentos que se mapea al conocimiento (ontologías), en el artículo de (Ju, 2006) el repositorio de documentos se denomina sistema de gestión de documento.

Si bien en la arquitectura de (Ju, 2006), hay usuarios que interactúan con el conocimiento almacenado, no se precisa claramente a través de qué herramientas y/o módulos interactúan con la capa 2 (repositorio de conocimiento).

#### 5.2.5.5. Modelo Propuesto - Arquitectura de para representar conocimiento de naturaleza heterogénea Onto-DOM (Ale et al., 2008)

En la arquitectura de (Ale et al., 2008), los usuarios presentan la necesidad de recuperar información relevante y esta no se debe de perder en el proceso de búsqueda, con la cual presenta una necesidad de usuario similar al presente trabajo de investigación.

La arquitectura Onto-DOM propuesta por (Ale et al., 2008), se divide en 3 capas: Capa de Recuperación y procesamiento de información, capa de representación del conocimiento y Capa de Interfaz de otros dominios. Si bien el procesamiento del análisis léxico (*tokenization*) y por consiguiente la carga de información a la capa de representación de conocimiento se lleva de forma automática, en la arquitectura propuesta en el presente trabajo de investigación se hace énfasis en la representación de conocimiento y no en la forma como se actualiza el conocimiento.

La capa de recuperación y procesamiento de información de (Ale et al., 2008), tiene un propósito similar al componente de Visualización propuesto en el presente trabajo. Sin embargo las reglas heurísticas deberían ser parte de la capa de representación del conocimiento. La capa de representación del conocimiento está compuesta principalmente por una ontología de dominio, conceptos y relaciones, diccionarios; si bien se puede cargar información al repositorio de conocimiento a través de documentos semi estructurados, bases de datos, perfiles e internet, no existe un enlace entre la fuente y el repositorio de conocimiento que permita recuperar el documento o contenido original; y como parte de este trabajo se realizó un mapeo semántico que enlaza el repositorio de conocimiento con el repositorio de documentos que permite recuperar un documento específico.

Además la arquitectura Onto-DOM de (Ale et al., 2008) cuenta con una capa de

Interfaz a otros dominios mientras en el presente trabajo no se cuenta con ninguna interfaz similar, puesto que el trabajo de investigación está limitado a una necesidad particular: recuperación de recursos de investigación en una institución de educación superior.

#### **5.2.5.6. Modelo Propuesto - Arquitectura de un sistema de gestión del conocimiento para una universidad (Adil Laoufi, 2011)**

La arquitectura de (Adil Laoufi, 2011), hace énfasis en la capacitación de personal y captura de conocimiento, asimismo sobre los mecanismos de ingreso y actualización del conocimiento a diferencia de la arquitectura propuesta que se focaliza en la representación del conocimiento. La arquitectura de (Adil Laoufi, 2011), presenta un capa de exploración que tiene por objetivo personalizar el conocimiento a ser presentado a los diferentes actores como docentes, administrativos y estudiantes; y con diferente nivel de detalle, sin embargo la arquitectura propuesta no tiene por objetivo recuperar información para docentes y estudiantes con algún tipo de detalle o nivel de acceso por tipo de actor, toda vez que se busca representar el conocimiento de los recursos de investigación.

La base de conocimientos de (Adil Laoufi, 2011), tiene una base de conocimientos el cual está compuesto por 04 ontologías: Ontología de perfil, ontología de dominio MUS, ontología de contexto y ontología de información; sin embargo para las tareas intensivas de conocimientos: transferencia de conocimiento e investigación solo se ha considerado una ontología de dominio.

#### **5.2.5.7. Modelo Propuesto - Arquitectura conceptual para sistemas de memoria organizacional (Toledo et al., 2011)**

La arquitectura de (Toledo et al., 2011), está basada en la arquitectura de (Ale et al., 2008) y se divide en 5 módulos: módulo de representación de conocimiento, módulo de recuperación del conocimiento, módulo de evolución de ontología, módulo de comunicación entre dominios y herramientas de procesamiento de lenguaje natural. El módulo de representación del conocimiento tiene como objetivo obtener conocimiento de diferentes fuentes como documentos, bases de datos, entre otros; asimismo se encarga de almacenarlo formateado para que pueda ser recuperado de forma adecuada. Además cuenta con un ingeniero de ontología que se encarga de realizar el mapeo semántico entre todos los documentos fuentes y los conceptos representados en una ontología. En la arquitectura propuesta del presente trabajo, como se ha indicado antes, se centra en la representación del conocimiento y no en como se cargará la información al repositorio, ni como se actualizará la información, ni quién estará a cargo de actualizar las ontologías.

Por otro lado, en la arquitectura de (Toledo et al., 2011), la capa de recuperación de conocimiento tiene una finalidad similar al componente de visualización de la arquitectura propuesta, sin embargo el procesamiento de lenguaje natural no

debería estar en una capa de presentación al usuario, y debería estar en el módulo de representación del conocimiento. La arquitectura de (Toledo et al., 2011) cuenta con una ontología de dominio que es similar a la definida al repositorio de conocimiento del presente trabajo, (Toledo et al., 2011) definió interfaces a otros dominios, sin embargo en la arquitectura propuesta esta interfaz no es necesaria puesto que ya se está trabajando en un dominio específico y focalizado en representación de recursos de investigación.



## Capítulo 6

# Conclusiones y Trabajos Futuros

El presente trabajo de investigación permitió definir una propuesta de arquitectura para sistemas de memoria organizacional en instituciones de educación superior. Para apoyar la definición de la arquitectura se revisó el estado del arte, se detallaron las técnicas y métodos de la ingeniería del conocimiento para los procesos de memoria organizacional, finalmente se usó el modelo CESM y de la metodología CommonKADS para estructurar la arquitectura propuesta.

En la revisión de literatura se puede encontrar numerosos estudios sobre gestión del conocimiento y memoria organizacional, asimismo se han propuesto e implementado muchos sistemas de gestión del conocimiento, sin embargo pocos están relacionados a arquitecturas de software conceptuales para construcción de sistemas de memoria organizacional en instituciones de educación superior. La misión principal de una institución de educación superior es la transferencia de conocimiento e investigación y debería ser un lugar donde se tenga como prioridad la gestión del conocimiento, sin embargo hay pocos estudios de memoria organizacional en una institución de educación superior que requiere instituciones de educación superior o universidades.

De la revisión del estado del arte se puede concluir que a través de una revisión sistemática no se puede definir los procesos de una institución de educación superior, que requieran actividades intensivas en conocimiento. Por lo cual las actividades intensivas en conocimiento se determinaron en base a su definición y a las actividades típicas que se realizan en una universidad, dando como resultado que la investigación es un proceso de una institución de educación superior que requiere actividades intensivas en conocimiento.

Se detallaron diversos métodos y técnicas de la Ingeniería del Conocimiento que pueden ser usados en los procesos de Memoria Organizacional. Primero se identificaron los procesos de una memoria organizacional (adquisición, codificación y recuperación del conocimiento). Luego se describieron los métodos y técnicas de la ingeniería del conocimiento por cada proceso. Para la adquisición del conoci-

miento se tiene la extracción de información, resumen de textos, datawarehouse y descubrimiento de conocimiento en bases de datos y bases de datos textuales. Para la codificación de conocimiento se tiene los vocabularios controlados, taxonomías, tesauros y ontologías. Para la recuperación del conocimiento se puede usar la recuperación de información, *Question - Answering* y razonamiento basado en casos.

Para explicar los procesos de memoria organizacional en el contexto de una institución educación superior, se limitó el alcance de los retenedores del conocimiento, lo que permitió determinar los métodos y técnicas de ingeniería del conocimiento en el contexto de instituciones de educación superior. Para adquisición del conocimiento se puede utilizar la extracción de información, el descubrimiento de conocimiento en bases de datos y bases datos textuales. Para la codificación del conocimiento se puede utilizar las ontologías, que de acuerdo a la revisión de literaturas es la técnica más utilizada. Para la recuperación del conocimiento se puede usar de técnicas de recuperación de información.

Se diseñó la arquitectura de software para sistemas de memoria organizacional en instituciones de educación superior, la arquitectura está compuesta por 04 componentes: usuarios, repositorio de conocimiento, repositorio de documentos y el componente de visualización. El análisis de una institución de educación superior desde la perspectiva de la teoría de sistemas de (Bunge, 2006), proporciona una visión global de estas instituciones. El análisis desde el punto de vista de la metodología *CommonKADS* (Schreiber and Akkermans, 1999) permite enfocarse en las actividades intensivas de conocimiento y en el conocimiento como activo.

La validación de la arquitectura se realizó comparándola con las arquitecturas de la revisión del estado del arte. La arquitectura propuesta presenta muchas similitudes con las arquitecturas de la revisión del estado del arte, ya que comparten elementos similares como los usuarios, las herramientas para interactuar con el sistema, repositorio de documentos y todas hacen uso de ontologías para agregar contenido semántico a la información y/o documentos. Pero la arquitectura también se diferencia de las arquitecturas de otros autores, puesto que se centra en la representación del conocimiento. Algunos autores de la revisión del estado del arte hacen énfasis en el proceso adquisición del conocimiento y en la actualización de ontologías.

## 6.1. Trabajos Futuros

El siguiente objetivo es diseñar un framework basado en la arquitectura propuesta para manejar proyectos investigación en un departamento académico de una universidad, en el caso de ciencias de la computación se puede usar la ontología de (Cassel et al., 2007). Después se debe desarrollar un prototipo para validar las características de la arquitectura de software propuesta.

Es conveniente realizar una revisión sistemática para clarificar y ver como se complementan la definición de gestión del conocimiento y memoria organizacional puesto que hay confusión sobre su definición, incluso se usan indistintamente.

También se podría identificar las tareas intensivas en conocimiento en una institución de educación superior a través de un estudio cuantitativo, como el realizado por (Andreas Rath and Lindstaedt, 2013).

Además, se podría ampliar la arquitectura para contemplar el proceso de adquisición de conocimiento, tal como se ha realizado en algunas de las arquitecturas de la revisión del estado del arte.

Finalmente, se podría probar la arquitectura propuesta en otros entornos fuera del ámbito académico, como la industria, gobierno, salud, entre otros.





# Bibliografía

- (2001). An ontology-driven organisational memory for managing group competencies. *The University of York, Department of Computer Science*, 1(1):1 – 306.
- Aamodt, A. and Plaza, E. (1994). Case-based reasoning; foundational issues, methodological variations, and system approaches. *AI COMMUNICATIONS*, 7(1):39–59.
- Abecker, A., Bernardi, A., Hinkelmann, K., K O., #252, hn, and Sintek, M. (1998). Toward a technology for organizational memories. *IEEE Intelligent Systems*, 13(3):40–48.
- Abecker, A., Bernardi, A., van Elst, L., Herterich, R., Houy, C., Muller, S., Dioudis, S., Mentzas, G., and Ernst, P. (2001). Workflow-embedded organizational memory access: The decor project.
- Ackerman, M. S. (1998). Augmenting organizational memory: a field study of answer garden. *ACM Trans. Inf. Syst.*, 16(3):203–224.
- Ackerman, M. S. and Hadverson, C. A. (2000). Reexamining organizational memory. *Commun. ACM*, 43(1):58–64.
- Ackerman, M. S. and Hadverson, C. A. (2004). Organizational memory as objects, processes, and trajectories: An examination of organizational memory in use. *University of Michigan and IBM Research*, (1):1–28.
- Adil Laoufi, S. Mouhim, E. H. M. C. C. (2011). An ontology based architecture to support the knowledge management in higher education. *Multimedia Computing and Systems (ICMCS), 2011 International Conference*, page 6.
- Alavi, M. and Leidner, D. E. (2001). Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS Q.*, 25(1):107–136.
- Ale, M. A., Gerarduzzi, C., Chiotti, O., and Galli, M. R. (2008). Organizational knowledge sources integration through an ontology-based approach: The ontology architecture. volume 5288 LNAI of *1st World Summit on the Knowledge Society, WSKS 2008*, pages 441–450. Athens.
- Andreas Rath, D. D. and Lindstaedt, S. (2013). An ontology based solution for knowledge management and elearning integration.

- Atwood, M. E. (2002). Organizational memory systems - challenges for information technology. *Hawaii International Conference on Systems Sciences*, page 9.
- Baeza-Yates, R. A. and Ribeiro-Neto, B. (1999). *Modern Information Retrieval*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA.
- Bass, L., Clements, P., and Kazman, R. (2003). *Software Architecture in Practice, Second Edition*. Addison-Wesley Professional.
- Bechhofer, S. (2002). Ontology language standardisation efforts.
- Borst, W. N. (1997). Construction of engineering ontologies for knowledge sharing and reuse. *Thesis Doctoral Centre for Telematicas and Information Technology, Netherlands*, page 227.
- Boury-Brisset, A. C. and Tourigny, N. (2000). Knowledge capitalization through case bases and knowledge engineering for road safety analysis. *Knowledge-Based Systems*, 13(5):297–305.
- Bunge, M. (2006). Chasing reality - strife over realism. *Toronto Buffalo London*, page 357.
- Cassel, L. N., Davies, G., Fone, W., Hacquebard, A., Impagliazzo, J., LeBlanc, R. J., Little, J. C., McGettrick, A. D., and Pedrona, M. (2007). The computing ontology: application in education. *ACM SIGCSE Bulletin*, 39(4):171–183.
- Chaudhuri, S. and Dayal, U. (1997). An overview of data warehousing and olap technology. *SIGMOD Rec.*, 26(1):65–74.
- Conklin, J. (2001). Designing organizational memory: Preserving intellectual assets in a knowledge economy. *Conexus Institute*, page 41.
- Cunningham, H. (1997). Information extraction a user guide.
- Fayyad, U., Piatetsky-shapiro, G., and Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI Magazine*, 17:37–54.
- Feldman, R. and Dagan, I. (1995). Knowledge discovery in textual databases (kdt). In *In Proceedings of the First International Conference on Knowledge Discovery and Data Mining (KDD-95)*, pages 112–117. AAAI Press.
- Gaizauskas, R. and Wilks, Y. (1998). Information extraction: Beyond document retrieval.
- Garlan, D., Bachmann, F., Ivers, J., Stafford, J., Bass, L., Clements, P., and Merson, P. (2010). *Documenting Software Architectures: Views and Beyond*. Addison-Wesley Professional - Software Engineering Institute, 2nd edition.
- Garshol, L. M. (2004). *Metadata? Thesauri? Taxonomies? Topic maps! Making sense of it all*.
- Grishman, R. (1997). Information extraction: Techniques and challenges.

- Group, I. A. W. (2000). Ieee std 1471-2000, recommended practice for architectural description of software-intensive systems. Technical report, IEEE.
- Gruber, T. R. (1993). Toward principles for the design of ontologies used for knowledge sharing. *Int. J. Hum.-Comput. Stud.*, 43(5-6):907–928.
- Grunstein, M. and J.-P, B. (1996). An industrial view of the process of capitalizing knowledge. In *J. F. Schreinemakers ed, Knowledge Management: Organization, Competence and Methodology, Proc. of the 4th Int. Symposium on the Management of Industrial and Corporate Knowledge (ISMICK'96)*, page 6.
- Guarino, N. (1998). Formal ontology and information systems. pages 3–15. IOS Press.
- Hepp, M. and Bruijn, J. D. (2007). *A Generic Methodology for Deriving OWL and RDF-S Ontologies from Hierarchical Classifications, Thesauri, and Inconsistent Taxonomies*.
- Heravizadeh, M. and Edmond, D. (2008). Making workflows context aware a way to support knowledge intensive tasks.
- Hirschman, L. and Gaizauskas, R. (2001). Natural language question answering: The view from here. *Nat. Lang. Eng.*, 7(4):275–300.
- Ingwersen, P. (1992). Information retrieval interaction.
- Inmon, W. H. (1992). *Building the Data Warehouse*. John Wiley & Sons, Inc., New York, NY, USA.
- Jackson, P. (2008). An exploratory survey of the structure and components of organizational memory. In Klobas, J., editor, *Becoming Virtual*, Contributions to Management Science, pages 89–110. Physica-Verlag HD.
- Ju, T. L. (2006). Representing organizational memory for computer-aided utilization. *Journal of Information Science*, page 15.
- Jung, J., Choi, I., and Song, M. (2007). An integration architecture for knowledge management systems and business process management systems. *Computers in Industry*, 58(1):21 – 34.
- Kingston, J. (2007). Multi-perspective modeling for knowledge management and knowledge engineering. *Thesis Doctoral*, page 449.
- Kitchenham, B. (2004). Procedures for performing systematic reviews. *Software Engineering Group, Department of Computer, Science Keele University and Empirical Software Engineering National ICT Australia Ltd*, page 33.
- Kodratoff, Y. (1999). Knowledge discovery in texts: A definition, and applications. In *Proc. of the 11th International Symposium on Foundations of Intelligent Systems (ISMIS-99)*, pages 16–29. Springer.
- Kruengkrai, C. and Jaruskulchai, C. (2003). Generic text summarization using local and global properties of sentences. In *Web Intelligence, 2003. WI 2003. Proceedings. IEEE/WIC International Conference on*, pages 201–206.

- Kuhn, O. and Abecker, A. (1997). Corporate memories for knowledge management in industrial practice: Prospects and challenges. *Journal of Universal Computer Science*, 3(8).
- Lehner, F. and Maier, R. K. (2000). How can organizational memory theories contribute to organizational memory systems? *Information Systems Frontiers*, 2(3-4):277–298. Cited By (since 1996): 13 Export Date: 26 April 2012 Source: Scopus.
- Meenachi, N. M. and Baba, M. S. (2012). Article: A survey on usage of ontology in different domain. *International Journal of Applied Information Systems*, 4(2):46–55. Published by Foundation of Computer Science, New York, USA.
- Melgar Sasieta, H. A. (2011). *UM MODELO PARA A VISUALIZACAO DE CONHECIMENTO BASEADO EM IMAGENS SEMANTICAS*.
- Molla, D. and Vicedo, J. L. (2007). Question answering in restricted domains: An overview. pages 41–61.
- Nevo, D. and Wand, Y. (2005). Organizational memory information systems: A transactive memory approach. *Decision Support Systems*, 39(4):549–562.
- Nonaka, I. (1994). A dynamic theory of organizational knowledge creation. pages 3 – 42.
- Nonaka, I. (2000). Chapter 1 - a dynamic theory of organizational knowledge creation. In Smith, D. E., editor, *Knowledge, Groupware and the Internet*, pages 3 – 42. Butterworth-Heinemann, Boston.
- Ramesh, B. (1997). Towards a meta-model for representing organizational memory. In *System Sciences, 1997, Proceedings of the Thirtieth Hawaii International Conference on*, volume 2, pages 320–329 vol.2.
- Ramshaw, L. and Weischedel, R. (2005). Information extraction.
- Rees, R. V. (2003). Clarity in the usage of the terms ontology, taxonomy and classification.
- Roussinov, D. and Robles-Flores, J.-A. (2004). Web question answering: Technology and business applications. *AMCIS 2004 Proceedings*.
- Schreiber, G. T. and Akkermans, H. (1999). *Knowledge engineering and management: the CommonKADS methodology*. MIT Press, Cambridge, MA, USA.
- Sinkula, J. M. (1994). Market information processing and organizational learning.
- Stein, E. (1995). Organization memory: Review of concepts and recommendations for management. *International Journal of Information Management*, 15:15.
- Studer, R., Fensel, D., Decker, S., and Benjamins, V. R. (1999). *Knowledge Engineering: Survey and Future Directions*.

- Toledo, C. M., Ale, M. A., Chiotti, O., and Galli, M. R. (2011). An ontology-driven document retrieval strategy for organizational knowledge management systems. *Electronic Notes in Theoretical Computer Science*, 281:21–34.
- Uschold, M., Uschold, M., King, M., King, M., House, S. B. R., Moralee, S., Moralee, S., Zorgios, Y., and Zorgios, Y. (1995). The enterprise ontology. *The Knowledge Engineering Review*, 13:31–89.
- Walsh, J. P. and Ungson, G. R. (1991). Organizational memory. *Academy of Management Review*, 16.
- Watson, I. (2003). *Chapter 1 - Knowledge management and organizational memory*, pages 3–21. Morgan Kaufmann, San Francisco.
- Widom, J. (1995). Research problems in data warehousing. In *Proceedings of the Fourth International Conference on Information and Knowledge Management, CIKM '95*, pages 25–30, New York, NY, USA. ACM.
- Zelle, J. (2002). *Python Programming: An Introduction to Computer Science*. 1st edition.
- Zouaq, A., Frasson, C., and Nkambou, R. (2006). An ontology-based solution for knowledge management and elearning integration.