



UNIVERSIDAD TECNOLÓGICA NACIONAL

Facultad Regional Santa Fe

Doctorado en Ingeniería

Mención Ingeniería en Sistemas de Información

TESIS DOCTORAL

**“MODELO CONCEPTUAL DE GESTIÓN DEL
CONOCIMIENTO EMPRESARIAL”**

Mariel Alejandra Ale

DIRECTORA:

Dra. Maria Rosa Galli

CO-DIRECTOR:

Dr. Omar Chiotti

Santa Fe, Argentina.

Septiembre de 2009

Ale, Mariel Alejandra

Modelo conceptual de gestión del conocimiento empresarial. - 1a ed. - Santa Fe : el autor, 2010.
343 p. ; 21x15 cm.

ISBN 978-987-05-8131-4

1. Ingeniería. 2. Tesis del Doctorado. I. Título
CDD 620.711

UNIVERSIDAD TECNOLÓGICA NACIONAL

Facultad Regional Santa Fe

Comisión de Posgrado

Se presenta esta Tesis en cumplimiento de los requisitos exigidos por la Universidad Tecnológica Nacional para la obtención del grado académico de Doctor en Ingeniería, mención Sistemas de Información

“MODELO CONCEPTUAL DE GESTIÓN DEL
CONOCIMIENTO EMPRESARIAL”

por

Mariel Alejandra Ale

DIRECTORA: Dra. Maria Rosa Galli

CO-DIRECTOR: Dr. Omar Chiotti

JURADOS DE TESIS:

Dr. Luis Olsina

Dr. Héctor Enrique Salomone

Dra. Inés Friss de Kereki

Santa Fe, Argentina.

Septiembre de 2009

A mis padres y a Jorge.

Sin ellos nada de esto hubiera sido posible...

Índice

| | |
|--|--------------|
| <i>Índice</i> | <i>vii</i> |
| <i>Lista de Figuras</i> | <i>xii</i> |
| <i>Lista de Tablas</i> | <i>xv</i> |
| <i>Lista de Algoritmos</i> | <i>xvii</i> |
| <i>Prólogo</i> | <i>xix</i> |
| <i>Resumen</i> | <i>xxv</i> |
| <i>Reconocimientos</i> | <i>xxvii</i> |
| Capítulo I: El Conocimiento Organizacional | 29 |
| 1. Introducción | 29 |
| 2. El Conocimiento Organizacional | 30 |
| 3. La Jerarquía del Conocimiento | 37 |
| 4. El Conocimiento y el Contexto | 45 |
| 5. La Jerarquía Inversa | 46 |
| 6. Información como Conocimiento Explícito | 48 |
| 7. Clasificaciones del Conocimiento | 52 |
| 7.1. Conocimiento Formal vs. Conocimiento Informal..... | 52 |
| 7.2. Conocimiento Tácito vs. Conocimiento Focal..... | 53 |
| 7.3. Conocimiento Técnico (Know-what) vs. Conocimiento Laboral (Know-how) | 53 |
| 7.4. Conocimiento de la Compañía vs. Conocimiento Corporativo..... | 54 |
| 7.5. Conocimiento Distribuido vs. Conocimiento Centralizado | 54 |
| 8. Conclusiones | 55 |
| Capítulo II: La Gestión del Conocimiento | 57 |
| 1. Introducción | 57 |
| 2. La Necesidad de la Gestión del Conocimiento | 58 |
| 2.1. Necesidades Relacionadas con la Actividad Organizacional..... | 59 |
| 2.1.1 Recurso de negocio fundamental | 59 |
| 2.1.2 Recurso transversal..... | 59 |
| 2.2. Necesidad debido a Factores Organizacionales Externos..... | 60 |
| 2.2.1. Ventaja competitiva..... | 61 |
| 2.2.2. Globalización | 61 |
| 2.2.3. Cambios organizacionales y sociales..... | 61 |
| 2.3. Necesidad debido a Factores Organizacionales Internos | 62 |
| 2.3.1. Incorporación de conocimiento | 62 |
| 2.3.2. Necesidad de conciencia | 62 |
| 2.3.3. Gestión orientada a proyectos | 62 |
| 2.3.4. Internacionalización (dispersión geográfica) | 63 |
| 2.4. Necesidad debido a la Naturaleza de la Información y del Conocimiento | 63 |
| 2.4.1. Cambios tecnológicos..... | 63 |
| 2.4.2. Sobrecarga de información | 63 |
| 2.4.3. Heterogeneidad de la información..... | 64 |

| | |
|---|------------|
| 3. La Gestión del Conocimiento..... | 65 |
| 4. Conclusiones..... | 70 |
| <i>Capítulo III: Paradigmas de Gestión del Conocimiento</i> | <i>73</i> |
| 1. Introducción | 73 |
| 2. Modelos Descriptivos de la Gestión del Conocimiento..... | 75 |
| 2.1. Modelo de Pilares para la Gestión del Conocimiento | 76 |
| 2.2. Modelo de las Habilidades Centrales y Construcción del Conocimiento | 77 |
| 2.3. Modelo de Gestión de Conocimiento Organizacional..... | 78 |
| 2.4. Modelo de la Organización que Conoce | 79 |
| 2.5. Modelo de los Estadios de la Gestión del Conocimiento | 80 |
| 3. Modelos Prescriptivos de la Gestión del Conocimiento | 82 |
| 3.1. Modelo de Bienes Intangibles | 82 |
| 3.2. Modelo de Capital Intelectual..... | 83 |
| 3.3. Modelo de Conversiones de Conocimiento | 84 |
| 3.4. Modelo de la Transferencia del Conocimiento | 85 |
| 3.5. Modelo del Proceso de Gestión del Conocimiento | 86 |
| 4. Enfoques para la Implementación de Modelos Conceptuales de Gestión del Conocimiento Distribuido..... | 88 |
| 4.1. Enfoque de Repositorio..... | 88 |
| 4.2 Enfoque de Red | 92 |
| 4.3 Modelo de Red vs. Modelo de Repositorio | 96 |
| 5. Actividades de Gestión del Conocimiento | 97 |
| 5.1 Crear Conocimiento | 100 |
| 5.2 Compartir Conocimiento..... | 101 |
| 5.2.1. Contexto de Transferencia | 101 |
| 5.2.2. Comunidades de Práctica y de Conocimiento | 102 |
| 5.3 Representar y Recuperar Conocimiento..... | 106 |
| 6. Conclusiones..... | 108 |
| <i>Capítulo IV: Memorias Organizacionales y Ontologías.....</i> | <i>111</i> |
| 1. Introducción | 111 |
| 2. Tipología y Facetas de las Memorias Organizacionales..... | 111 |
| 3. Ontologías – Modelado de Conocimiento | 122 |
| 4. Conclusiones..... | 134 |
| <i>Capítulo V: Requisitos para un Modelo Conceptual de Gestión del Conocimiento</i> | <i>137</i> |
| 1. Introducción | 137 |
| 2. Requisito I: Alineación de las iniciativas de Gestión del Conocimiento con la Estrategia Organizacional | 137 |
| 3. Requisito II: Identificación del Conocimiento Organizacional (toma de conciencia)..... | 138 |
| 4. Requisito III: Estructuración de las Actividades de Gestión de Conocimiento | 138 |
| 5. Requisito IV: Consideración de las Principales Actividades Relacionadas con la Gestión del Conocimiento | 139 |

| | |
|--|-------------------|
| 5.1. Requisito IV.1: En Relación a la Creación del Conocimiento..... | 139 |
| 5.2. Requisito IV.2: En Relación a la Distribución del Conocimiento..... | 139 |
| 5.3. Requisito IV.3: En Relación a la Representación y Recuperación del Conocimiento | 140 |
| 6. Requisito V: Gestión Distribuida del Conocimiento | 141 |
| 7. Requisito VI: Balance entre los Aspectos Sociales y Tecnológicos de la Gestión del Conocimiento..... | 141 |
| 8. Requisito VII: Cambio en la Cultura Organizacional | 142 |
| 9. Satisfacción de Requisitos por parte de los Modelos de Gestión del Conocimiento | 142 |
| 10. Problemática a Resolver | 143 |
| 11. Conclusiones..... | 145 |
| <i>Capítulo VI: Modelo Conceptual de Gestión del Conocimiento Distribuido. 147</i> | |
| 1. Introducción | 147 |
| 2. Procesos que Componen el Modelo Conceptual..... | 148 |
| 2.1. Proceso de Creación de Conocimiento..... | 148 |
| 2.2. Proceso para Compartir Conocimiento | 154 |
| 2.3. Proceso de Representación y Recuperación de Conocimiento | 156 |
| 3. Modelo Conceptual de Gestión del Conocimiento Distribuido | 158 |
| 4. Adecuación del Modelo Propuesto a los Requisitos | 160 |
| 5. Conclusiones..... | 163 |
| <i>Capítulo VII: Arquitectura del Sistema de Memoria Organizacional Distribuida</i> | <i>165</i> |
| 1. Introducción | 165 |
| 2. Necesidad de una Gestión Distribuida..... | 166 |
| 3. Arquitectura de un Sistema de Memoria Organizacional Distribuida..... | 168 |
| 3.1. Componente de Representación del Conocimiento..... | 172 |
| 3.2. Componente de Procesamiento y Recuperación de Información..... | 181 |
| 3.3. Interfaz a Otros Dominios | 188 |
| 4. Conclusiones..... | 189 |
| <i>Capítulo VIII: Onto – DOM.....</i> | <i>191</i> |
| 1. Introducción | 191 |
| 2. Arquitectura Onto-DOM | 192 |
| 2.1. Proceso de Inicialización de Onto-DOM..... | 196 |
| 2.2. Componente de Representación de Conocimiento (KRL)..... | 200 |
| 2.2.1. Proceso de Etiquetado | 202 |
| 2.2.2. Proceso de Búsqueda de Ocurrencias Exactas | 202 |
| 2.2.3. Proceso de Expansión Semántica | 203 |
| 2.2.4. Proceso de Inclusión de Superclases | 204 |
| 2.2.5. Persistencia | 204 |
| 2.3. Componente de Recuperación de Información (IRPL) | 205 |
| 2.4. Componente de Propagación a otros Dominios (ODI)..... | 207 |
| 2.4.1 Propagación Completa..... | 208 |
| 2.4.2 Propagación Selectiva..... | 208 |

| | |
|--|------------|
| 3. Conclusiones | 212 |
| Capítulo IX: Análisis de Desempeño de Onto – DOM | 213 |
| 1. Introducción | 213 |
| 2. Caso de Estudio | 214 |
| 3. Análisis de Desempeño | 218 |
| 3.1 Componente de Representación de Conocimiento..... | 218 |
| 3.1.1 Procedimiento | 218 |
| 3.1.2 Métricas | 218 |
| 3.1.3 Anotado manual | 219 |
| 3.1.4 Anotado automático | 219 |
| 3.1.5 Resultados de la comparación | 219 |
| 3.2 Componente de Recuperación de Información..... | 224 |
| 3.2.1 Procedimiento | 224 |
| 3.2.2 Métricas | 224 |
| 3.2.3 Búsqueda manual | 225 |
| 3.2.4 Búsqueda automática..... | 225 |
| 3.2.5 Resultados de la comparación | 226 |
| 3.3 Componente de Propagación a otros Dominios | 232 |
| 3.3.1 Procedimiento | 233 |
| 3.3.2 Métricas | 234 |
| 3.3.3 Resultados de la comparación | 234 |
| 4. Conclusiones | 238 |
| Capítulo X: Conclusiones y Trabajos Futuros | 241 |
| 1. Introducción | 241 |
| 2. Principales Contribuciones | 241 |
| 2.1. Modelo Conceptual de Gestión del Conocimiento Distribuido | 242 |
| 2.2. Sistema de Memoria Organizacional Distribuida | 244 |
| 2.3. Onto-DOM: Prototipo de Memoria Organizacional..... | 245 |
| 3. Trabajos Futuros | 246 |
| Anexo A | 249 |
| 1. Proceso de Etiquetado | 249 |
| 2. Proceso de Búsqueda de Ocurrencias Exactas | 251 |
| 3. Proceso de Expansión Semántica | 253 |
| 4. Proceso de Inclusión de Superclases | 256 |
| 5. Persistencia | 257 |
| Anexo B | 259 |
| 1. Proceso General | 259 |
| 2. Proceso de Análisis Lingüístico | 261 |
| 2.1. Análisis Lingüístico | 263 |
| 2.2. Obtención de Palabras de Definición de un Verbo | 264 |
| 2.3. Obtención de los Sustantivos de Definición de un Verbo | 265 |
| 3. Proceso de Identificación de Sustantivos | 266 |
| 4. Proceso de Identificación de Relaciones | 267 |
| 4.1. Obtención de Relaciones Ontológicas..... | 269 |

| | |
|--|------------|
| 5. Identificación de Sustantivos de Definición de un Verbo..... | 271 |
| 6. Proceso de Preparación de Consulta | 272 |
| 6.1. Proceso de Validación Inicial | 274 |
| 7. Proceso de Recuperación | 275 |
| <i>Anexo C.....</i> | 281 |
| 1. Valoración de los Dominios..... | 281 |
| 2. Propagación..... | 282 |
| 3. Valoración de la consulta | 283 |
| <i>Anexo D</i> | 285 |
| 1. Ontología Europa-Travel..... | 285 |
| 2. Ontología África-Travel..... | 292 |
| 3. Consultas y Resultados de la Recuperación | 302 |
| <i>Anexo E.....</i> | 309 |
| 1. Herramientas de Gestión de Ontologías | 309 |
| 1.1 Protégé | 312 |
| 2. Lenguajes de Definición de Ontologías..... | 314 |
| 2.1. OWL..... | 317 |
| 3. Herramientas de Anotado Basadas en Ontologías | 318 |
| <i>Referencias.....</i> | 321 |

Lista de Figuras

| | |
|--|-----|
| Figura I.1. Datos, Información, Conocimiento y Sabiduría..... | 38 |
| Figura I.2. Jerarquía Inversa..... | 47 |
| Figura III.1. Modelo de los Pilares de Gestión del Conocimiento (Wiig)..... | 76 |
| Figura III.2. Modelo de las Habilidades Centrales y la Construcción de Conocimiento (Leonard-Barton)..... | 77 |
| Figura III.3. Modelo de Gestión del Conocimiento Organizacional (Arthur Andersen & APQC)..... | 79 |
| Figura III.4. Modelo de la Organización que Conoce (Choo)..... | 80 |
| Figura III.5. Modelo de los Estadios de Gestión de Conocimiento (Van der Spek-Spijkervet)..... | 81 |
| Figura III.6. Modelo de Bienes Intangibles (Sveiby)..... | 83 |
| Figura III.7. Modelo de Capital Intelectual (Petrash)..... | 84 |
| Figura III.8. Modelo de Conversiones de Conocimiento (Nonaka)..... | 85 |
| Figura III.9. Modelo del Proceso de Gestión del Conocimiento (Alavi)..... | 87 |
| Figura III.10. Flujo de Conocimiento en el Modelo de Repositorio..... | 88 |
| Figura III.11. Flujo de Conocimiento en el Modelo de Red..... | 93 |
| Figura III.12. Contextos de Transferencia de Conocimiento..... | 102 |
| Figura IV.1. Polisemia y Sinonimia..... | 134 |
| Figura VI.1. Proceso de Creación de Conocimiento..... | 149 |
| Figura VI.2. Círculo Virtuoso de Enseñanza..... | 150 |
| Figura VI.3. Red de Conocimiento..... | 156 |
| Figura VI.4. Modelo Conceptual de Gestión del Conocimiento Distribuido... | 160 |
| Figura VII.1. Arquitectura de Gestión de Conocimiento Distribuida a través de un Esquema de Memorias Organizacionales..... | 169 |
| Figura VII.2. Representación del Conocimiento y Recuperación de Información a través de una Memoria Organizacional en un Dominio de Conocimiento..... | 171 |
| Figura VII.3. Proceso de Etiquetado..... | 175 |
| Figura VII.4. Proceso de Búsqueda de Ocurrencias Exactas..... | 177 |
| Figura VII.5. Proceso de Expansión Semántica..... | 179 |
| Figura VII.6. Proceso de Inclusión de Superclases..... | 180 |
| Figura VII.7. Proceso de Etiquetado de la Consulta..... | 184 |
| Figura VII.8. Posible Representación de la Consulta..... | 185 |
| Figura VII.9. Puntos de Contacto entre los Verbos..... | 186 |
| Figura VII.10. Descriptores de la Consulta..... | 187 |

| | |
|---|-----|
| Figura VIII.1. Arquitectura Onto-DOM | 192 |
| Figura VIII.2. Modelo de Persistencia..... | 195 |
| Figura VIII.3. Diseño de Tablas..... | 196 |
| Figura VIII.4. Proceso de Inicialización de la Herramienta Onto-DOM..... | 198 |
| Figura VIII.5. Pantalla de Inicio de Onto-DOM..... | 200 |
| Figura VIII.6. Interfaz KRL..... | 201 |
| Figura VIII.7. Interfaz IRPL..... | 205 |
| Figura VIII.8. Módulo de Procesamiento y Recuperación de la Información | 207 |
| Figura VIII.9. Módulo de Interfaz a otros Dominios..... | 211 |
| Figura VIII.10. Resultados del Proceso de Propagación..... | 212 |
| Figura IX.1. Ontología Travel Extendida | 214 |
| Figura IX.2. Valor Promedio del Recall de la Estrategia de Anotado..... | 222 |
| Figura IX.3. Valor Promedio del Precision de la Estrategia de Anotado..... | 222 |
| Figura IX.4. Valor Promedio del Fallout de la Estrategia de Anotado..... | 223 |
| Figura IX.5. Valor Promedio del Accuracy de la Estrategia de Anotado..... | 223 |
| Figura IX.6. Gráfica de la Métrica Precision(Rjk) vs. Recall(Rjk)..... | 230 |
| Figura IX.7. Valores de MAP en Función del Número de Consultas | 231 |
| Figura IX.8. Valores de MAP en Función del Número de Consultas para África-Travel | 232 |
| Figura IX.9. Valores de MAP en Función del Número de Consultas para Europa-Travel | 232 |
| Figura IX.10. Evolución de los Valores de Accuracy | 238 |
| Figura A.1. Proceso de Etiquetado..... | 250 |
| Figura A.2. Proceso de Búsqueda de Ocurrencias Exactas..... | 252 |
| Figura A.3. Proceso de Expansión Semántica..... | 253 |
| Figura A.4. Proceso de Inclusión de Superclases..... | 257 |
| Figura A.5. Persistencia..... | 258 |
| Figura B.1. Proceso General de Recuperación de Información..... | 260 |
| Figura B.2. Proceso General de Análisis Lingüístico..... | 262 |
| Figura B.3. Análisis Lingüístico..... | 263 |
| Figura B.4. Obtención de Palabras de Definición de un Verbo..... | 264 |
| Figura B.5. Obtención de Sustantivos de Definición de un Verbo..... | 265 |
| Figura B.6. Proceso de Identificación de Sustantivos..... | 266 |
| Figura B.7. Proceso de Identificación de Relaciones..... | 268 |
| Figura B.8. Obtención de Relaciones Ontológicas..... | 270 |

| | |
|---|-----|
| Figura B.9. Identificación de Sustantivos de Definición de un Verbo..... | 271 |
| Figura B.10. Proceso de Preparación de Consulta..... | 273 |
| Figura B.11. Proceso de Recuperación..... | 275 |
| Figura C.1. Recuperación de Ratings..... | 281 |
| Figura C.2. Propagación de la Consulta..... | 283 |
| Figura C.3. Valoración de la Consulta..... | 284 |
| Figura D.1. Valores de MAP en Función del Número de Consultas para Europa | 307 |
| Figura D.2. Valores de MAP en Función del Número de Consultas para África | 307 |

Lista de Tablas

| | |
|--|-----|
| Tabla III.1. Comparación de los Modelos de Red y Repositorio de Conocimiento..... | 96 |
| Tabla III.2. Actividades y Componentes de la Gestión de Conocimiento..... | 98 |
| Tabla III.3. Capacidades Organizacionales para la Gestión del Conocimiento..... | 99 |
| Tabla III.4. Tipos de Comunidades..... | 105 |
| Tabla IV.1. Clasificación de las Memorias Organizacionales en función de los Procesos de Gestión que Implementan..... | 114 |
| Tabla V.1. Adecuación de los Modelos a los Requisitos..... | 143 |
| Tabla VI.1. Comparación entre Tutores y Entrenadores..... | 152 |
| Tabla VI.2. Adecuación del Modelo Conceptual Propuesto a los Requisitos ... | 162 |
| Tabla IX.1. Resultados del Anotado Automático en Comparación del Anotado Manual | 220 |
| Tabla IX.2. Resultados del Análisis de Desempeño de la Estrategia de Anotado..... | 223 |
| Tabla IX.3. Consultas Analizadas para Evaluar la Estrategia de de Recuperación | 225 |
| Tabla IX.4. Resultados Cuantitativos de la Búsqueda | 226 |
| Tabla IX.5. Valores de Rjk | 227 |
| Tabla IX.6. Valores de Precision (Rjk)..... | 228 |
| Tabla IX.7. Valores de Recall((Rjk) | 229 |
| Tabla IX.8. Consultas Propagadas | 234 |
| Tabla IX.9. Índices de Similitud | 235 |
| Tabla IX.10. Resultados de Propagación Selectiva | 236 |
| Tabla IX.11. Valores Promedios de la Accuracy para la Propagación Selectiva | 237 |
| Tabla A.1. Reglas Morfológicas de WordNet..... | 254 |
| Tabla B.1. Clasificación Tipo de Respuesta..... | 278 |
| Tabla D.1. Consultas analizadas en el Dominio Europa-Travel | 302 |
| Tabla D.2. Consultas analizadas en el Dominio África-Travel | 302 |
| Tabla D.3. Resultados Cuantitativos de la Búsqueda en el Dominio Europa-Travel | 303 |
| Tabla D.4. Resultados Cuantitativos de la Búsqueda en el Dominio África-Travel | 303 |

| | |
|---|-----|
| Tabla D.5. Valores Rkj del Dominio Europa-Travel..... | 304 |
| Tabla D.6. Valores Rkj del Dominio África-Travel..... | 304 |
| Tabla D.7. Valores Precision(Rjk) del Dominio Europa-Travel | 305 |
| Tabla D.8. Valores Precision(Rjk) del Dominio África-Travel | 305 |
| Tabla D.9. Valores Precision(Rjk) del Dominio Europa-Travel | 306 |
| Tabla D.10. Valores Precision(Rjk) del Dominio África-Travel | 306 |

Lista de Algoritmos

| | |
|---|-----|
| Algoritmo VII.1. Lógica del Componente de Representación del Conocimiento..... | 174 |
| Algoritmo VII.2. Lógica del Componente de Procesamiento y Recuperación de Información..... | 182 |
| Algoritmo VII.3. Lógica del Componente de Interfaz a Otros Dominios..... | 188 |

Prólogo

Existen varias formas de ver y definir a las organizaciones. Estas entidades abstractas y vivientes estructuran y regulan a la sociedad humana y su función principal es responder a las necesidades individuales y colectivas de sus miembros mediante la realización de un conjunto de actividades. Estas actividades requieren de un esfuerzo de gestión cuyo objetivo es una coordinación efectiva de los esfuerzos individuales para el alcance de los objetivos organizacionales.

La realización de las actividades individuales requiere de conocimiento y la sumatoria de este conocimiento individual conforma lo que se denomina el *conocimiento organizacional*. Una falla en la gestión de este conocimiento colectivo puede resultar en una disfunción organizacional. Más aún, en las últimas décadas se observa un cambio de paradigma organizacional en el cual el capital y el trabajo ya no son los pilares fundamentales de éxito, sino que la capacidad de utilizar el conocimiento organizacional surge como factor determinante en la obtención de ventajas competitivas.

En la actualidad, a medida que los mercados se hacen cada vez más dinámicos y competitivos, la pérdida de información, indefectiblemente, se traduce en pérdida de oportunidades. Las organizaciones deben reaccionar rápidamente a los cambios del entorno anticipando, al mismo tiempo, las necesidades a las que deben responder.

En este contexto, el conocimiento surge como un activo organizacional fundamental para la competitividad y la supervivencia haciendo indispensable que la gestión organizacional deba incluir a la Gestión del Conocimiento. La función principal de la Gestión del Conocimiento apunta a resolver problemas relacionados con la identificación, adquisición, almacenamiento, acceso, difusión, reutilización y mantenimiento tanto del conocimiento individual como colectivo.

El enfoque propuesto en esta tesis para gestionar el conocimiento organizacional es establecer una Memoria Organizacional Distribuida basada en Ontologías de Dominio. Esta Memoria Organizacional cumple una doble función: asegurando, por una parte, el almacenamiento persistente y el indexado del conocimiento organizacional y, por otro lado, capturando objetos de conocimiento relevantes para proveerlos a las personas que los requieren. Las fuentes de conocimiento asociadas a la Memoria Organizacional son heterogéneas y están conformadas por los diferentes objetos de conocimiento (documentos, bases de datos, perfiles, imágenes, video, etc.) que pertenecen a la organización. Esta tesis se enfoca a los documentos como fuente de conocimiento. La implementación de una Memoria Organizacional de estas características requiere del desarrollo de un marco teórico que le sirva de referencia, una serie de estrategias y una herramienta que le de soporte.

En esta tesis se mostrará que: (1) existe la necesidad de un modelo conceptual para la Gestión del Conocimiento organizacional más abarcativo que los existentes; (2) las ontologías son medios efectivos y aplicables para dar soporte a espacios de conocimiento distribuido y heterogéneo; (3) el modelo de Memoria Organizacional Distribuido propuesto es una arquitectura aplicable y efectiva para la Gestión del Conocimiento organizacional. Con este fin, el presente documento está dividido de la siguiente forma: en el *Capítulo I* se introducirá al conocimiento como bien organizacional, las diferentes definiciones y clasificaciones encontradas en la literatura y la jerarquía entre los datos, la información y el conocimiento. En el *Capítulo II*, se establece la necesidad de la Gestión de Conocimiento y se presentan las diferentes iniciativas de Gestión del Conocimiento Organizacional presentes en la literatura.

En el *Capítulo III*, se presentan los enfoques para la Gestión del Conocimiento, los diferentes modelos presentes en la literatura y las actividades que los mismos involucran, y en el *Capítulo IV* se discute la noción de Memoria Organizacional y su capacidad para responder a las necesidades de la Gestión del Conocimiento Organizacional. Se analizan, a su vez, los últimos avances en modelado de conocimiento, enfocándose especialmente en la noción de ontología

que es la parte del modelo de Memoria Organizacional Distribuida (presentado en el capítulo VII) que captura la semántica de los objetos de conocimiento de un dominio dado.

En el *Capítulo V* se derivan una serie de requisitos que debe satisfacer un modelo conceptual para la Gestión de Conocimiento. En el *Capítulo VI* se presenta el Modelo Conceptual para la Gestión del Conocimiento Distribuido propuesto y en el *Capítulo VII* se desarrolla una arquitectura propia de Memoria Organizacional Distribuida que permite la implementación de las actividades de Gestión del Conocimiento descritas en el capítulo anterior. En el *Capítulo Capítulo VIII* se presenta una herramienta denominada Onto-DOM que implementa las estrategias propuestas en la arquitectura de Memoria Organizacional Distribuida. El análisis de desempeño de los componentes de dicha herramienta se presenta en el *Capítulo IX*.

Finalmente, en el *Capítulo X* se presentan las conclusiones y trabajos futuros.

Algunos resultados de la presente tesis que han sido publicados son los siguientes:

- ✓ Ale, Mariel Alejandra; Gerarduzzi, Cristian; Chiotti, Omar y Galli, Maria Rosa; *Capítulo: Organizational Knowledge Sources Integration through an Ontology-Based Approach: The Onto-DOM Architecture*; Libro: "Emerging Technologies and Information Systems for the Knowledge Society"; Lecture Notes in Computer Science; M.D. Lytras et al. (Eds.): WSKS 2008, LNAI 5288, pp. 441–450; ISBN 978-3-540-87780-6; 2008.
- ✓ Ale, Mariel Alejandra; Gerarduzzi, Cristian; Chiotti, Omar y Galli, Maria Rosa; *Onto-DOM: Organizational Knowledge Sources Integration through an Ontology-Based Approach*; Proceedings de la XXXIV Conferencia Latinoamericana de Informática; pp. 569-578; ISBN 978-950-9770-02-7; 2008.
- ✓ Ale, Mariel Alejandra; Chiotti, Omar y Galli, Maria Rosa; *Capítulo 8: "Enterprise Knowledge Management for Emergent Organizations: An Ontology Driven Approach"*; Libro: "Knowledge Management Strategies: A Handbook Of Applied Technologies", Editores: Lytras, M. D., Russ, M., Maier, R., Naeve, A.; Ed. IGI Publishing. ISBN 978-1-59904-603-7; 2008.

- ✓ Ale, Mariel Alejandra; Gerarduzzi, Cristian; Chiotti, Omar y Galli, Maria Rosa; *Onto-Dom: A Question-Answering Ontology-Based Strategy For Heterogeneous Knowledge Sources*; ICFAI Journal of Knowledge Management; ICFAI University Press; pp. 54-68; ISSN 0972-9216; 2007.
- ✓ Ale, Mariel Alejandra; Gerarduzzi, Cristian; Chiotti, Omar y Galli, Maria Rosa; *Onto-Dom: A Question-Answering Ontology-Based Strategy For Heterogeneous Knowledge Sources*; Proceedings de las VI Jornadas Iberoamericanas de Ingeniería de Software e Ingeniería del Conocimiento (JIISIC 2007) realizadas del 31 de Enero al 02 de Febrero de 2007 en Lima, Perú, ISBN 978-9972-2885-1-7; Page(s) 79-86; 2007.
- ✓ Ale, Mariel Alejandra; Chiotti, Omar y Galli, Maria Rosa; *Gestión Del Conocimiento Empresarial En Organizaciones Emergentes: Un Enfoque Basado En Ontologías*; Proceedings del 4° Simposio de Sociedad de la Información (JAIIO 2006), realizadas el 06 y 07 de Septiembre de 2006 en Mendoza, Argentina; 2006.
- ✓ Ale, Mariel Alejandra; Chiotti, Omar y Galli, Maria Rosa; *Semantic Document Representation In An Ontology-Driven Distributed Organizational Memory*; Proceedings de las V Jornadas Iberoamericanas de Ingeniería de Software e Ingeniería del Conocimiento (JIISIC 2006) realizadas del 01 al 03 de Febrero de 2006 en Puebla, México, ISBN 970-84770-0-5, Page(s) 191-198; 2006.
- ✓ Ale, Mariel Alejandra; Gerarduzzi, Cristian; Chiotti, Omar y Galli, Maria Rosa; *A Distributed Knowledge Management Conceptual Model For Knowledge Organizations*; ICFAI Journal of Knowledge Management; ICFAI University Press; Vol. 3 Nro. 4; ISSN 0972-9216; Page(s) 27-39; December 2005.
- ✓ Ale, Mariel Alejandra; Chiotti, Omar y Galli, Maria Rosa; *A Strategy For Semantic Document Classification In An Ontology Driven Knowledge Management System*; Proceedings de la IADIS International Conference realizadas del 15 al 17 de Diciembre en Porto, Portugal, 2005, ISBN 972-8924-06-2, Page(s) 410-414; 2005.
- ✓ Ale, Mariel Alejandra; Chiotti, Omar y Galli, Maria Rosa; *Semantic Document Indexing In Ontology - Driven Organizational Memories*; Proceedings del XI Congreso Argentino de Ciencias de la Computación (CACIC) realizado en la ciudad de Concordia, Entre Rios, del 17 al 21 de Octubre de 2005.
- ✓ Ale, Mariel Alejandra; Chiotti, Omar y Galli, Maria Rosa; *A Distributed Knowledge Management Conceptual Model For Knowledge Organizations*; Proceedings de las Jornadas Iberoamericanas de Ingeniería de Software e Ingeniería del Conocimiento (JIISIC 2004),

realizadas del 03 al 05 de Noviembre de 2004, Madrid, España. ISBN 84-688-8957-1; 2004.

- ✓ Ale, Mariel Alejandra; Chiotti, Omar y Galli, Maria Rosa; *Agent-Supported Ontology-Based Knowledge Management System*; Proceedings del Simposio Argentino de Sistemas de Información (ASIS 2004) dentro del marco de las 33° Jornadas Argentinas de Informática e Investigación Operativa (JAIIO 2004), realizado el 20 al 24 de Septiembre de 2004, Córdoba, Argentina. Nro. páginas: 11; 2004.

Finalmente es de destacar que las publicaciones más importantes de los resultados de esta tesis se esperan obtener en el futuro inmediato, producto de haber concluido el trabajo.

Resumen

Para hacer frente a los desafíos que presenta el mercado actual las organizaciones deben ser capaces de gestionar eficientemente el conocimiento que poseen. Sin embargo, con frecuencia, los gerentes organizacionales no pueden identificar dónde reside el valor del conocimiento que poseen, ni cómo usarlo como ventaja competitiva. En la literatura asociada se describen una multiplicidad de modelos e iniciativas, cada uno de ellos se enfoca en ciertos elementos de la Gestión del Conocimiento pero ninguno los resume a todos. Una estrategia de Gestión del Conocimiento debe estar basada en un entendimiento exhaustivo de lo que implica la Gestión del Conocimiento.

En esta tesis se identifican un conjunto de requisitos que un modelo conceptual de Gestión del Conocimiento organizacional debería satisfacer para constituirse en marco de referencia para una implementación de Gestión de Conocimiento y para el desarrollo de tecnologías de información, y se muestra que ninguno de los modelos conceptuales de gestión del conocimiento propuestos en la bibliografía satisface todos estos requisitos. Surge entonces una problemática concreta que es la falta de un modelo conceptual unificado y más abarcativo que satisfaga todos los requisitos identificados en esta tesis que sirva como marco de referencia para las iniciativas de Gestión de Conocimiento y de desarrollo de tecnologías de información para su implementación concreta.

Con el propósito de solucionar esta problemática, el objetivo de esta tesis es proponer un Modelo Conceptual para la Gestión del Conocimiento Organizacional que, cumpliendo con todos los requisitos identificados, integre tanto los aspectos tecnológicos como los sociales de este fenómeno. Sustentada en este modelo como marco de referencia, se propone una arquitectura de Memoria Organizacional Distribuida que se implementa en un sistema de tres capas (Onto-DOM) que aborda dos problemas comunes en implementaciones de estas características: la sobrecarga de documentación que implica para los trabajadores la elicitación de conocimiento para contribuir a los repositorios y la descontextualización del conocimiento producto de su conversión entre sus

formas tácita y explícita. Se presentan además, una estrategia de anotado y una de recuperación basadas en ontologías, que permiten un tratamiento semántico automático de las fuentes de conocimiento organizacionales heterogéneas dentro de esta Memoria Organizacional.

Reconocimientos

Quisiera agradecer a todos aquellos que han ayudado a que esta tesis fuera posible. En primer lugar, quiero agradecer a mi directora de tesis, la Dra. Maria Rosa Galli, por su excelente dirección y ayuda constante, en especial por su orientación metodológica y por su continuo estímulo a lo largo de este trabajo de investigación hasta el final del mismo. El agradecimiento a mi codirector, el Dr. Omar Chiotti, en especial por fomentar un pensamiento crítico indispensable para este tipo de tarea y por sus acertadas correcciones.

Quiero agradecer también a la Universidad Tecnológica Nacional por haber hecho posible la realización de esta tesis con el otorgamiento de una Beca para Docentes; y a la Facultad Regional Santa Fe por el espacio y material brindado.

A los integrantes del Centro de Investigación y Desarrollo en Ingeniería en Sistemas de Información (CIDISI) de la UTN, por brindarme un ambiente de excelencia tanto a nivel de investigación como académico para el desarrollo de este trabajo. En especial a Cristian por su asistencia en la implementación del prototipo.

Agradezco también a todas aquellas personas, revisores, colegas, etc., que en forma directa o indirecta contribuyeron a que este trabajo de investigación pudiera llevarse a cabo mediante sus valiosas sugerencias.

A mis amigas, por su continuo soporte. A Jorge, por su paciencia y apoyo. Por último, mi más profundo agradecimiento a mis padres, sin ellos nada de esto hubiera sido posible.

Finalmente, quiero agradecer a la Universidad Tecnológica Nacional (UTN) y a la Agencia de Promoción de la Ciencia y la Tecnología (ANPCyT) por el apoyo económico brindado para la realización de este trabajo a través de subsidios a diversos proyectos de los cuales esta investigación formó parte.

Capítulo I

El Conocimiento Organizacional

1. Introducción

Hoy en día, muy pocos cuestionarían la afirmación de que vivimos en una sociedad basada en el conocimiento. Los expertos organizacionales coinciden en que el conocimiento es uno de los recursos más importantes que contribuyen a las ventajas competitivas de una organización, pero reconocen al mismo tiempo, que es un concepto multifacético con significados variados. Con frecuencia, se afirma que sólo aquellas organizaciones que puedan desarrollar las mejores prácticas para gestionar este complejo concepto de conocimiento serán las que puedan navegar con éxito la actual ola competitiva (Karp y col., 1999).

Sin embargo, a pesar del reconocimiento del conocimiento como activo organizacional, un primer desafío que enfrentan las empresas es la falta de claridad al momento de definir qué es conocimiento y diferenciarlo de los datos y la información. En la sección 2 se presentan las diferentes definiciones de estos conceptos y en la sección 3 se analiza la jerarquía que los relaciona y que ofrece alguna luz sobre la posibilidad de emplear las tecnologías de información para gestionar el conocimiento.

De las definiciones analizadas se desprende que un concepto crucial para comprender la distinción entre datos, información y conocimiento es el de contexto, el cual se analiza en la sección 4.

En la sección 5 se presenta un patrón de relación alternativo para estos conceptos y en la sección 6 se analiza a la información como una forma de conocimiento explicitado. A partir de esta noción de conocimiento explícito, en la sección 7 se presentan otras clasificaciones de conocimiento. Finalmente, en la sección 8 se presentan las conclusiones del presente capítulo.

2. El Conocimiento Organizacional

El comienzo del nuevo siglo ha traído un cambio de paradigma en el cual el capital y el trabajo ya no son las únicas bases fundamentales para una gestión exitosa. La globalización es otro factor que tiene implicancias significativas para el tratamiento del conocimiento organizacional. En este escenario global, la tendencia hacia productos intensivos en conocimiento hace que una gestión eficiente del conocimiento sea crítica para incorporar conocimiento en el diseño, producción y distribución de bienes y servicios de forma ágil y continua.

Actualmente, las organizaciones dependen fuertemente de su habilidad para identificar y usar adecuadamente el conocimiento que poseen y es por ello que, a lo largo de las dos últimas décadas, la posibilidad de hacer del conocimiento un bien organizacional ha capturado la atención de las empresas como una de las formas más prometedoras para alcanzar el éxito en la era de la información. En este contexto, las compañías comenzaron a entender la importancia del conocimiento como bien organizacional que hace posible obtener una ventaja competitiva sustentable y ven a su gestión como un requerimiento para sobrevivir en un entorno altamente competitivo (Bolloju y col., 2002; Desouza, 2003).

Hoy en día existe un gran número de actividades relacionadas con el tratamiento del conocimiento implementadas en las organizaciones, pero

carecen, en general, de una perspectiva estratégica. Como consecuencia, la gestión de este activo organizacional parece incluir toda clase de enfoques teóricos, actividades prácticas, medidas y tecnologías sin demasiada consideración a su relación con la estrategia de negocios. Por otro lado, existe un conjunto de autores que sugieren pragmáticamente una serie de actividades, esfuerzos o estrategias para tratar el conocimiento sin ninguna diferenciación entre estos conceptos. La mayoría de estos autores basan sus propuestas en estudios empíricos sobre iniciativas de gestión del conocimiento organizacional en las organizaciones durante los últimos años (Maier y Remus, 2001).

Lo que es más, aunque existe un creciente reconocimiento de la importancia del conocimiento organizacional, generalmente los gerentes no pueden identificar dónde reside el valor del conocimiento o cómo usarlo como una ventaja competitiva (Ale y col., 2006b). Como resultado, el conocimiento organizacional contenido en las personas y las comunidades formadas dentro de la organización raramente está suficientemente detallado para ser valioso y frecuentemente se pierde cuando las personas abandonan la organización.

Por otro lado, muchas organizaciones no han encarado actividades de gestión de una forma estructurada porque todavía no comprenden el alcance de este tipo de iniciativa o qué actividades implica. La razón de esto es la falta de claridad al identificar las principales características de los procesos necesarios para gestionar el conocimiento. Más aún, existe una importante brecha entre la investigación académica y las iniciativas prácticas en este campo y, como consecuencia, muchas organizaciones no tienen una estrategia consolidada y explícita para el tratamiento del conocimiento organizacional. En su lugar, muchos intentos de gestión sólo se han basado en la implementación de nuevas tecnologías de sistemas de información para plasmar todo el conocimiento posible de una organización en base de datos que serían de fácil acceso para todos los empleados (King, 1999; Levine, 2001). Esta filosofía de tratar el conocimiento como una “cosa” que puede ser gestionada como cualquier otro bien físico no ha sido exitosa por varias razones. Una de ellas es la aparente dificultad concerniente a la captura del conocimiento y la conversión entre sus

formas tácita y explícita, las cuales se desarrollarán con más detalle en los Capítulos VI y VII.

A pesar de la reconocida importancia de la necesidad de gestionar el conocimiento organizacional, no existe un consenso de lo que el mismo significa. Peor aún, como concluye Spiegler, se sufre de una falta de acuerdo aún en la definición del concepto mismo de “conocimiento”, confundiéndolo a menudo con los datos o la información (Spiegler, 2003). Para clarificar estos términos, Spiegler sugiere un modelo recursivo y en espiral que relaciona los tres conceptos (datos, información y conocimiento). En este modelo, los datos de ayer son la información de hoy y el conocimiento de mañana, y este último servirá como retroalimentación para futuros datos e información. Se discutirán a continuación diferentes visiones del conocimiento como así también la naturaleza de su relación con los datos y la información organizacional.

Bellinger y col. proponen una diferenciación simple donde los datos representan un hecho o declaración de un evento sin ninguna relación con otra cosa (Bellinger y col., 2004). La información abarca la comprensión de una relación de algún tipo entre los datos, posiblemente una relación causa-efecto. Finalmente, el conocimiento representa un patrón que conecta y generalmente provee un nivel más alto de predictibilidad.

Estos autores afirman que a pesar de que la gestión del conocimiento organizacional es un área muy activa en cuanto a investigación y desarrollo hay pocas referencias que definen sólo lo que es el conocimiento o cómo distinguirlo de conceptos relacionados tales como datos, información y sabiduría. Una revisión de la literatura específica muestra que se han ensayado algunas definiciones de conocimiento que van desde lo conceptual, a lo práctico o a lo filosófico y desde un alcance estrecho a uno más amplio. Se presentan aquí algunas breves definiciones de conocimiento relevantes para el alcance de esta tesis:

“El conocimiento es información organizada aplicable a la resolución de problemas” (Woolf, 1990).

“El conocimiento es información que ha sido organizada y analizada para hacerla comprensible y aplicable a la resolución de problemas o la toma de decisiones” (Turban y Frenzel, 1992).

“El conocimiento abarca las restricciones implícitas y explícitas puestas sobre los objetos (entidades), las operaciones y relaciones junto con heurísticas específicas y generales y procedimientos de inferencia en la situación que está siendo modelada” (Sowa, 1984).

“El conocimiento consiste de verdades y creencias, perspectivas y conceptos, juicios y expectativas, metodologías y conocimiento laboral (know-how)” (Wiig, 1993a).

“El conocimiento es el conjunto completo de entendimientos, experiencias y procedimientos que son considerados correctos y verdaderos y que por lo tanto guían los pensamientos, comportamientos y comunicación de las personas dentro de una organización” (van der Spek y Spijkervet, 1997).

“El conocimiento es el proceso de razonar sobre la información y los datos para permitir activamente el funcionamiento, la solución de problemas, la toma de decisiones, el aprendizaje y la enseñanza” (Ackoff, 1996).

Estas definiciones ofrecen alguna luz sobre lo que es el conocimiento a un alto nivel de abstracción, sin embargo hacen poco para ayudar a entender la verdadera naturaleza de la relación entre los datos, la información y el conocimiento o sobre cómo las tecnologías de información pueden usarse como soporte para ayudar a gestionar el conocimiento en una organización.

De acuerdo a Weggeman, conocimiento e información son términos primitivos, es decir, términos que son entendidos aunque no puedan ser exactamente definidos y cuyo significado recae en el correcto uso del concepto que sólo puede ser aprendido con la práctica de su uso (Weggeman, 1996). En la literatura específica se ha señalado con frecuencia que la relación entre datos, información y conocimiento es importante y generalmente, malinterpretada; siendo esta mala interpretación la génesis de la mayoría de los problemas en el

diseño de sistemas de información que dan soporte a las actividades de gestión del conocimiento organizacional.

Comenzando desde los datos se presentarán definiciones propias junto con las de los principales referentes en el área. Por ejemplo, Davenport y Prusak establecen que el conocimiento no es ni datos ni información aunque está relacionado con ambos, y la diferencia entre estos términos es generalmente una cuestión de grado (Davenport y Prusak, 1998). La confusión sobre lo que son los datos, la información y el conocimiento (en qué se diferencian, qué significan estas palabras) ha resultado en enormes gastos en iniciativas tecnológicas que raramente resultan en lo que las empresas que realizan la inversión necesitan o pensaron que obtendrían. La clave está en entender que existen considerables diferencias entre los datos, la información y el conocimiento.

Sveiby sostiene que parte de la presente confusión relacionada con cómo hacer negocios en la era del conocimiento probablemente se eliminaría si se tuviera un mejor entendimiento de las formas en las que la información y el conocimiento son tanto similares como diferentes (Sveiby, 1997a). La presunción ampliamente difundida y mayormente inconsciente de que la información es igual al conocimiento y que la relación entre una computadora y la información es equivalente a la relación entre el cerebro humano y el conocimiento humano pueden llevar a errores peligrosos y costosos.

La idea intuitiva de que el conocimiento es algo más que la información ha llevado a muchos autores a hacer distinciones entre los datos en bruto, la información y el conocimiento. A primera vista, estos conceptos parecen casi obvios para el sentido común, y aun así (o tal vez debido a ello) han sido una constante fuente de confusión. Por ejemplo, de acuerdo a algunos autores, los datos son símbolos que no han sido todavía interpretados, la información son datos con significado y el conocimiento es lo que permite a las personas asignar significado y por lo tanto generar información (van der Spek y Spijkervet, 1997).

Según Weggeman los datos son representaciones simbólicas de números, hechos, cantidades; un ítem de dato es lo que un sensor natural o artificial indica

sobre una variable (Weggeman, 1996). La información, a su vez, es el resultado de la comparación de estos datos, los cuales están situacionalmente estructurados para poder arribar a un mensaje que es significativo en un contexto dado. Para Fukuda los datos están hechos de símbolos y figuras que reflejan una percepción del mundo experimental (Fukuda, 1995). La información se obtiene a partir de los datos a los cuales se les ha dado un significado y han sido seleccionados como útiles. Por su parte, el conocimiento es información sistematizada, lo que debe entenderse como información que ha sido arreglada de acuerdo a un plan o esquema, es decir, por conocimiento existente. Dicho de otra manera, el conocimiento es una información que ha sido interpretada en un contexto (es decir, que se ha decidido su significado deseado) y cuyo significado fue articulado con conocimiento ya adquirido.

Una definición de datos y conocimiento desde un punto de vista lógico la presenta Fernández, quien afirma que un ítem de dato es la representación de un hecho particular, donde los datos son puramente extensionales (Fernández, 2000). Por extensional se entiende que la representación no denota más que un solo hecho, y por lo tanto sólo dice algo sobre la ocurrencia de un único evento. Un ítem de conocimiento es la representación de muchos hechos particulares, en el sentido de que con un único ítem de conocimiento se pueden explicitar varios hechos particulares mediante la aplicación de reglas de inferencia. El conocimiento es primariamente intencional, lo que quiere decir que la representación denota más de un hecho y por lo tanto dice algo sobre la ocurrencia de varios eventos.

Según Davenport los datos son simples observaciones de los estados del mundo, la información son datos provistos de relevancia y propósito y el conocimiento es información valiosa (Davenport, 1997). Para Sveiby la información no tiene significado, pero se convierte en conocimiento significativo cuando es interpretado (Sveiby, 1997a). Wiig afirma que la información consiste de hechos o datos que están organizados para describir una situación particular o condición mientras que el conocimiento consiste en verdades o creencias, perspectivas y conceptos, juicios y expectativas, metodologías y conocimiento de

trabajo (know-how) (Wiig, 1993b). Finalmente Nonaka y Takeuchi establecen que la información es un flujo de mensajes significativos para comenzar, pero se convierte en conocimiento cuando se crean compromisos y creencias como resultado de estos mensajes (Nonaka y Takeuchi, 1995).

Subyacente a todas estas definiciones de conocimiento como *una forma más alta de información* está la idea de que el conocimiento tiene que ser extraído de material en bruto, y en este proceso, se le tiene que agregar significado. Aunque Nonaka y Takeuchi, Wiig y Sveiby puntualizan que el conocimiento tiene que ver con la acción, la mayoría de las veces el conocimiento es conceptualizado como una representación útil, exacta y significativa de hechos en un contexto (Nonaka y Takeuchi, 1995) (Wiig, 1993b) (Sveiby, 1997a). La concepción subyacente también asume continuidad; es decir, un modelo de proceso donde algo simple se convierte en algo más complejo y valioso.

Si se asume que el objeto del conocimiento es una realidad externa que puede ser estudiada empíricamente para aprender su estructura y estados, es intuitivamente claro que primero se necesita observar algunos hechos simples antes de poder crear conocimiento. Sin embargo, es comúnmente sabido que los *datos en bruto* no existen, y que aún la percepción más elemental está ya influenciada por los potenciales usuarios, sus expectativas, el contexto y construcciones teóricas (Kuhn, 1996). El proceso cognitivo humano no puede ver hechos simples sin que estos hechos se conviertan en parte de su correspondiente estructura de significados.

Lo que es más, la mayor parte de esta estructura de significado es un trasfondo no articulado contra la cual se realizan las explicaciones y articulaciones. Es por ello que los sistemas de información de soporte para el conocimiento organizacional necesitan también tener en cuenta este componente no articulado de significado. Como Polanyi puntualiza, es imposible articular todo el significado al mismo tiempo. El conocimiento tácito o no explicitado, siguiendo la terminología de Polanyi, es ese *halo de conciencia* o trasfondo contra el cual emerge el significado de forma intencional, conciente y focal

(Polanyi, 1997). El componente tácito es dinámico en el sentido que cada vez que el foco del conocimiento cambia, algún significado tácito previo se convierte en focal y el resto de la estructura pasa a ser tácita. Un tema clave en el diseño de sistemas de información soporte para la gestión de conocimiento es, por lo tanto, tener en cuenta esos almacenes de significado compartidos socialmente que permanecen tácitos en el proceso de articulación.

El conocimiento, en definitiva, es una mezcla fluida de experiencias, valores, información contextual y entendimiento experto que provee un marco para evaluar e incorporar nuevas experiencias e información. Se origina y es aplicado en la mente de los conocedores. En las organizaciones, generalmente está embebido no sólo en documentos o repositorios sino también en rutinas, procesos, prácticas y normas organizacionales (Davenport y Prusak, 1998).

3. La Jerarquía del Conocimiento

De acuerdo a las definiciones vistas en la sección anterior, el conocimiento y algunos conceptos relacionados, tales como datos, información, sabiduría y entendimiento pueden ser organizados en una jerarquía que comienza a ofrecer alguna luz sobre cómo emplear las tecnologías de información para gestionar el conocimiento. Ackoff describe estos conceptos como contenidos de un proceso mucho más abarcativo que es el de aprendizaje y sugiere que forman una jerarquía de valor incremental (Ackoff, 1996).

Comúnmente los datos se ven como hechos simples que pueden estructurarse para convertirlos en información. La información, a su vez, se convierte en conocimiento cuando es interpretada, puesta en contexto, o cuando se le agrega significado. Si bien existen algunas variaciones de este esquema ampliamente adoptado, la idea común es que los datos son algo de menor jerarquía que la información, y que la información es menos que el conocimiento. Lo que es más, se asume que primero son necesarios los datos antes de que la información pueda crearse, y sólo cuando tenemos la información

puede surgir el conocimiento. Una representación de esta visión puede verse en la Figura I.1 donde los datos son vistos como hechos simples aislados. Cuando esos hechos son puestos en contexto y combinados en una estructura emerge la información. Eventualmente, cuando a la información se le da un significado mediante la interpretación, la información se convierte en conocimiento. En este punto, los hechos existen dentro de una estructura mental que la conciencia puede procesar, por ejemplo, para producir consecuencias futuras o para hacer inferencias. A medida que la mente humana usa este conocimiento para elegir entre varias alternativas el comportamiento se hace cada vez más inteligente. Finalmente, cuando las creencias y los valores guían este comportamiento inteligente, se dice que el comportamiento se basa en la sabiduría.

Esta visión incremental ve la construcción del conocimiento de forma similar a usar letras como átomos para la construcción de palabras que son subsecuentemente combinadas en oraciones con significado. La curva que se observa en la figura intenta simbolizar que el valor de las varias formas de los datos, la información y el conocimiento se incrementa a través del aprendizaje. Podría decirse que en este proceso los datos son progresivamente *refinados* ganando significado y alcance.

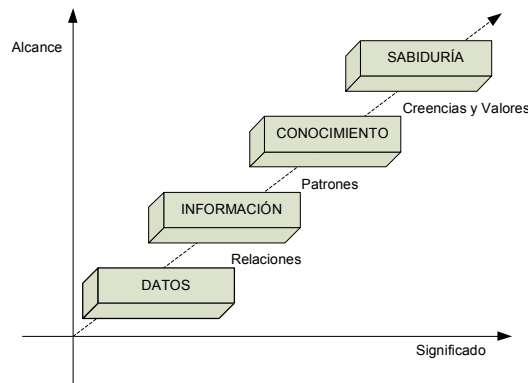


Figura I.1. Datos, Información, Conocimiento y Sabiduría

Analizando con más detalle esta jerarquía puede verse que esta visión es ampliamente compartida por una serie de autores, aunque difieren en algunos detalles. Por ejemplo, Davenport y Prusak establecen que los datos son un conjunto de referencias objetivas y discretas sobre eventos (Davenport y Prusak, 1998). En este sentido, los datos sólo describen una parte de lo que sucede; no proveen juicio o interpretación y ninguna base sustentable de acción, es decir, no dicen nada sobre su propia importancia o relevancia. Sin embargo, de acuerdo a estos autores, los datos se convierten en información tan pronto como se les da significado. La información debe comunicar, son datos que hacen la diferencia.

Otros autores puntualizan que los datos son generalmente símbolos en bruto que simplemente existen y no tienen ningún significado más allá de esta existencia. Ackoff describe a los datos como productos de observaciones hechas por personas o instrumentos (Ackoff, 1989). Bellinger, por su parte, afirma que los datos en bruto carecen de contexto y por lo tanto no tienen un significado por sí mismos ni tienen ninguna relación significativa a ninguna otra cosa (Bellinger y col., 2004). Para asignar significado a los datos las personas crean un contexto asociando los datos con otros objetos o proveyendo alguna interpretación.

Ackoff compara los datos con las minas de metales preciosos en el sentido de que ambos tienen poco o ningún valor hasta que son procesados en una forma usable (Ackoff, 1989). Los datos pueden existir en cualquier forma, ya sea usable o no. Una colección de datos no es información si no hay relación entre ellos (Fleming, 1996). Un ejemplo puede ser una planilla de cálculo que contiene datos previos a que se le apliquen fórmulas para organizarlos y resumirlos en relaciones tales como sumatorias, promedios, etc.

Subiendo un escalón en la jerarquía, la información consiste en datos organizados en relaciones y estructuras significativas. Son datos a los que se le ha agregado significado a través de conexiones relacionales y este significado puede ser útil, pero esto no es necesario. Contrariamente, Ackoff afirma que la diferencia entre datos e información no es de carácter estructural sino que está relacionada con la utilidad funcional, es decir, la información es más útil que los datos.

Para comprender la información deben entenderse las relaciones entre los elementos de datos. La información es usualmente estática y lineal y no provee ninguna comprensión de por qué los datos existen de la forma en que lo hacen. La información es factual por naturaleza e implica una descripción, definición o perspectiva y puede proveer respuestas a preguntas del tipo qué, quién, cuándo y dónde. En este punto cabe aclarar que una colección de información no es conocimiento si no hay patrones o procesos repetibles entre los elementos de información. Un ejemplo de generación de información a partir de datos podría ser cómo una base de datos relacional construye información a partir de los datos que almacena.

Estos autores señalan que se tiene, en general, una sensación intuitiva de que el conocimiento es algo más que los datos o la información, algo más amplio, profundo y rico. Sostienen que el conocimiento se deriva de la información como la información se deriva de los datos. Ven al conocimiento como información refinada donde la cognición humana ha adicionado valor, así, la información se convierte en conocimiento a través de un esfuerzo cognitivo. Por ejemplo, la mente humana puede comparar información sobre una situación específica con otras situaciones que ha conocido, anticipar consecuencias para las decisiones y las acciones, relacionar porciones de conocimiento con otras porciones de conocimiento y compartir interpretaciones con otras personas. Como resultado, y a diferencia de los datos y la información, el conocimiento contiene juicio. Los valores y creencias son integrales al conocimiento, determinando en gran medida lo que el conocedor ve, absorbe y concluye de sus observaciones.

Para Bateson, el conocimiento es información organizada en patrones significativos y procesos repetibles (Bateson, 2002). Los patrones son más que relaciones de relaciones, ejemplifican la completitud y consistencia de las relaciones y hasta cierto punto crean su propio contexto. Los patrones también tienen capacidades de predicción y repetición implícitas y por lo tanto sirven como arquetipos (Senge, 2006).

El conocimiento representado a través de patrones tiene una tendencia a ser más auto-contextualizable que la información o los datos. Los patrones raramente son estáticos por naturaleza y si uno entiende un patrón, exhibirá un alto grado de predictibilidad y confiabilidad en términos de cómo evolucionará en el tiempo. Los patrones que representan el conocimiento tienen un grado mayor de completitud que la información pero para entender el conocimiento uno debe entender los patrones que existen entre la información que lo comprende.

El conocimiento, además, es una colección de información apropiada en el sentido de que su intención es ser útil. Es procedural por naturaleza y comprende estrategia, práctica, método o enfoque y puede responder a preguntas sobre cómo realizar procedimientos en términos que pueden tener un orden o temporalidad específicos.

Como puede concluirse por lo expresado hasta ahora, a pesar de que parece existir un consenso sobre la idea de que el conocimiento es más que la información, existen varias visiones sobre su relación exacta. Una de las descripciones más detalladas de la jerarquía conceptual del conocimiento es la ofrecida por Earl (Earl, 1994). La misma difiere de la mayoría de las jerarquías existentes ya que la característica distintiva del conocimiento es su aceptación social. Esto refleja la idea de que el conocimiento tiene que ser interpersonal y objetivo.

De acuerdo a Earl, existen en realidad cuatro niveles de conocimiento para entender la información organizacional, cada uno de los niveles representando una cantidad mayor de estructura, certidumbre y validación. En primer lugar, los eventos organizacionales se representan, se recolectan y procesan para generar datos. Los datos son manipulados, presentados e interpretados aún más para generar información. La información luego conduce al conocimiento al ser testeada, validada y codificada. Earl enfatiza la idea de que el conocimiento emerge a través de una validación interpersonal. El concepto subyacente, sin embargo, es todavía aquel basado en la visión de los datos como materia prima desde la cual se crea el conocimiento.

La razón por la que se ve el conocimiento como resultado final de un proceso en el cual es creado a partir de datos e información es que, en la mayoría del trabajo hecho dentro de la visión tradicional, la epistemología subyacente está basada en presunciones empíricas y realistas. En la historia de la epistemología occidental, esta línea de pensamiento puede rastrearse hasta Aristóteles. En este sentido, se asume que el mundo existe independientemente de sus observadores, y que el verdadero conocimiento tiene que ser independiente del conocedor, potencialmente disponible para todos los observadores cuidadosos.

Esta visión aristotélica del conocimiento también es adoptada por Nonaka y Takeuchi, quienes sostienen que el conocimiento es una creencia verdadera y justificada (Nonaka y Takeuchi, 1995). Esta es una consecuencia lógica de su posición de que el conocimiento es, hablando estrictamente, creado sólo por individuos, pero que el proceso de conversión de conocimiento es esencialmente social. Como consecuencia, la verdad se convierte en un aspecto necesario, conectando el conocimiento intrapersonal con la realidad interpersonal y objetiva. A pesar de su esfuerzo por subestimar la importancia de la “verdad” como un factor constitutivo del conocimiento su concepción del conocimiento individual hace tal objetividad inevitable.

De más está decir que las definiciones de conocimiento abundan y también sus clasificaciones, tipos y variaciones de acuerdo al campo de experticia que se esté estudiando. Sin embargo, una de las más útiles y pragmáticas definiciones para el propósito de la comprensión de la influencia cultural en la construcción del conocimiento fue establecida por Joseph C. Pitt. Este autor afirma que los individuos producen aseveraciones candidatas a conocimiento y estos candidatos se convierten en conocimiento una vez que son respaldadas por la comunidad apropiada usando acuerdos realizados bajo estándares compartidos (Pitt, 2001).

Esto define un proceso que establece claramente la creación del conocimiento a través de la interacción social dependiendo de los estándares y una comunidad para su validación. De acuerdo a esta visión, los modelos

mentales se *construyen* a través de un proceso de desarrollo evolutivo que se bosqueja sobre la experiencia personal y cultural compartida. Estos estadios del proceso son jerárquicos, describiendo la realidad en pasos incrementales con un significado y alcance cada vez mayores.

Analizando los puntos en común de las visiones descritas hasta aquí puede afirmarse que los *datos* son observaciones en bruto extraídas del entorno a través de la percepción sensorial individual. Pueden haber sido acumulados personalmente o almacenados por otros y pueden ser cualitativos o cuantitativos, discretos o continuos; pero permanecen sin procesar, sin relacionar y no proveen ningún tipo de inteligencia más allá de su valor como observación.

La *información* es, en su mayoría, datos descriptivos y correlacionados que ofrecen historia, ubicación, condiciones, perfiles y otras perspectivas contextuales que construyen un marco para las situaciones y circunstancias que existen en el entorno. En este nivel también se infiere algún juicio sobre el significado del contenido de la información, generalmente con respecto a su validez o a su urgencia. Puede decidirse además rechazarla o aceptarla, y si es aceptada, si requiere algún tipo de acción, sólo es necesario comunicarla o simplemente debe ser archivada.

El *conocimiento* requiere que se agregue tanto el significado como la utilidad a la información para hacerla funcional. Para tener conocimiento, deben entenderse las relaciones y patrones implícitos en los datos y la información. También debe tenerse la habilidad explícita para aplicar la información y lograr resultados predecibles. El conocimiento también contiene las bases para su propia expansión proveyendo enlaces cognitivos y asociaciones que pueden usarse individualmente o en grupos, para enmendar, corregir o mejorar los modelos actuales de la realidad. Sin embargo, no todo el conocimiento está necesariamente basado en la realidad. Esto se debe a que el conocimiento extrae entendimiento de la cultura, historia y tradición además de la experiencia cognitiva, y por lo tanto puede incluir elementos conceptuales que caen fuera de los límites del entorno observable.

Finalmente, el conocimiento se convierte en *sabiduría* mediante un proceso de asimilación cultural. A los nuevos conceptos y modelos de la realidad se le asignan valores y éstos son relacionados a principios, buenas costumbres e ideales en contexto con los sistemas de creencias establecidos. La sabiduría puede ser incrementada o desafiada por nuevo conocimiento, pero tiene una fuerte base emocional y no siempre está sujeta a cambio por argumentos estrictamente lógicos y racionales. Lo que la gente cree es una función que depende fuertemente de sus prejuicios, preferencias, tradiciones y normas sociales.

Con respecto a este último estadio puede decirse que la sabiduría puede definirse como conocimiento intemporal que se desarrolla cuando se entienden los principios fundamentales que subyacen los patrones. Aún más que el conocimiento, la sabiduría tiende a ser auto-contextualizable y es un proceso no determinístico y no probabilístico. Se basa en tipos especiales de razonamiento humano (creativo, innovativo) y a diferencia de los niveles previos puede realizar preguntas para las cuales no hay respuestas simples y en algunos casos ninguna respuesta.

Ackoff afirma que hay una diferencia entre eficiencia y eficacia que diferencia a la sabiduría de los niveles más bajos de la jerarquía (Ackoff, 1996). Los datos, la información y el conocimiento contribuyen primariamente a la eficiencia, mientras que la sabiduría es necesaria para asegurar la eficacia. Los datos, la información y el conocimiento tienen valor en el sentido de que facilitan el proceso de persecución de objetivos y salidas deseadas. La sabiduría se utiliza para seleccionar los objetivos correctos y por lo tanto está asociada a la eficacia, es decir, a la elección de las metas tomando en cuenta el valor de la salida.

4. El Conocimiento y el Contexto

Un concepto relacionado con los datos, la información y el conocimiento y que con frecuencia es utilizado para diferenciar uno del otro es el contexto. Por *contexto* se entiende la colección de condiciones relevantes e influencias circundantes que hacen que una situación sea única y comprensible. Bellinger y col. argumentan que a medida que uno se mueve hacia arriba en la jerarquía de los datos hacia la sabiduría, el nivel de independencia del contexto decrece (Bellinger y col., 2004). Tomando una interpretación ligeramente diferente se puede considerar cuán embebido está el contexto en la construcción. En este caso a medida que nos movemos hacia arriba en la jerarquía hay un creciente nivel de *significado implícito* o contexto embebido dentro de las relaciones, patrones y principios asociados con los diferentes niveles.

Los datos por sí mismos no tienen contexto hasta que alguien los interpreta y le asigna un contexto. La información tiene algo de contexto embebido por la naturaleza de las relaciones que resumen y organizan los datos que la comprenden. El conocimiento tiene más contexto embebido a través de los patrones que son formados mediante el arreglo de la información que contiene. Finalmente, la sabiduría tiene la mayor cantidad de contexto embebido dentro de los principios que forman sus valores y arquetipos.

La experiencia también es un constructor del contexto, en el sentido que los principiantes pueden encontrar más difícil asignar el contexto o interpretar los datos de una manera significativa o desde perspectivas útiles y variadas que los expertos de un dominio dado. Esto es porque los principiantes tienen mayor dificultad en reconocer y comprender relaciones, patrones y principios que comprenden información, conocimiento y sabiduría respectivamente y, como resultado, pueden tener más dificultad en inferir el contexto apropiado para un tema dado que aquellos que tienen más experiencia. A medida que el contexto está más embebido en las construcciones de más alto nivel los principiantes pueden encontrar mayores obstáculos que los expertos para generalizar conocimiento (patrones) y sabiduría (principios) a otros contextos.

Son más escasas las definiciones que parten desde el conocimiento para llegar a los datos estableciendo una serie de jerarquía inversa aunque en la manipulación individual y colectiva del conocimiento vamos desde los datos hacia el conocimiento para analizar y en el sentido contrario para comunicar (Tuomi, 1999). En la siguiente sección se presenta esta jerarquía inversa.

5. La Jerarquía Inversa

Según esta visión, los datos emergen últimos, sólo después que hay información y conocimiento disponibles. No hay *porciones aisladas de hechos simples* a menos que alguien las haya creado usando su conocimiento. Los datos sólo pueden emerger si una estructura significativa o semántica es primero establecida y luego usada para representar información. Esto pasa, por ejemplo como se representa en la Figura I.2., cuando la información es almacenada en una base de datos semánticamente bien definida. En este caso especial, se tiene que descontextualizar el conocimiento y estructurarlo de acuerdo a una semántica predefinida en entradas de base de datos aisladas e independientes. Idealmente, los datos así producidos pueden ser completamente separados de cualquier significado, de forma tal que puedan ser automáticamente procesados usando sistemas de información. Los datos, por lo tanto, existen solamente como una solución práctica a un problema: cómo dividir la información de forma que pueda ser modelada, representada y procesada separadamente.

La estructura de significado que subyace al conocimiento de un individuo se articula a través de un esfuerzo cognitivo para convertirse en focal y estructurado. Si el significado es articulado dentro de un contexto conceptual y lingüístico, puede convertirse en verbal o textual. A ese punto se lo denomina conceptualmente información (puede estar presente en un documento y puesto en un archivo). Cuando ese conocimiento articulado es almacenado en la memoria de una computadora para manipularlo automáticamente debe representarse el significado de la información. Esto quiere decir que la información debe

separarse en “átomos” que no tienen significado que deba ser tenido en cuenta en su procesamiento automático. En este punto se han creado datos y para arribar a este estadio, se necesita una gran cantidad de esfuerzo cognitivo y trabajo de diseño.



Figura I.2. Jerarquía Inversa

Los datos, por lo tanto, no se convierten en información después de que se les haya agregado significado sino, por el contrario, el dato es creado a partir de la información poniendo a la información en una estructura de datos predefinida que define completamente su significado. En lugar de ser la materia prima de la información, los datos emergen como resultado de agregar valor a la información poniéndola en una forma en la que pueda ser automáticamente procesada.

Visto desde otra perspectiva, puede decirse que cuando existe una herramienta de medición, ya no podemos reinterpretar libremente lo que significan sus lecturas. En el proceso de creación de una herramienta de

medición se incorporan aspectos importantes del conocimiento en la estructura del mecanismo de medición. Otra forma de decir esto es que la herramienta usada para recolectar los datos (por ejemplo, temperatura) fija el significado de las relaciones (lo que es realmente la temperatura). Los datos, por lo tanto, existen sólo después de que se ha hecho ese prejuizgamiento.

6. Información como Conocimiento Explícito

La jerarquía inversa presentada conceptualiza la información como un producto que es creado a partir del conocimiento. Como un producto, es externalizado y hecho interpersonal. Subyacente a esta visión está la idea de que la información es creada a partir del conocimiento a través de un proceso de articulación. Cuando se estructura el conocimiento y es embebido en artefactos se convierte en un objeto de información físico que puede ser compartido. En algunos casos esta articulación ocurre a través del lenguaje, en otros casos puede ocurrir a través de la creación de diseños, imágenes y herramientas.

En general, cuando se diseñan sistemas de información el conocimiento articulado y explícito es sólo la punta del iceberg. Por ejemplo, para entender un documento almacenado en un sistema se necesita una gran cantidad de información contextual y generalmente este conocimiento no está almacenado dentro del sistema. En su lugar, los diseñadores de sistemas de información implícitamente confían en los almacenes de conocimiento tácito acumulados en los individuos y culturalmente compartidos para que faciliten un contexto de interpretación.

La articulación del conocimiento en las organizaciones requiere como prerequisites de un considerable esfuerzo y de almacenes de conocimiento tácito compartido socialmente bien desarrollados. Aunque la distribución del conocimiento ha sido entendida como distribución de documentos, y el conocimiento como texto articulado, en la práctica, está claro que simplemente incrementando la distribución de documentos dentro de una organización no se

logra la solución final para los problemas de gestión del conocimiento organizacional.

Una respuesta natural a este desafío es esforzarse para agregar más información contextual a los sistemas de gestión de documentos, o tratar de representar en forma más completa ese conocimiento organizacional tácito que fue dejado sin articular. Esto no es, sin embargo, una solución robusta o completa al problema. En realidad, apunta a poner más tecnología para resolver un problema que originalmente fue creado usando esta misma tecnología. En lugar de hacer más de lo mismo, es necesario hacer algo diferente.

Por ejemplo, se tiene que reconsiderar la relación entre conocimiento tácito y explícitamente articulado. Michael Polanyi afirma que *podemos conocer más de lo que podemos decir* (Polanyi, 1997). En la terminología de Polanyi, el conocer emerge de una interacción dinámica entre los componentes focales y secundarios del significado. De acuerdo a Polanyi, el conocimiento secundario consiste en pistas contextuales y subliminales de las cuales no estamos concientes. En lugar de eso, estas pistas marginales y subliminales proveen el contexto contra el cual el conocimiento focal obtiene su forma. A este componente *de fondo* Polanyi lo denomina tácito, argumentando que actúa como el trasfondo no articulado contra el cual todo significado focal es distinguido.

Siguiendo a Polanyi, Nonaka y Takeuchi basan su modelo de creación del conocimiento en la interacción dinámica entre estos dos tipos de conocimiento. El conocimiento tácito, de acuerdo a Nonaka y Takeuchi, es personal, específico de contexto y por lo tanto, difícil de formalizar y comunicar. El conocimiento explícito, por el contrario, se refiere a conocimiento transmisible en un lenguaje sistemático y formal (Nonaka y Takeuchi, 1995).

De acuerdo a estos autores, el conocimiento tácito incluye elementos técnicos y cognitivos. Los elementos cognitivos incluyen modelos mentales, tales como esquemas de datos, paradigmas, perspectivas, creencias y puntos de vista y ayudan a los individuos a percibir y definir su mundo. Los elementos técnicos, por otro lado, incluyen conocimiento laboral (know-how) concreto,

habilidades manuales y destrezas. La idea central en el modelo Nonaka-Takeuchi es que el nuevo conocimiento es creado mediante la articulación de modelos mentales tácitos en una especie de *proceso de movilización*. En este proceso, el conocimiento tácito es convertido a una forma explícita (Nonaka y Takeuchi, 1995).

Como Nonaka y Takeuchi comienzan con una distinción primaria entre conocimiento tácito y explícito, es importante notar las diferentes formas en las que Nonaka-Takeuchi y Polanyi se refieren a estos conceptos (Polanyi, 1997) (Nonaka y Takeuchi, 1995). Para Polanyi, el conocimiento tácito es una precondition para el conocimiento focal significativo y no hay conocimiento explícito sin una estructura significativa tácita, marginal y secundaria que subyace a todo conocimiento focal (Polanyi, 1997). Es por lo tanto imposible separar dos almacenes de conocimiento diferentes, uno tácito y otro focal. En su lugar, el almacén de conocimiento tácito es el trasfondo desde el cual el conocedor da forma al conocimiento focal.

Usando el concepto de *calidad de tácito* definida por Polanyi, puede verse que, a nivel individual, el conocimiento no es convertido en un conjunto separado de conocimiento explícito. En su lugar, la estructura de significado cambia de forma tal que parte de ella se convierte en focal en relación al resto que provee la periferia y el trasfondo (Polanyi, 1997).

En el nivel social, esencialmente ocurre el mismo proceso cuando el conocimiento tácito individual se convierte en conocimiento tácito compartido colectivamente. En este proceso de socialización el trasfondo tácito es provisto por una estructura de significado socialmente compartido, construida a través de un proceso social y cultural que es internalizado por los miembros de una sociedad durante su desarrollo cognitivo y social.

A diferencia de Polanyi, Nonaka y Takeuchi usan la distinción tácito/explicito para diferenciar almacenes de conocimiento articulados y no articulados. A menudo consideran la articulación y la verbalización como equivalentes. Por lo tanto, su modelo de conversión de conocimiento

específicamente hace alusión a aquellos procesos cognitivos y sociales que subyacen a la aparición de símbolos, conceptos, lenguaje y sistemas conceptuales (Nonaka y Takeuchi, 1995) (Polanyi, 1997). Sin embargo, el modelo Nonaka-Takeuchi debería ser clarificado agregándole un componente crucial: el almacén de conocimiento compartido comunalmente que hace posible la socialización, la articulación y la externalización en primer lugar. Este almacén es construido y reabastecido a través de un proceso de significado social. Dicho proceso conecta a las personas en sociedad y comunidades a través de un desarrollo socio-cultural y también restringe las posibilidades del procesamiento de significado. Parte del nuevo conocimiento creado es fácilmente conectado a la estructura de significado existente, mientras que parte del conocimiento cae en desuso sin dejar rastros apreciables en la memoria colectiva de la organización.

Usando esta distinción entre calidad de tácito y explícito, se pueden distinguir dos tipos de información. En primer lugar, la información puede ser conocimiento potencial. Este es un punto de vista del *observador* quien percibe el entorno como una fuente potencial de eventos significativos que hacen la diferencia. En segundo lugar, la información puede articularse como artefactos cuya intención es servir como medio de comunicación. Este es el punto de vista del *comunicador*, el cual asume un procesamiento del significado interpersonal. En el primer caso, la información es todo aquello que tenga sentido, en el segundo, es algo que se intenta tenga sentido. En ambos casos, sin embargo, la información emerge contra un almacén de conocimiento tácito que es necesario para interpretar esta información.

En la práctica, el problema del trasfondo tácito no sólo está contenido en un único individuo, sino que se presenta una situación ligeramente más complicada. En la mayoría de los sistemas de gestión del conocimiento el problema no es sólo aumentar o extender las capacidades cognitivas de un solo individuo. En su lugar, el foco del problema es la creación de sistemas que puedan dar soporte a la gestión del conocimiento organizacional interpersonal.

La problemática de la interacción entre el conocimiento tácito y explícito será abordada con más detalle en capítulos subsiguientes, pero existen otras clasificaciones de conocimiento que se detallan a continuación. La mención de estas clasificaciones es importante ya que cada una de ellas hace énfasis en aspectos claves a tener en cuenta al momento de gestionar el conocimiento.

7. Clasificaciones del Conocimiento

Se han identificado varias categorizaciones de conocimiento en la literatura específica, las cuales resaltan diferentes aspectos focales en relación a lo que es el conocimiento y lo que el mismo comprende:

7.1. Conocimiento Formal vs. Conocimiento Informal

En su estado natural el contexto del conocimiento incluye una situación de interpretación y un actor que interpreta esa información. Sin embargo, cuando se concibe una representación del conocimiento persistente, corpórea y explícita, significa que es un conocimiento artificial, o más conocido como conocimiento formalizado donde la información y el contexto son capturados en un sistema simbólico y su modelo adjunto, que permiten respectivamente un procesamiento automático y una interpretación no ambigua de los resultados. Este enfoque, que proviene de la inteligencia artificial, está cercano a la visión del área de gestión para la cual el conocimiento es un recurso artificial que sólo tiene valor dentro de un contexto apropiado, que puede ser revaluado o devaluado, que es inextinguible y usable, pero no consumido. Desafortunadamente gran cantidad de conocimiento viene en un contenedor (persona, documento, etc.) que no obedece a esto último.

Euzenat afirma que aunque el conocimiento debe ser establecido lo más formalmente posible, no todo el conocimiento puede o debe ser formalizado y

aún si esto fuera posible, las representaciones formales podrían sufrir de serias limitaciones relacionadas con la complejidad o incompletitud (Euzenat, 1996). Además, para enfatizar la idea de que las categorías formal e informal no son mutuamente excluyentes establece que es posible envolver un esqueleto de conocimiento formal con elementos más informales tales como texto, imágenes, animaciones, etc. En este sentido, es posible asociar a un cuerpo formal aquel conocimiento que no ha alcanzado todavía un estado formal, como por ejemplo, comentarios sobre la producción de conocimiento, o explicaciones informales. Estas afirmaciones parecen reconocer el hecho de que en los repositorios de conocimiento de una organización habrá diferentes grados de formalización del conocimiento y que el sistema de gestión debería ocuparse de esa característica.

7.2. Conocimiento Tácito vs. Conocimiento Focal

Sveiby discute el conocimiento tácito vs. el focal (Sveiby, 1997a). Según este autor, en cada actividad hay dos dimensiones de conocimiento, las cuales son mutuamente exclusivas: el conocimiento focal sobre el objeto o fenómeno que está en foco y el conocimiento tácito que es usado como herramienta para gestionar o mejorar lo que está en foco. Las dimensiones tácita y focal son complementarias. El conocimiento tácito funciona como conocimiento de fondo que asiste en el cumplimiento de la tarea que está en foco. Lo que es tácito varía de una situación a otra.

7.3. Conocimiento Técnico (Know-what) vs. Conocimiento Laboral (Know-how)

Seely Brown y col. afirman que el conocimiento organizacional para constituir una *competencia central* requiere de conocimiento técnico (know-what) y de conocimiento laboral (know-how) (Seely Brown y Duguid, 1998). El conocimiento técnico es conocimiento explícito que puede ser compartido por

varias personas. El conocimiento laboral es la habilidad particular para poner el conocimiento técnico en práctica. Mientras que ambos trabajan juntos, circulan separadamente. El conocimiento técnico circula con relativa facilidad y por lo tanto es difícil de proteger. Por el contrario, el conocimiento laboral puede ser difícil de diseminar, coordinar, comparar o cambiar ya que consiste en una disposición manifestada en la práctica. Como consecuencia, el conocimiento laboral es crítico para hacer que el conocimiento sea accionable y operacional.

7.4. Conocimiento de la Compañía vs. Conocimiento Corporativo

Esta clasificación está asociada al nivel organizacional donde se aplica el conocimiento. El conocimiento de la compañía es conocimiento técnico usado dentro de la compañía, sus unidades de negocio, departamentos, subsidiarias (es el conocimiento que necesitan los trabajadores en sus actividades diarias); el conocimiento corporativo es conocimiento estratégico usado por la gerencia a nivel corporativo (conocimiento sobre la compañía) (Grunstein y Barthés, 1996).

7.5. Conocimiento Distribuido vs. Conocimiento Centralizado

La distribución del conocimiento en una organización, o en una sociedad como un todo, refleja la división del trabajo. Como fue establecido por Adam Smith, la división del trabajo es una gran fuente de dinamismo y eficiencia (Smith, 1776). Las divisiones jerárquicas del trabajo distinguen a los pensadores de los hacedores, la labor mental de la manual y la estrategia (el conocimiento requerido en el tope de la jerarquía) de la táctica (el conocimiento usado en los niveles inferiores). Sin embargo, una división mental-manual predispone a las organizaciones a ignorar un bien central, el valor del conocimiento laboral (know-how) creado por todas sus partes. Con frecuencia se trata de contrarrestar la natural distribución del conocimiento organizacional con la creación de

grandes repositorios para centralizar dicho conocimiento lo cual, como se analizará más adelante, presenta serias desventajas.

8. Conclusiones

En la última década, una gestión efectiva del conocimiento organizacional se ha convertido en uno de los factores críticos para el éxito en el mercado actual. En el presente capítulo se analizó la importancia del conocimiento como bien organizacional dejando la fundamentación de la necesidad de su gestión para el Capítulo II.

Aunque parecen existir numerosas definiciones de conocimiento y una multiplicidad de categorizaciones, un consenso subyacente a la mayoría de ellas es la idea de que no todo el conocimiento puede ser comunicado de forma explícita. En esta tesis se distingue entre conocimiento tácito (incorporado en la mente de las personas) y conocimiento explícito (que puede ser representado por diferentes tipos de artefactos como textos, videos, etc.) y se trabajará sobre los procesos de conversión entre los mismos.

En la literatura asociada al tratamiento del conocimiento organizacional se ha señalado con frecuencia que la relación entre datos, información y conocimiento es importante y generalmente, malinterpretada; siendo esta mala interpretación la génesis de la mayoría de los problemas en el diseño de sistemas de información que dan soporte a las actividades para gestionar el conocimiento (Tuomi, 1999). Es por ello que en este capítulo se trabajó en diferenciar estos términos y en este sentido se encontraron dos posturas, en principio, contrapuestas. Una que parte de los datos para obtener conocimiento, y la otra que utiliza como punto de partida el conocimiento. A los efectos de esta tesis y desde el punto de vista de las tecnologías de información ambos enfoques son válidos. En el enfoque que va desde el conocimiento a los datos, se plantea que mediante un proceso de articulación o explicitación, el conocimiento se puede estructurar en información y ésta en datos. Por otro lado, mediante un proceso de

adquisición o aprendizaje, los datos son convertidos en información y ésta en conocimiento. En el Capítulo VII se analizarán con detalle estos procesos de conversión y se describirá su rol en el Modelo Conceptual de Gestión del Conocimiento propuesto.

En el siguiente capítulo se analiza qué se entiende por Gestión del Conocimiento Organizacional.

La Gestión del Conocimiento

1. Introducción

Gran parte de los estudios de la Gestión de Conocimiento han sido basados en el trabajo de Polanyi (Polanyi, 1997). Polanyi usa como ejemplo la habilidad de integrar el significado de una palabra con la cosa externa que denota esta palabra para ilustrar que *podemos conocer más de lo que podemos decir*. El análisis de Polanyi enfatiza varios conceptos claves. En primer lugar, la habilidad de identificar los objetos externos, y por lo tanto conocer, es adquirida a través de un proceso de experiencia personal. Segundo, la calidad de tácito o explícito son dimensiones independientes, el incremento de una no significa la disminución de la otra. Finalmente, ya que el conocimiento tácito es un elemento esencial de cualquier tipo de conocimiento y es adquirido a través de experiencia personal, cualquier esfuerzo de alcanzar un conocimiento absolutamente objetivo y separado está mal encaminado y destinado al fracaso (Zhang y Faerman, 2002). Polanyi es ampliamente citado en la literatura de Gestión del Conocimiento, pero como la mayor parte de su obra es anterior a 1970, su trabajo está situado en un contexto filosófico y se enfoca en la definición del conocimiento pero no en el esfuerzo sistemático de gestionarlo.

La conceptualización de la Gestión del Conocimiento no se desarrolló, sin embargo, hasta que el conocimiento se convirtió en un factor central para la producción y la innovación en los 90s. Peter Drucker fue uno de los primeros que abogó por el advenimiento de la sociedad del conocimiento. Drucker documentó la transformación desde el Capitalismo a una Sociedad del Conocimiento, la cual comienza poco después de la Segunda Guerra Mundial, puntualizando que el recurso más importante ya no era el capital, la tierra o el trabajo, sino más bien, el conocimiento (Drucker, 1988). En la sección 2 se analiza a la Gestión del Conocimiento como una necesidad desde diversos puntos de vista.

El campo de la Gestión del Conocimiento también ha sido moldeado por la experiencia y filosofía de la sociedad oriental. Nonaka y Takeuchi presentaron un trabajo pionero en el mapeo del conocimiento tácito y explícito para demostrar la dinámica de la creación de conocimiento. De acuerdo a estos autores, las organizaciones no solamente son capaces de procesar información adquirida desde el exterior, sino también crear nuevo conocimiento desde adentro redefiniendo los problemas y las soluciones, como así también el entorno. Ellos argumentan que la creación del conocimiento organizacional es una interacción dinámica y continua entre el conocimiento tácito y el explícito (Nonaka y Takeuchi, 1995). En la sección 3 se presentan algunas definiciones de Gestión del Conocimiento y finalmente en la sección 4 se presentan las conclusiones del presente capítulo.

2. La Necesidad de la Gestión del Conocimiento

Numerosas son las razones para la implementación de iniciativas de Gestión del Conocimiento y están relacionadas con factores diversos tales como la actividad de la organización, factores internos y externos y a la propia naturaleza del conocimiento.

2.1. Necesidades Relacionadas con la Actividad Organizacional

2.1.1 Recurso de negocio fundamental: Drucker advertía que el conocimiento era un recurso primario de la sociedad y pronosticaba que las implicancias de este cambio de paradigma serían increíblemente significativas para las organizaciones (compañías comerciales, instituciones públicas, etc.) (Drucker, 1988). El surgimiento de trabajadores del conocimiento y organizaciones intensivas en conocimiento durante la última década probó que estaba en lo cierto. El conocimiento es ahora considerado como un capital que tiene un valor económico, es un recurso estratégico para incrementar la productividad, es un factor de estabilidad en un entorno competitivo dinámico e inestable y es una ventaja competitiva decisiva (Ermine, 2000). Lo anterior implica, además, la necesidad de asegurar la protección de este recurso vital abordando el problema de la seguridad del conocimiento existente y del nuevo, como así también su almacenamiento para hacerlo persistente a través del tiempo. Además, como la creación del conocimiento se ha convertido en una capacidad organizacional crítica, los gerentes deben encarar esta tarea de una forma más reflexiva apoyándose en métodos y herramientas que los asistan en los procesos de gestión (Seely Brown y Duguid, 1998).

2.1.2 Recurso transversal: el conocimiento atraviesa numerosas áreas organizacionales: I&D, gerencia, producción, recursos humanos, etc. y en general, es necesitado lejos de su fuente de creación. La disponibilidad de conocimiento conduce a la toma de decisiones informadas por parte de los individuos, promoviendo a su vez la efectividad y viabilidad de la toma de decisiones descentralizada (Nagendra Prasad y Plaza, 1996).

2.2. Necesidad debido a Factores Organizacionales Externos

Prusak argumentó que el incremento del interés en el conocimiento está relacionado con la globalización y la toma de conciencia del valor del conocimiento a nivel social, como también a nivel de las unidades de producción individuales (Prusak, 1998). Los avances en el transporte, las tecnologías de telecomunicación y cómputo, como también el decaimiento de las economías centralizadas han propulsado la creación de un mercado global. La ganancia potencial de distribuir productos y servicios a mercados más amplios y utilizar completamente un mercado laboral ilimitado es atractiva para los gerentes, aunque requiere que sean capaces de manejar una complejidad sin precedentes. En este contexto, es necesario conocer qué se sabe y qué no, quién lo sabe y qué debería saberse. Lo que es más, es imperioso abordar el problema de compartir conocimiento y comprender una nueva forma de trabajo a través de locaciones distribuidas (Koenig y Srikantaiah, 2000).

La Gestión del Conocimiento, por lo tanto ha crecido desde una comprensión del valor crítico del conocimiento para la prosperidad y el desarrollo económico. Desde que Drucker previó en primer lugar una economía basada en conocimiento ha sido ampliamente aceptado que el conocimiento es un factor de producción crítico que necesita ser contabilizado y fomentado (Drucker, 1988). Neef notó que el crecimiento de una economía basada en servicios es altamente dependiente del conocimiento y las habilidades que permiten la solución de problemas complejos, de la innovación tecnológica, la exploración creativa de nuevos mercados y del desarrollo de nuevos productos o la oferta de nuevos servicios (Neef, 1998). A un nivel micro-económico, las compañías reconocen que el conocimiento, generalmente en la forma de capital intelectual y conocimiento laboral (know-how), es necesario para las ventajas competitivas a largo plazo (Hall, 1998). El poder competitivo de las compañías basadas en el conocimiento es generalmente derivado de la habilidad de explotar mercados y de innovar.

Puntualmente, los cambios en el contexto económico que fomentaron el cambio de paradigma del capital al de conocimiento son:

2.2.1. Ventaja competitiva: según Nonaka, en una economía donde lo único cierto es la incertidumbre, la única fuente segura de ventaja competitiva es el conocimiento (Nonaka, 1994). La globalización está conduciendo a una fuerte competencia internacional que se expresa mediante rápidos desarrollos de mercado y ciclos de vida más cortos en productos y servicios. Durante las actividades organizacionales se genera conocimiento que si no es almacenado, indexado y recuperado cuando es necesario, se pierde. En un mercado global como el actual, donde la información viaja a la velocidad de Internet, la pérdida de memoria en una organización es peligrosa; una memoria eficiente se convierte, por lo tanto, en un activo organizacional de vital importancia.

2.2.2. Globalización: este fenómeno fue impulsado por el avance en las tecnologías de información y comunicación, pasando de ser un fenómeno social a ser un factor dominante de la supervivencia organizacional en el mercado actual.

2.2.3. Cambios organizacionales y sociales: Drucker afirmaba que en la sociedad Post-Capitalista, la distinción social entre burguesía y proletariado desaparece y es reemplazada por dos grupos de trabajadores: los trabajadores de conocimiento y los de servicio (Drucker, 1988). Con el crecimiento de la población de trabajadores de conocimiento y su capacidad de producción, es cada vez más importante identificar, fomentar y compartir el conocimiento.

2.3. Necesidad debido a Factores Organizacionales Internos

2.3.1. Incorporación de conocimiento: el conocimiento ganado en el transcurso de las actividades diarias generalmente se pierde en un entorno de negocios altamente dinámico (Dzbor y col., 2000). La incorporación de conocimiento en artefactos que se modifican (por ejemplo, la evolución de productos o software) y en miembros de la organización que pueden abandonar la organización (retiro, movilidad interna o externa, etc.) son las principales causas de los problemas en la capitalización del conocimiento.

2.3.2. Necesidad de conciencia: es común que los miembros de una organización generalmente no estén concientes de los recursos críticos que permanecen ocultos en los grandes repositorios. Esta es una de las razones por las cuales mucho del conocimiento es olvidado al corto tiempo de ser desarrollado. La presión competitiva requiere de reacciones rápidas y efectivas a las situaciones siempre cambiantes del mercado. La brecha entre los recursos de conocimiento e información colectivos continuamente evolucionando y cambiando de una organización y la conciencia de los empleados de la existencia de tales recursos y de sus cambios pueden conducir a una pérdida en la productividad (Nagendra Prasad y Plaza, 1996). Muchas veces una parte de la organización repite el trabajo de otra parte simplemente porque es imposible llevar el rastro y hacer uso del conocimiento generado. Las organizaciones necesitan saber cuáles son sus bienes de conocimiento corporativos y cómo gestionarlos y usarlos para capitalizarlos. Los gerentes se están dando cuenta de lo importante que es *saber cuánto saben* para hacer un uso óptimo de su conocimiento (Kogut y Zander, 1996).

2.3.3. Gestión orientada a proyectos: esta nueva forma de gestión apunta a que las organizaciones se conviertan en sistemas de aprendizaje que evitan repetir los mismos errores. La información sobre proyectos anteriores:

protocolos, especificaciones de diseño, documentación de experiencias: tanto de fallas como de éxitos, alternativas exploradas, etc., todo es útil para transferir experticia dentro y a través de la organización.

2.3.4. Internacionalización (dispersión geográfica): el otro lado de la moneda de la globalización es la expansión de las compañías a consorcios globales. Como las compañías se expanden internacionalmente, las barreras geográficas pueden afectar el intercambio de conocimiento y evitar un fácil acceso a la información (O'Leary, 1998). Se amplía el problema ya existente de la circulación y distribución del conocimiento entre los artefactos y las personas dentro de una organización a un nivel global. Lo que es más, la internacionalización trae otros problemas como por ejemplo hacer que el conocimiento cruce culturas y fronteras del lenguaje.

2.4. Necesidad debido a la Naturaleza de la Información y del Conocimiento

2.4.1. Cambios tecnológicos: el desarrollo de las tecnologías de información, especialmente de las telecomunicaciones y redes, facilita el intercambio y el almacenamiento de información y conocimiento. Con el desarrollo omnipresente de poder de cómputo e infraestructuras de red cada vez más grandes, es posible explotar el conocimiento combinando grandes cantidades de datos desde varias fuentes. Muchas tecnologías usadas en Gestión del Conocimiento son extensiones naturales de soluciones tecnológicas en otras áreas, tales como data warehouses, técnicas de minería de datos, agentes de software, web, intranets e Internet.

2.4.2. Sobrecarga de información: con la cantidad creciente de datos en línea (base de datos, intranets, data warehouses, Internet, etc.) las organizaciones se están ahogando en información pero carecen del conocimiento necesario para el

desarrollo de sus actividades (Hussain y Wahba, 2002). Los problemas a resolver en este sentido tienen que ver con la relevancia, la evolución y la anticipación de las necesidades de conocimiento. En la actualidad, la disponibilidad de recursos de información aumenta exponencialmente sobrecargando a los trabajadores que usan conocimiento ya que demasiada información obstaculiza la toma de decisiones tanto como lo hace la insuficiencia de conocimiento.

2.4.3. Heterogeneidad de la información: los bienes de conocimiento residen en diferentes lugares tales como base de datos, base de conocimiento, documentos y la mente de las personas. Están distribuidos a través de toda la organización y las intranets e Internet aumentan la complejidad del problema (Kogut y Zander, 1996). Una consecuencia de lo anterior es que las fuentes diversas, fragmentadas y dispersas aumentan la sobrecarga cognitiva de los trabajadores del conocimiento que necesitan vistas homogéneas e integradas para interpretarlos como conocimiento útil.

Más allá de la serie de factores mencionados, debe puntualizarse que el desarrollo de la Gestión del Conocimiento ha sido conducido, además, por prácticas y desarrollos en la gestión de datos e información. Se desarrolló primero en áreas intensivas en conocimiento, tales como la consultoría, investigación y desarrollo, y compañías farmacéuticas siendo luego adoptada por otras áreas (Davenport y Grover, 2001).

En resumen, el desarrollo de este campo ha sido impulsado por los cambios en el desarrollo económico y las estructuras organizacionales, sociales y tecnológicas. Las empresas enfrentan en la actualidad un entorno cada vez más competitivo y a medida que las compañías reducen su tamaño para adaptarse a este entorno deben ser capaces de reducir costos. Sin embargo, a menos que hayan capturado el conocimiento de sus empleados, una consecuencia directa de la reducción de tamaño puede ser la pérdida de información crítica. A medida que la tasa de rotación de empleados crece en el mercado laboral actual, es

probable que las organizaciones pierdan acceso a grandes cantidades de conocimiento crítico. Además, como las compañías se vinculan internacionalmente, las barreras geográficas pueden afectar el intercambio de conocimiento y dificultar el fácil acceso a la información. Estos y otros factores están forzando a las empresas a explorar mejores métodos para la Gestión del Conocimiento.

Por todo lo anterior, podría decirse que uno de los raros consensos en el dominio de la Gestión del Conocimiento es que el conocimiento ahora es percibido como un bien organizacional y de producción, un patrimonio valioso a ser gestionado y por lo tanto existe la necesidad de mejores herramientas y métodos para asistir en dicha gestión.

Una vez establecidos los factores que hacen necesaria una Gestión del Conocimiento eficiente, se analizará la naturaleza de este fenómeno.

3. La Gestión del Conocimiento

Debido a que la Gestión del Conocimiento es todavía una disciplina emergente, aún se están presentando trabajos representativos en diversas facetas del área. Dentro de la literatura que ha proliferado en este campo se encuentran algunas piezas que arrojan cierta luz sobre las teorías y la práctica. Kogut, Zander y Grant enlazaron la Gestión del Conocimiento con la teoría organizacional y desarrollaron la teoría de la empresa basada en el conocimiento (Grant, 1996; Kogut y Zander, 1992, 1996). Teece ata la Gestión del Conocimiento a la gestión estratégica y traduce las teorías de las competencias claves y las teorías organizacionales/industriales a una visión basada en conocimiento de la estrategia para la obtención de ventajas competitivas (Teece, 2001). Cook y Brown tratan el conocimiento explícito, tácito, individual y grupal como cuatro formas distintas y equivalentes de conocimiento, generando un nuevo interés en estudiar epistemología en el contexto de gestión (Cook y Brown, 1999).

La Gestión del Conocimiento organizacional formalmente conlleva gestionar los recursos de conocimiento para facilitar el acceso y reutilización del mismo, típicamente usando tecnologías de información de avanzada. Es formal en el sentido de que el conocimiento es clasificado y categorizado de acuerdo a un modelo preespecificado, pero al mismo tiempo evolutiva en bases de conocimiento y de datos estructurados y semi estructurados. El propósito dominante de la Gestión del Conocimiento organizacional es hacer que el conocimiento sea accesible y reutilizable para la organización.

Los sistemas de Gestión de Conocimiento integran componentes existentes tanto a nivel de contenido como de infraestructura, poniendo en contacto a las personas y a los sistemas de información asociados con tareas intensivas de conocimiento colaborativas. Las herramientas que manejan contenido codificable y explícito (sistemas de gestión de proyectos, data warehouses, etc.) como también componentes que permiten compartir y distribuir contenido contextualizado o tácito (por ejemplo, pizarrones virtuales, herramientas de razonamiento basados en casos y sistemas de mapeo conceptual) deben ser temporal y espacialmente integrados. El conocimiento contenido en un lugar debería ser accesible en forma transparente desde otras locaciones independientemente de los protocolos o el formato (Tiwana, 2001).

Los recursos de conocimiento varían de acuerdo a las industrias y aplicaciones particulares, pero generalmente incluyen manuales, cartas, resúmenes de documentos a clientes, noticias, información de clientes, inteligencia sobre competidores, conocimiento derivado de los procesos de trabajo, etc. Un amplio rango de tecnologías se están usando para implementar soluciones tendientes a la Gestión de Conocimiento: e-mail, base de datos y data warehouses; sistemas de soporte de grupos; navegadores (browsers) y motores de búsqueda; intranets e Internet, sistemas expertos y basados en conocimiento y agentes inteligentes. Como ejemplo pueden mencionarse Autonomy Inc. (www.autonomy.com), BackWeb Technologies (www.backWeb.com), Communispace (www.communispace.com), Docushare Xerox Corporation (www.xerox.com), Eugenia Unity Desktop (www.engenia.com), Excalibur

Technologies Corporation (www.excalib.com), Hummingbird (www.pcdocs.com), Intraspect Software, Inc. (www.intraspect.com), MAGI (www.projectmagi.com), Open Text Corporation (www.opentext.com), Plumtree Software (www.plumtree.com), SageMaker (www.sagemaker.com) y Tacit Knowledge (www.tacit.com).

En inteligencia artificial, las bases de conocimiento son generadas para ser usadas por expertos y sistemas basados en conocimiento, donde las computadoras usan inferencias basadas en reglas para contestar las consultas de los usuarios. Aunque la adquisición de conocimiento para la inferencia por computadora es todavía importante, los esfuerzos de Gestión del Conocimiento más recientes tienen como objetivo hacer disponible el conocimiento para consumo directo de las personas o el desarrollo de software para procesar ese conocimiento.

Históricamente, la Gestión del Conocimiento ha sido orientada a un solo grupo (gerentes) a través de lo que se ha referido como Sistemas de Información Ejecutivos. Un Sistema de Información Ejecutivos contiene un conjunto de herramientas cuyo objetivo es dar soporte a la toma de decisiones empresariales.

Sin embargo, si los ejecutivos necesitan acceso a la información y al conocimiento, es probable que también sus empleados tengan interés y necesidad de esos mismos recursos. Además, la tecnología de Gestión del Conocimiento es idealmente conveniente para grupos no gerenciales tales como soporte a clientes, I&D, etc.

Entre otras cosas los sistemas organizacionales deben comenzar a diferenciar entre la gestión de la información y la Gestión del Conocimiento. La gestión de la información se realiza generalmente estructurando la información, usualmente en la forma de una base de datos. En estas estructuras, la información puede ser almacenada, ordenada y recuperada. El conocimiento, por otro lado, requiere un sistema que no solamente pueda almacenar el conocimiento existente como información, sino que también pueda recuperar y usar esa información como conocimiento cuando sea necesario, especialmente durante la formulación

de un problema. De esta manera, puede crearse nuevo conocimiento a partir de conocimiento existente en combinación con nueva información. Adicionalmente, tal sistema abarcativo de Gestión de Conocimiento debe dar soporte a la utilización del conocimiento tácito de los miembros de una organización manteniendo información del área de experticia de los miembros y proveyendo canales de comunicación entre los miembros de la organización (Hall y col., 2001).

De forma similar a lo que sucede cuando se intenta encontrar una definición de conocimiento, generalmente no está claro lo que implica la Gestión del Conocimiento y esta es la razón por la que se encuentran tantas definiciones (Hlupic y col., 2002; Snyder y Wilson, 2002). De acuerdo a O'Leary la Gestión del Conocimiento puede ser definida como los esfuerzos hechos por la organización para capturar el conocimiento, convertir el conocimiento personal (tácito) en conocimiento disponible grupalmente (explícito), conectar personas con personas, personas con conocimiento, conocimiento con conocimiento, y medir el conocimiento para facilitar la gestión de recursos y para ayudar a entender su evolución (O'Leary, 2002). Estas actividades son restringidas o facilitadas por factores que influyen la Gestión del Conocimiento (Joshi, 2001).

Ponelis y col. ven a la Gestión del Conocimiento como una nueva dimensión de la Gestión de Información Estratégica (Ponelis y Fair-Wessels, 1998). Smith, por su parte, describe al objetivo de la Gestión del Conocimiento como el mejoramiento del desempeño organizacional permitiendo que los individuos capturen, compartan y apliquen su conocimiento colectivo para tomar decisiones óptimas en tiempo real (Smith y Farquhar, 2000). Thomas y col. mencionan la necesidad de aumentar la visión típica de la Gestión del Conocimiento como un problema de captura, organización y recuperación de información con una perspectiva que también reconozca el rol de la cognición humana cuando el conocimiento está situado en contextos de trabajo sociales (Thomas y col., 2001). King y col. reconocen el fuerte rol que juega la tecnología de información en la Gestión del Conocimiento e identifican

aplicaciones clave en esta área que están basadas en infraestructuras de tecnologías de información existentes (King y col., 2002): repositorios de conocimiento, sistemas de mejores prácticas y lecciones aprendidas, redes de expertos, y comunidades de práctica (Hackbarth y Grover, 1999; McDermott, 1999; Zack, 1999). Skyrme sugiere que la Gestión del Conocimiento es la gestión intencionada y sistemática del conocimiento vital junto con sus procesos asociados de creación, recolección, organización, difusión, uso y explotación de ese conocimiento (Skyrme, 2003).

La Gestión del Conocimiento es, en definitiva, la gestión de los activos intangibles que generan valor para la organización. La mayoría de estos intangibles tienen que ver con procesos relacionados de una u otra forma con la captación, estructuración y transmisión de conocimiento. Está constituida por el conjunto de procesos y sistemas que permiten que el capital intelectual (experiencia aplicada, tecnología organizacional, destrezas profesionales, o aportaciones no materiales que se entienden como el principal activo de las organizaciones) de una organización aumente de forma significativa, mediante la gestión de sus capacidades de resolución de problemas de forma eficiente, con el objetivo final de generar ventajas competitivas sustentables en el tiempo.

Más específicamente, la Gestión del Conocimiento tiene por objeto obtener, organizar, compartir y enriquecer el conocimiento relevante fundamentalmente con miras al rediseño personal y organizacional. Además se persigue la creación de nuevo conocimiento para generar visiones creativas y transformadoras.

La Gestión del Conocimiento no es gestionar la información, aunque se le parece y sobre todo, gestionar el conocimiento no es implementar una plataforma informática, aunque ésta puede convertirse en un gran facilitador de la Gestión del Conocimiento. En este sentido, un Sistema de Gestión del Conocimiento se define como una infraestructura creada por la organización para implantar los procesos y los procedimientos que, actuando sobre una base de datos, de información y de conocimientos, permitan la utilización del capital intelectual de

la organización para generar los servicios y los productos necesarios para contribuir al aumento de la competitividad de la organización.

A los efectos del alcance de esta tesis, se establece que la Gestión del Conocimiento organizacional es *una disciplina emergente que tiene como objetivo generar, compartir y utilizar el conocimiento tácito y explícito existente en un determinado espacio para dar respuesta a las necesidades de los individuos y de las comunidades dentro de la organización.*

4. Conclusiones

En este capítulo se han analizado los diferentes factores que hacen de la Gestión del Conocimiento organizacional una necesidad. Los mismos van desde el reconocimiento del conocimiento como un recurso de negocios generador de ventajas competitivas hasta el análisis de factores externos relacionados con la globalización y los cambios sociales y tecnológicos.

El ámbito empresarial demandaba y demanda un alto índice de renovación para ser competitivo con respecto al resto de las empresas que existen en el mercado, por lo que la innovación se hace condición indispensable para la supervivencia y crecimiento empresarial de una firma.

En este contexto, la Gestión del Conocimiento consiste en la capacidad de generar nuevos conocimientos a partir de datos y experiencias, además de organizar, distribuir y ordenar los conocimientos ya existentes en la misma organización. Es el proceso por el que una organización innova y compite con las demás de manera más eficiente.

La Gestión del Conocimiento no es un fin en sí misma sino un medio para lograr objetivos. La implementación de soluciones para la Gestión del Conocimiento debe responder a una estrategia corporativa de esta época, que se caracteriza por un entorno tremendamente cambiante, una sobreabundancia de información y oportunidades, limitación de recursos y un incremento notable de la inversión en empleados y en información.

En el capítulo siguiente se analizan los paradigmas y modelos para la Gestión del Conocimiento propuestos en la bibliografía, los cuales deberían dar el marco adecuado para la implementación de sistemas de gestión de conocimiento.

Paradigmas de Gestión del Conocimiento

1. Introducción

Las organizaciones necesitan entender cómo los diferentes enfoques de Gestión del Conocimiento se relacionan con la estrategia de negocios e interactúan con otro tipo de iniciativas para aprovechar los efectos sinérgicos de esta actividad. Algunos marcos estratégicos existentes para la Gestión del Conocimiento proveen conceptualizaciones relacionadas con el conocimiento como fuente de ventaja competitiva pero no consolidan las implicaciones prácticas (Von Krogh y col., 2000).

Es natural que diferentes grupos e individuos dentro de la organización sostengan visiones diversas y generalmente conflictivas sobre qué tipo de conocimiento es importante, quién debería controlarlo, y cómo debería ser gestionado y distribuido este conocimiento. Tales tensiones están implícitas en la vida organizacional pero tienden a ser exacerbadas por los esfuerzos e iniciativas para gestionar recursos de conocimiento. Si no se legitiman y reconocen las diferentes perspectivas de conocimiento, las iniciativas de Gestión del Conocimiento se convierten en un “pararrayos” que atrae conflictos y confusión conceptual en lugar de valor para la organización (de Long y Seeman, 2000).

Existe una gran y creciente cantidad de literatura que enfatiza la importancia del conocimiento y aprendizaje organizacional. En esta tesis se han identificado cuatro enfoques paradigmáticos en la literatura asociada a la Gestión del Conocimiento organizacional:

- ✓ *Sistémico*: las organizaciones pueden ser entendidas como sistemas de conocimiento, que consisten en un *conjunto de procesos* de conocimiento interrelacionados y socialmente validados: creación, organización, almacenamiento, distribución y aplicación del conocimiento. Aquí, la Gestión del Conocimiento trata con preguntas sobre el diseño y reingeniería de estos procesos y sus interacciones y sobre el desarrollo de mecanismos de validación del conocimiento. El conocimiento organizacional es visto como una construcción social, embebido en el conocimiento tácito de los individuos, en las rutinas, la tecnología y la cultura (Holzner y Marx, 1979).

- ✓ *Orientado a recurso*: este enfoque se orienta a medir y auditar el conocimiento como *recurso organizacional*. Es en este paradigma donde surge la noción de Capital Intelectual, reconociendo la importancia tanto de los componentes estructurales como los humanos (Sveiby, 1997a).

- ✓ *Técnico*: en este caso el conocimiento es visto en un sentido *objetivo*, como algo que podemos recolectar, representar, manipular y transmitir usando computadoras o algún otro dispositivo técnico (Wijnhoven, 1998). Esta visión contrasta con la visión sistémica donde el conocimiento puede ser transmitido sólo en el sentido en que el mismo es “reconstruido” en el extremo receptor.

- ✓ *Organizacional*: en este paradigma se presentan teorías en donde el conocimiento es visto como una *herramienta* que permite organizarse para entornos de negocios cada vez más turbulentos, donde la globalización de los productos y mercados implica una necesidad de innovación constante y flexibilidad a través de toda la organización (Drucker, 1988; Nonaka y Takeuchi, 1995).

Estos paradigmas suelen constituir el trasfondo para el desarrollo de modelos para la Gestión del Conocimiento y, en el campo académico, los investigadores han propuesto una variedad de modelos para ayudar a entender la problemática asociada. Pueden distinguirse dos tipos principales de modelos: descriptivos y prescriptivos. Los modelos descriptivos tratan de caracterizar la naturaleza de la Gestión del Conocimiento, mientras que los modelos prescriptivos presentan metodologías a seguir en la conducción del proceso de Gestión del Conocimiento. Estos modelos se describen con más detalles en las secciones 2 y 3. En relación a los enfoques comúnmente utilizados en la implementación de estos modelos en la sección 4 se describen el de repositorio y el de red. Derivados del análisis de los modelos y de los enfoques de implementación, en la sección 5, se presentan las actividades básicas de la Gestión del Conocimiento. Finalmente, en la sección 6, se presentan las conclusiones del capítulo.

2. Modelos Descriptivos de la Gestión del Conocimiento

Entre los modelos descriptivos pueden mencionarse los de Wiig, Leonard-Barton, Arthur Andersen & APQC, Choo y Van der Spek-Spijkervet (Wiig, 1993a) (Leonard-Barton, 1995) (Andersen, 1996) (Choo, 1996) (Van der Spek y Spijkerver, 1997). Estos modelos generales difieren no sólo en su foco sino también en la amplitud y profundidad con la que caracterizan la naturaleza de la Gestión del Conocimiento. A continuación se presentan los modelos en orden cronológico.

2.1. Modelo de Pilares para la Gestión del Conocimiento

El modelo de Wiig se enfoca en la gestión e identifica las principales funciones necesarias para la Gestión del Conocimiento presentando lo que él denomina los tres pilares de la Gestión del Conocimiento (Wiig, 1993a). Estos pilares representan las tres funciones principales y necesarias para gestionar el conocimiento. Como se observa en la Figura III.1., estos pilares se basan en un entendimiento amplio de lo que es la creación, manifestación, uso y transferencia de conocimiento.

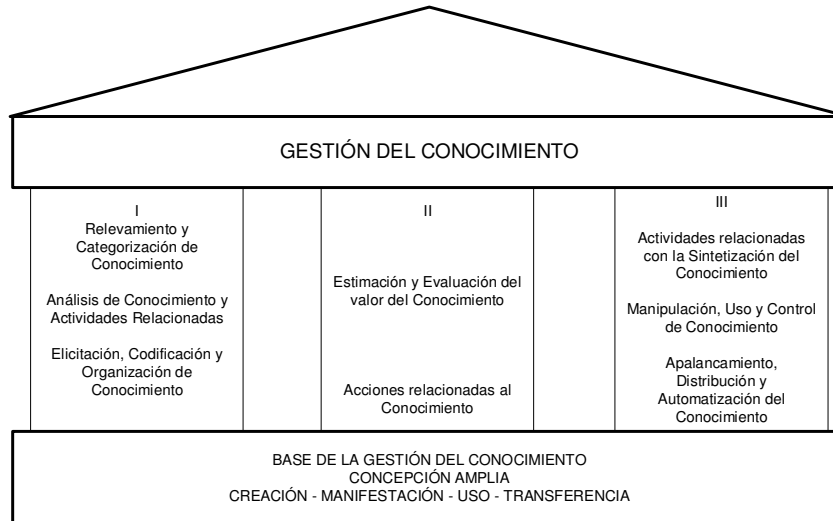


Figura III.1. Modelo de los Pilares de Gestión del Conocimiento (Wiig)

El pilar I está relacionado con la exploración del conocimiento y su adecuación. El modelo identifica varios componentes de esta función: investigar y categorizar el conocimiento; analizar el conocimiento y otras actividades relacionadas con el conocimiento tales como la elicitación, codificación, organización, etc. El pilar II involucra tasar y evaluar el valor del conocimiento y de sus actividades relacionadas. El pilar III se enfoca en la conducción de las actividades de Gestión de Conocimiento. Esta función tiene tres

componentes/actividades relacionadas con la síntesis del conocimiento; manipulación, uso y control de conocimiento y apalancamiento, distribución y automatización del conocimiento.

2.2. Modelo de las Habilidades Centrales y Construcción del Conocimiento

El modelo de Leonard-Barton se centra en la interacción de las capacidades tecnológicas de la organización y las actividades de desarrollo de conocimiento (Leonard-Barton, 1995). El marco comprende cuatro habilidades principales y cuatro actividades de construcción del conocimiento las cuales se afirma, son esenciales en una organización basada en conocimiento.

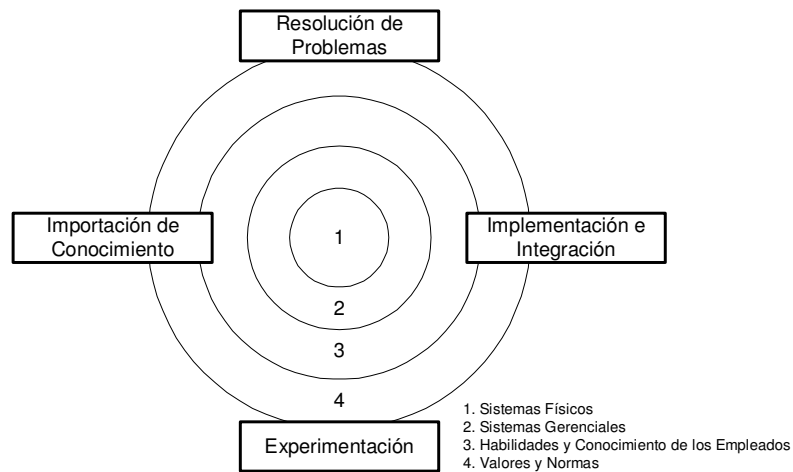


Figura III.2. Modelo de las Habilidades Centrales y la Construcción de Conocimiento (Leonard-Barton)

Como se observa en la Figura III.2, existen cuatro actividades de construcción de conocimiento que rodean las cuatro habilidades principales: solución creativa y compartida de problemas (para la producción de los productos actuales); implementación e integración de nuevas metodologías y

herramientas (para mejorar las operaciones internas); experimentación y prototipado (para construir habilidades futuras) y la importación y absorción de tecnologías más allá de los límites del conocimiento de la organización. Estas son actividades de creación y difusión de conocimiento.

Las actividades de construcción de conocimiento están influenciadas por lo que Leonard-Barton denomina habilidades centrales. Afirma que *las habilidades centrales constituyen una ventaja competitiva para una organización, ellas se construyen a través del tiempo y no pueden ser imitadas fácilmente*. Las cuatro habilidades centrales identificadas en este modelo son sistemas físicos (competencias acumuladas en sistemas materiales que son construidas a través del tiempo tales como bases de datos, maquinaria y software), habilidades y conocimiento de los empleados, sistemas gerenciales (rutinas organizadas que dirigen la acumulación y distribución de recursos, creando los canales a través de los cuales es accedido el conocimiento) y los valores y normas de la organización (que determinan el tipo de conocimiento buscado y alimentado, y los tipos de actividades de creación de conocimiento toleradas y fomentadas dentro de la organización). Las primeras dos involucran reservorios de conocimiento dinámico más los medios para manipularlos. Los últimos dos son mecanismos de canalización y control de conocimiento.

2.3. Modelo de Gestión de Conocimiento Organizacional

El modelo de Arthur Andersen & APQC provee las bases para la conducción de un proceso de benchmarking de Gestión de Conocimiento entre organizaciones y también dentro de la propia organización (Andersen, 1996). Esta organización ha propuesto un modelo que comprende siete procesos de Gestión del Conocimiento que pueden operar sobre el conocimiento organizacional. Como puede verse en la Figura III.3 estos procesos son creación, identificación, recolección, adaptación, aplicación, organización y distribución. La naturaleza del conocimiento que procesan no es caracterizada en este modelo,

ni tampoco se caracteriza la naturaleza de los procesos en sí mismos. El modelo identifica cuatro facilitadores organizacionales que promueven la implementación de los procesos de Gestión del Conocimiento: liderazgo, medición, cultura y tecnología. El modelo tampoco detalla la naturaleza de estos facilitadores.

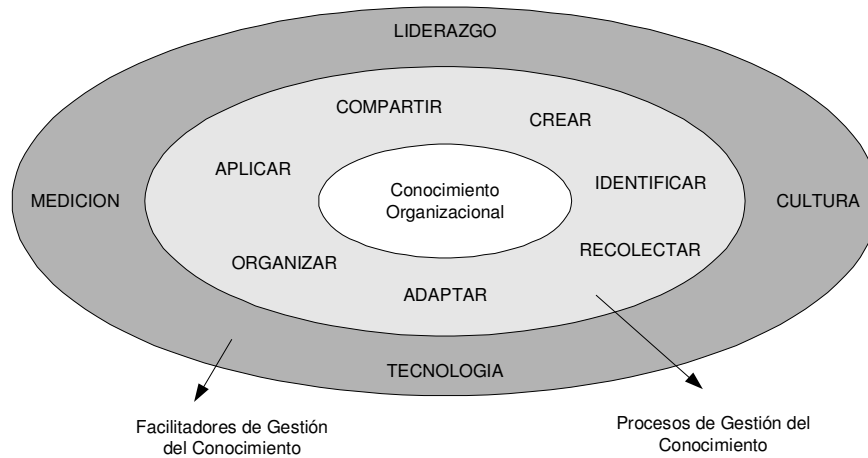


Figura III.3. Modelo de Gestión del Conocimiento Organizacional (Arthur Andersen & APQC)

2.4. Modelo de la Organización que Conoce

Choo presenta en su modelo a la *organización que conoce* tal y como se observa en la Figura III.4. (Choo, 1996). De acuerdo a este modelo, una organización usa información de forma estratégica para la comprensión, la creación del conocimiento y la toma de decisiones. Estos tres procesos están enlazados como un continuo de actividades de información anidadas que definen a una organización que posee la información y el conocimiento para actuar inteligentemente. El modelo usa los términos conocimiento e información sin comentar la existencia o naturaleza de cualquier distinción entre ellos.

Durante la fase de comprensión, una organización trata de entender su entorno cambiante. Este proceso está relacionado con la comprensión de cómo la gente en la organización interpreta la información con el fin de superar la incertidumbre del entorno.

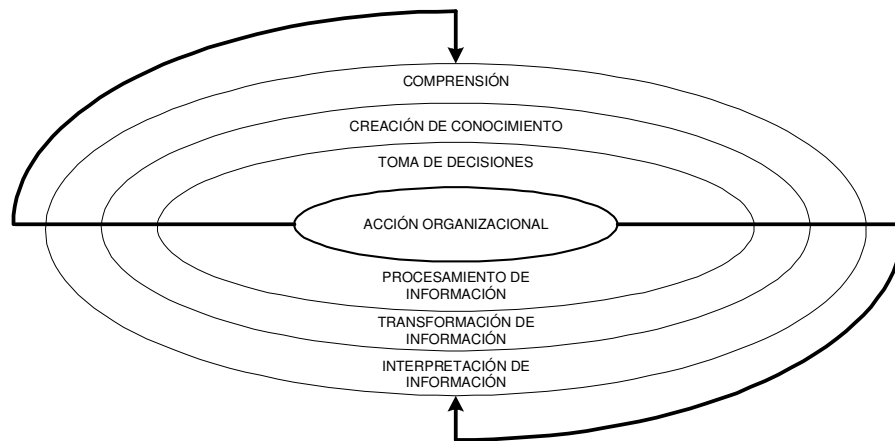


Figura III.4. Modelo de la Organización que Conoce (Choo)

Durante la creación del conocimiento, una organización crea nuevo conocimiento en el curso de la innovación. Este proceso está relacionado con la comprensión de cómo la información es transformada en nuevo conocimiento dentro de la organización. Finalmente, este modelo ve a la toma de decisiones como un proceso que trata de comprender cómo una organización usa la información para hacer frente a la incertidumbre en las tareas.

2.5. Modelo de los Estadios de la Gestión del Conocimiento

Finalmente, el modelo de Van der Spek y Spijkerver identifica un ciclo de cuatro estadios de Gestión del Conocimiento: conceptualización, reflexión, actuación y retrospección (Van der Spek y Spijkerver, 1997). Como puede verse

en la Figura III.5., estos estadios controlan las operaciones básicas del conocimiento. El estadio de conceptualización se enfoca en ganar comprensión de los recursos de conocimiento. Esto es alcanzado mediante la investigación, clasificación y modelado del conocimiento existente. Durante el estadio de reflexión, el conocimiento conceptualizado es evaluado usando una variedad de criterios, se establecen requerimientos de mejora y se planifica un proceso para llevarlos a cabo. Durante el estadio de actuación se toman acciones para mejorar el conocimiento. Esto involucra desarrollar nuevo conocimiento más la distribución y combinación de este conocimiento desarrollado. El último estadio, la retrospectiva, reconoce los efectos del estadio de actuación, evalúa los resultados alcanzados en esa etapa y compara la situación anterior y la nueva.

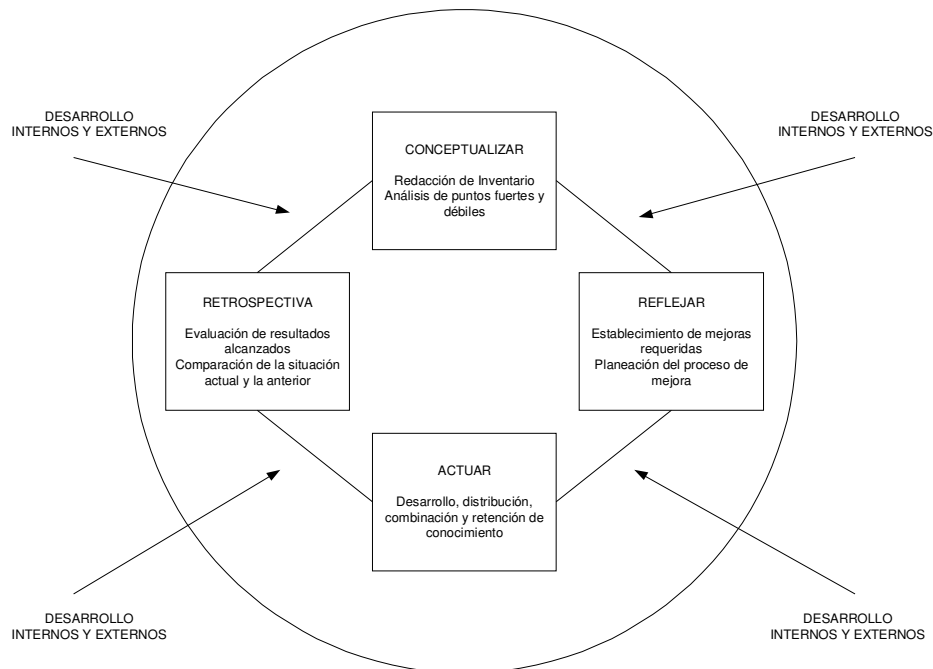


Figura III.5. Modelo de los Estadios de Gestión de Conocimiento (Van der Spek-Spijkervet)

La configuración de los estadios de Gestión del Conocimiento está orientada hacia un ciclo de solución de problemas. Por lo tanto, esta configuración puede ser vista como una forma de coordinar las actividades de manipulación del conocimiento dentro de un episodio que involucre la solución de un problema. Los estadios en el ciclo son impactados por desarrollos internos

y externos. Los factores internos que impactan la organización de la Gestión del Conocimiento son la cultura, la motivación de los empleados, la organización, la gestión y las tecnologías de la información. Los factores externos son reconocidos como influencias pero no se ofrecen ejemplos de los mismos en el marco.

3. Modelos Prescriptivos de la Gestión del Conocimiento

Entre los modelos prescriptivos, aquellos que intentan presentar metodologías a seguir en la conducción del proceso de Gestión del Conocimiento, se encuentran los propuestos por Sveiby, Petrash, Nonaka, Szulanski y Alavi (Sveiby, 1997b) (Petrash, 1996) (Nonaka, 1994) (Szulanski, 1996) (Alavi, 1997).

3.1. Modelo de Bienes Intangibles

El modelo de Sveiby se enfoca en la caracterización y medición de los bienes organizacionales intangibles (particularmente el conocimiento) (Sveiby, 1997b). Como se observa en la Figura III.6, el modelo comprende tres elementos: estructuras externas, estructuras internas y competencias de los empleados. Las estructuras externas incluyen relaciones con proveedores y clientes, nombres de marcas, marcas registradas y la imagen y reputación de la firma.

Las estructuras internas incluyen patentes, conceptos, modelos, sistemas de información y cultura organizacional. La competencia de los empleados consiste en las habilidades y bases de conocimiento de los individuos dentro de la organización.



Figura III.6. Modelo de Bienes Intangibles (Sveiby)

Los empleados usan sus habilidades y sus bases de conocimiento para actuar en una amplia variedad de situaciones para poder crear bienes tangibles e intangibles. Cuando las competencias de los empleados se dirigen hacia entidades fuera de la organización, entonces se consideran que resultan en estructuras externas, si esos esfuerzos se dirigen hacia adentro, se considera que crean estructuras internas.

3.2. Modelo de Capital Intelectual

El modelo de Petrash se centra en la caracterización y medición del *capital intelectual* de la organización, esto es, está orientado a la identificación de recursos de conocimiento (Petrash, 1996). El modelo involucra tres tipos de recursos organizacionales que se denominan capital intelectual: capital humano, capital organizacional y capital cliente. El capital humano es el conocimiento que cada individuo tiene y genera. El capital organizacional es el conocimiento que ha sido capturado/institucionalizado como la estructura, proceso y cultura de una organización. El capital cliente es la percepción del valor obtenido por los clientes cuando hacen negocios con un proveedor de bienes o servicios.

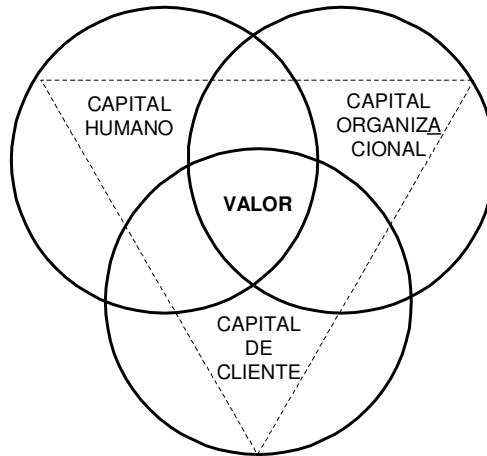


Figura III.7. Modelo de Capital Intelectual (Petrasch)

Como se observa en la Figura III.7, este modelo reconoce que las relaciones entre los tres tipos principales de capital intelectual conducen a resultados financieros, es decir, valor. Las líneas punteadas representan la gestión de los bienes intelectuales. Maximizando las interrelaciones entre estos tres tipos de capital intelectual se incrementa el espacio de creación de valor. En la Figura III.7 esto se representa creando el máximo solapamiento entre los tres anillos que representan el capital.

3.3. Modelo de Conversiones de Conocimiento

El modelo de Nonaka presenta el proceso de creación del conocimiento como una interacción entre dos tipos de conocimiento (tácito y explícito) (Nonaka, 1994). Como se observa en la Figura III.8, el modelo identifica cuatro tipos de conversión de conocimiento que conducen a la creación de conocimiento: socialización, externalización, internalización y combinación.

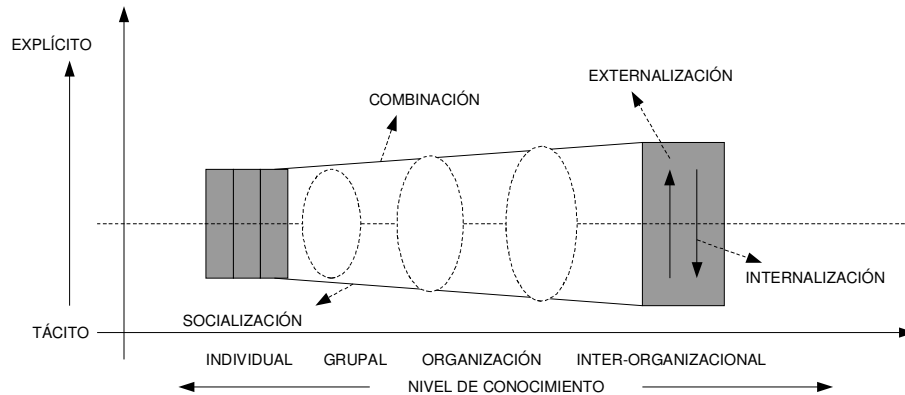


Figura III.8. Marco de Conversiones de Conocimiento (Nonaka)

Estas conversiones están basadas en una dicotomía entre los modos tácito y explícito de conocimiento y serán descritas con más detalle en el Capítulo VI.

3.4. Modelo de la Transferencia del Conocimiento

El modelo Szulanski se enfoca en la identificación de las barreras que existen para la transferencia de conocimiento en una organización (Szulanski, 1996). El modelo analiza la *rigidez interna* de la transferencia de conocimiento, con el foco en la transferencia de buenas prácticas. La rigidez interna se refiere a la dificultad de transferir conocimiento dentro de una organización. La transferencia de las mejores prácticas dentro de la organización es vista como un proceso desplegable que consiste en estadios en los cuales los factores característicos no sólo aparecen en mayor o menor medida sino en ocurrencia de orden de creación. El modelo identifica cuatro estadios involucrados en la transferencia de conocimiento: iniciación, implementación, crecimiento e integración. También identifica cuatro factores que impactan la dificultad de la transferencia de conocimiento: características de la transferencia de conocimiento (ambigüedad causal e improbabilidad), características de la fuente

de conocimiento (falta de motivación y falta de fiabilidad percibida), características de los recipientes de conocimiento (falta de motivación, falta de capacidad de absorción y falta de capacidad de retención) y características del contexto (contexto organizacional árido y relaciones dificultosas).

El estadio de iniciación está comprendido por todos los eventos que conducen a la decisión de transferir. Primero, se reconoce una necesidad de conocimiento que dispara una búsqueda para satisfacer esa necesidad. Una vez que se identificaron la necesidad y la solución potencial a esa necesidad, entonces se explora la factibilidad de transferir tal conocimiento. El estadio de implementación comienza una vez que se ha tomado la decisión de transferir el conocimiento. En este estadio, los recursos de conocimiento fluyen entre la fuente y el receptor, se establecen lazos sociales entre la fuente y el receptor, la transferencia es personalizada para cubrir las necesidades del receptor y se toman los recaudos para evitar los problemas de transferencias anteriores. Estas actividades cesan una vez que el recipiente comienza a usar el conocimiento recibido. En el estadio de crecimiento un recipiente comienza a usar el conocimiento, intenta identificar y resolver problemas no esperados que surgen mientras usan el nuevo conocimiento y cumple (o excede) las expectativas de desempeño post transferencia. En el estadio de integración, el conocimiento transferido se convierte gradualmente en rutinario e institucionalizado.

3.5. Modelo del Proceso de Gestión del Conocimiento

Finalmente, el modelo de Alavi se centra en el uso de la tecnología para el logro de la Gestión del Conocimiento (Alavi, 1997). Este modelo define a la Gestión del Conocimiento como la creación, apalancamiento y distribución del conocimiento laboral (know-how) y bienes intelectuales por todos los individuos a lo largo de la firma para mejorar el servicio a clientes. El modelo de proceso de Gestión del Conocimiento consiste en una secuencia de seis fases como se observa en la Figura III.9: adquisición, indexado, filtrado, enlazado, distribución

y aplicación. La adquisición se refiere a la creación del conocimiento y desarrollo de contenido. Esto es conseguido mediante el destilado de las experiencias y lecciones aprendidas, recolectando, sinterizando e interpretando una gran variedad de conocimiento. Las siguientes tres fases (indexado, filtrado y enlazado) son denominadas actividades de gestión de librería e incluyen la visualización, clasificación, catalogación, integración e interconexión del contenido tanto de fuentes internas como externas. La fase de distribución incluye el empaque y distribución de conocimiento. La aplicación se refiere al uso del conocimiento que ha sido recolectado, capturado y distribuido para producir productos y servicios.

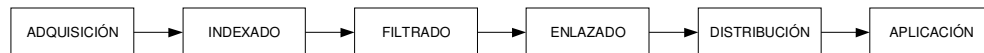


Figura III.9. Modelo del Proceso de Gestión del Conocimiento (Alavi)

Como se desprende de la descripción de los modelos realizada en este capítulo, el foco de cada modelo revela qué dimensión o aspectos se enfatizan o contemplan en cada caso. Cada uno de ellos se enfoca en ciertos elementos de la Gestión del Conocimiento pero ninguno parece resumirlos a todos. Esta cuestión se abordará en el capítulo siguiente presentando un modelo propio y más abarcativo para la Gestión del Conocimiento.

Ahora bien, además de modelos conceptuales para la Gestión del Conocimiento, existen enfoques relacionados a las tendencias en la implementación de dichos modelos, los cuales se analizan en la sección siguiente.

4. Enfoques para la Implementación de Modelos Conceptuales de Gestión del Conocimiento Distribuido

Existen dos tendencias relacionadas con el foco en la implementación de modelos de Gestión del Conocimiento: el enfoque de repositorio y el enfoque de red (Abidi, y col., 2005). En la siguiente sección se describe cada uno de ellos.

4.1. Enfoque de Repositorio

El enfoque de repositorio apunta a la codificación del conocimiento, es decir, la creación y mantenimiento de almacenes de conocimiento explícito. En este enfoque, el conocimiento es visto como un objeto que puede ser recolectado, almacenado, organizado y distribuido. Como tales, estos sistemas se orientan a la gestión del conocimiento explícito y primariamente a los aspectos de almacenamiento y recuperación de la Gestión del Conocimiento organizacional.

En el enfoque de repositorio de conocimiento, se establece una ubicación central para almacenar el conocimiento documentado. El conocimiento tácito fluye desde las personas al repositorio y desde el repositorio a las personas más que de persona a persona como se representa en la Figura III.10. Capturar el conocimiento tácito y explicitarlo es el objetivo principal del repositorio de conocimiento (Turban y col., 2002).

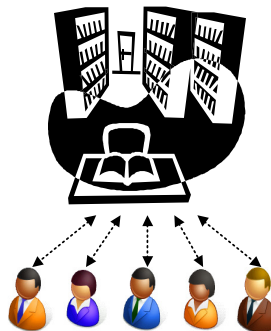


Figura III.10. Flujo de Conocimiento en el Modelo de Repositorio

Un repositorio de conocimiento puede tomar muchas formas y comprender la relación entre conocimiento e información es crucial para un repositorio de conocimiento. Como se mencionó en el Capítulo I, el conocimiento y la información comparten una cualidad intrínseca que los separa de los datos: el contexto. Marchand enfatiza aún más esta noción afirmando que el conocimiento se diferencia de la información en que siempre es personal (Marchand, 1998).

Las organizaciones que buscan preservar sus bienes intelectuales mediante la aplicación del enfoque de repositorio de conocimiento deben establecer tres procesos claves: selección, almacenamiento y actualización (Probst y col., 2000). En primer lugar, la organización debe seleccionar de entre la multitud de eventos, personas y procesos aquellos que vale la pena retener. Sólo el conocimiento que puede ser útil para otras personas en el futuro amerita la preservación. Como se mencionó antes en esta tesis, intentar almacenar toda clase de conocimiento sería antieconómico y en ocasiones imposible. Al mismo tiempo, debido a que no es posible predecir las necesidades de conocimiento futuras con absoluta precisión, la búsqueda de conocimiento útil no debería ser demasiado estrecha.

En segundo lugar, la organización debe almacenar su experiencia de una forma adecuada ya que la memoria humana puede ser transitoria y dinámica, cambiando con cada nueva experiencia. La tecnología ofrece una amplia variedad de opciones para el almacenamiento del conocimiento que son confiables y permanentes. Las organizaciones pueden en la actualidad aprovechar las ventajas de los artefactos digitales modernos para preservar su capital intelectual. Además, los costos decrecientes en los equipos de computación hacen que este almacenamiento sea asequible. Esta tendencia garantiza casi una capacidad de almacenamiento ilimitada, especialmente para que las grandes organizaciones construyan repositorios de conocimiento digitales. Las bases de datos y los data warehouses pueden convertirse en parte del sistema de Gestión de Conocimiento siempre y cuando puedan estructurarse de acuerdo a la idiosincrasia del conocimiento.

Finalmente, debe asegurarse que la memoria corporativa sea actualizada periódicamente. Si los trabajadores responsables fallan en mantener sus repositorios de conocimiento actualizados, la calidad de su contenido se hace cuestionable. Si esto ocurre, la confianza en el sistema de conocimiento decrece, reduciendo el soporte corporativo hasta que colapsa. Como consecuencia, la actualidad de los objetos almacenados en el repositorio es crítica para evitar el espiral descendente. Los repositorios de conocimiento requieren un staff dedicado para permanecer actuales y efectivos. Esta tarea nunca termina ya que el conocimiento organizacional está continuamente evolucionando y adaptándose a los cambios en el entorno.

Además de ciertas cuestiones de implementación que deben ser tenidas en cuenta relacionadas con la selección del conocimiento a almacenar, el tipo de almacenamiento y los medios y frecuencia de actualización del conocimiento almacenado, existen consideraciones tácticas que deben analizarse. Convertir el conocimiento tácito en artefactos de conocimiento explícito para una distribución más amplia requiere de confianza. Confiar a otros individuos con conocimiento personal, pensamientos y creencias debe ser percibido por todas las partes involucradas como un esfuerzo útil (Marchand, 1998).

Otro riesgo relacionado con la distribución del conocimiento tiene que ver con su actualidad. Una vez que el conocimiento personal es almacenado en artefactos tales como documentos, mensajes o base de datos, el mismo es congelado en el tiempo. Debido a la naturaleza dinámica del conocimiento, sin embargo, es altamente probable que el mismo caduque a través del tiempo. Decisiones y acciones importantes podrían ser ejecutadas basadas en conocimiento desactualizado o que puede ser inexacto luego de un período de tiempo. Esta amenaza acentúa la importancia de las prácticas de Gestión del Conocimiento dedicadas a monitorear la actualidad de los objetos almacenados.

Un desafío para los trabajadores relacionado con lo anterior es el tratamiento del conocimiento desactualizado. Los cambios ambientales pueden tener efectos impredecibles en los patrones operacionales de una organización. Cuando las experiencias, habilidades y prácticas previas ya no cubren los

desafíos actuales, las organizaciones deben olvidar la antigua forma de hacer negocios y adaptarse a la nueva situación. El problema aquí es decidir si los objetos de conocimiento almacenados, que en definitiva son bienes de negocios, pueden ser destruidos sin dañar el desempeño futuro. Además, considerando que el nuevo conocimiento organizacional solamente puede ser desarrollado sobre la base de conocimiento previo, la eliminación de conocimiento existente debe ser ejecutada cuidadosamente.

Por otro lado, la sobrecarga de conocimiento puede convertirse en uno de los peligros más frecuentes de los repositorios de conocimiento. Los avances tecnológicos en las comunicaciones han creado una red global de fuentes de información accesibles. Sin embargo, el factor principal en la actividad de creación de conocimiento (las personas) ha permanecido sin cambios. Las personas tienen una capacidad limitada para analizar grandes cantidades de información en periodos cortos de tiempo. Aún con la ayuda de dispositivos digitales tales como computadoras, la mayoría de las personas no tienen el tiempo o la atención para interpretar cada pieza de información relacionada a una tarea o tópico en particular. Sin el mantenimiento apropiado, los repositorios de conocimiento pueden crecer extremadamente rápido, dejando a los usuarios con la abrumadora tarea de buscar el conocimiento relevante en una pila de objetos almacenados.

De la misma manera en que los riesgos inherentes al enfoque de repositorio de conocimiento son evidentes, los beneficios asociados son palpables. El más notable es la preservación del conocimiento organizacional. Debido a que los bienes intangibles como el conocimiento se han convertido en esenciales para el éxito en los negocios, invertir en métodos confiables para asegurar estos recursos es ahora una obligación. Las experiencias y habilidades colectivas de las personas son el marco del capital intelectual en cualquier organización. El enfoque de repositorio de conocimiento puede aliviar o incluso eliminar la pérdida de capital intelectual.

Comparada con la memoria humana que puede ser limitada y transitoria, un repositorio de conocimiento bien diseñado puede ser extenso y confiable. El

desarrollo de una computación omnipresente hace de los artefactos digitales la elección natural para asegurar el capital intelectual. El costo decreciente de los dispositivos de almacenamiento permite que una sustancial cantidad de información pueda acumularse en repositorios de conocimiento digitales. Las computadoras permiten, al mismo tiempo, el procesamiento de grandes cantidades de información considerablemente más rápido de lo que las personas pueden hacerlo, reduciendo el impacto de la sobrecarga de información.

Finalmente, el enfoque de repositorio de conocimiento trae aparejado algunos dilemas éticos y desafíos de gestión. Surgen cuestiones relacionadas con la prescindibilidad de las personas una vez que su conocimiento tácito ha sido elicitado y asegurado por la organización. Un gerente utilitarista podría basarse en esto para conducir reducciones de personal, comprometiendo al mismo tiempo, la atmósfera de confianza esencial para cualquier comportamiento asociado a compartir conocimiento.

4.2 Enfoque de Red

La otra tendencia en la implementación de modelos de Gestión del Conocimiento, el enfoque de red, apunta a usar el poder de las tecnologías de información y comunicación para dar soporte al flujo de conocimiento en el ámbito organizacional y entre redes de dominios de conocimiento. En este enfoque el conocimiento permanece distribuido a través de los agentes que actúan dentro de los límites que han sido definidos de acuerdo a los requerimientos para un proyecto, un departamento o un negocio. Debido a que el conocimiento no está almacenado en una ubicación central, el énfasis en este enfoque está en las relaciones entre los actores involucrados, que pueden ser personas, grupos, uniones de organizaciones, comunidades o incluso sociedades (Seufert y col., 1999). Más que ser transferido verticalmente como en el enfoque de repositorio de conocimiento, el conocimiento se transfiere en dirección horizontal a través de interacción social y colaboración entre individuos dentro

de una estructura definida por el contenido y la forma de estas relaciones como se representa en la Figura III.11.



Figura III.11. Flujo de Conocimiento en el Modelo de Red

En este enfoque, las redes sociales actúan como una estructura dinámica donde el conocimiento que es valioso fluye de persona a persona en lugar de estar depositado en un repositorio central de forma estática. Como en el enfoque de repositorio, se requiere algo de conocimiento sobre el conocimiento (metadatos) para propósitos estructurales, pero esto es solamente necesario para crear un directorio para saber qué conocimiento detenta cada agente. Además, en lugar de basarse en tecnologías de data warehousing para capturar y almacenar el conocimiento, el enfoque de red enfatiza el uso de las tecnologías de comunicación (Turban y col., 2002).

De acuerdo a lo anterior, es evidente que el paso más importante en la implementación de un modelo basado en el enfoque de red es la codificación de la distribución del conocimiento a través de la red y el establecimiento de los medios para accederlo. En este enfoque, la principal fuente de conocimiento dentro de una organización es el conocimiento tácito de las personas involucradas. Sin embargo, este conocimiento tácito (tales como acciones individuales, experiencias, ideas y valores) no es fácilmente accesible independiente de la persona y el contexto. Por lo tanto, se necesita una red de relaciones para facilitar la transferencia de conocimiento de una persona, grupo u organización a la siguiente. Esta red debe ser establecida de forma tal que los razonamientos subjetivos, habilidades personales y conocimiento laboral (know how) puedan ser transmitidos tan fácilmente como los hechos objetivos.

Es evidente que el enfoque de red es dependiente de las relaciones e interacciones sociales entre los actores como el medio principal para distribuir conocimiento. Sin embargo, debido a la dificultad para transferir con precisión el conocimiento tácito, deben tenerse en cuenta varios factores cruciales cuando se implementa un enfoque basado en red. Como se mencionó antes en esta tesis, las ventajas estratégicas y competitivas del conocimiento para una organización dependen de la capacidad organizacional para generar, comprender y utilizar nuevo conocimiento. Esto trae aparejado la necesidad de que los gerentes se involucren con la tasa de disponibilidad, es decir, la tasa a la cual el nuevo conocimiento es accesible por todos, y la tasa de absorción, es decir, la tasa a la cual los que acceden al conocimiento son capaces de entenderlo, internalizarlo y actuar sobre ese conocimiento.

Como en el enfoque de repositorio, todas las partes involucradas deben percibir el hecho de confiar a otros individuos conocimiento personal, pensamientos y creencias como un esfuerzo útil. El gerente de conocimiento puede asegurar esta percepción mediante la creación de un entorno de confianza y también tomando un rol activo de facilitador en la guía de la calidad y utilidad del conocimiento (Marchand, 1998). Mediante la instrucción en el uso de imágenes, metáforas, ayudas visuales, etc. puede fomentar el refinamiento y la sistematización del conocimiento. Esto no sólo redonda en un conocimiento más útil para aquellos que lo reciben, ayudando en el proceso de internalización y acelerando la tasa de absorción, sino que también incrementa su aplicabilidad a través de toda la organización y lo hace más ampliamente disponible para su uso.

Otra consideración importante para la gestión en un enfoque de red es la integración de diferentes tipos de conocimiento. Debido a que la mayoría del conocimiento no está almacenado en un repositorio central, el gerente tendrá que combinar conocimiento previo con conocimiento nuevo en diferentes niveles. Los tipos individuales (tácito/explicito) tendrán que combinarse con tipos más grandes (grupales, organizacionales) como así también combinar las diferentes áreas de especialización (I&D, ventas, marketing, etc.) para cada uno de esos niveles. Cada una de estas áreas contribuirá con conocimiento con diferentes grados de

especificidad y relevancia, y con objeto de asegurar la utilidad, el gerente deberá estar constantemente evaluando y actualizando el flujo entre aquellos que están involucrados.

La ventaja principal del enfoque de red es que es fácil de usar y de implementar. En su nivel más simple, no son necesarios recursos tecnológicos adicionales porque el conocimiento permanece distribuido en las mentes de los individuos de la organización. Por lo tanto, la Gestión del Conocimiento en un enfoque de red consiste principalmente en mantener una especie de registro de quién sabe qué y cómo puede ser contactado si es necesario.

Otra ventaja del enfoque de red es la facilidad con la que una organización que ha implementado redes de conocimiento puede tomar ventaja de la mayor conectividad global entre los individuos y la organización. Con la creciente ubicuidad de Internet y de los dispositivos para acceder a ella, las redes de conocimiento ya no están confinadas a las intranets organizacionales. Mediante el enfoque en una estructura de conocimiento basada en la comunicación y las relaciones, una organización puede incorporar y combinar de forma más fácil conocimiento entrante desde fuentes externas. La habilidad para usar y explotar de forma más fácil el conocimiento, las capacidades y recursos distribuidos en redes, tanto dentro como fuera de la organización, le da a dicha organización una ventaja competitiva distintiva sobre aquellas que sólo están basadas en un enfoque de repositorio.

Sin embargo, el éxito de un enfoque de red es dependiente de varios factores que son mucho más difíciles de regular que en un enfoque de repositorio. Debido a que la tasa a la que es transmitido y utilizado el conocimiento es dependiente de las relaciones sociales, la Gestión del Conocimiento toma más un rol psicológico que técnico. Más que monitorear el ancho de banda de cables y routers, el gerente de conocimiento debe monitorear el lenguaje, la apertura, la creatividad y la tolerancia hacia las ideas de otros.

El gerente de conocimiento es responsable de mantener una cultura organizacional que fomente la voluntad de compartir conocimiento y que

requiere al mismo tiempo un entorno confortable y de confianza (Marchand, 1998). Para maximizar el flujo de conocimiento, el gerente también debe promover la participación en grupos multidisciplinarios eliminando los prejuicios normales y las líneas de división mentales que separan a los individuos de los diferentes departamentos previniendo, al mismo tiempo, la formación de islas de conocimiento.

4.3 Modelo de Red vs. Modelo de Repositorio

Resumiendo, la Tabla III.1 presenta una comparación de los dos modelos en las áreas claves.

| | REPOSITORIO DE CONOCIMIENTO | RED DE CONOCIMIENTO |
|---|-----------------------------|---------------------|
| Costo de Implementación | Alta | Baja |
| Longevidad Intelectual | Permanente | Transitoria |
| Flujo de Conocimiento | Vertical | Horizontal |
| Foco de la Gestión del Conocimiento | Técnico | Sociológico |
| Control de la Calidad de la Información | Normalizado | Variable |
| Capacidad de Almacenamiento | Ilimitada | Limitada |
| Configuración del Sistema | Centralizado | Distribuido |
| Interacción del Usuario | Impersonal | Personal |

Tabla III.1. Comparación de los Modelos de Red y Repositorio de Conocimiento

El mayor beneficio de un repositorio de conocimiento es su posteridad. A medida que los empleados, programas y productos vienen y van, el repositorio permanece tanto como un archivo institucional y como un recurso que puede ser usado a través de toda la organización. Esta ventaja viene con un precio. Debido a la naturaleza dinámica del conocimiento, cualquier sistema dedicado a la preservación del conocimiento está en riesgo de convertirse en obsoleto o lleno de conocimiento inútil. Los procedimientos de Gestión del Conocimiento son

esenciales para mantener la actualidad y la relevancia del repositorio de conocimiento, adicionando costos y complejidad no requeridos en otros modelos.

La centralización del conocimiento en un repositorio lo puede hacer más vulnerable al sabotaje y la destrucción si no se prevén las contramedidas adecuadas en el diseño. La seguridad del repositorio en general es incuestionable, pero los costos para prevenir o aliviar todas las amenazas deben evaluarse contra los costos de perder integridad de conocimiento.

Por otro lado, el beneficio principal del enfoque de red es su facilidad de implementación, especialmente dadas las modernas tecnologías de comunicación. La naturaleza distribuida del conocimiento implica que nuevos nodos pueden conectarse al sistema. La interacción con el sistema es más directa y personal, haciendo que sea más fácil para la organización juzgar si el sistema es exitoso.

Estos beneficios deben ser balanceados con factores tales como la capacidad de almacenamiento limitada de una red de conocimiento. Además, la naturaleza transitoria de los empleados, que pueden dejar la organización o retirarse, implica que el sistema no es permanente. La red debe ser capaz de actualizar sus registros a medida que los miembros organizacionales ganan nuevas habilidades. Existen también dificultades que pueden generarse tratando de codificar el conocimiento tácito de los empleados. La misma configuración distribuida que hace atractivo el sistema de red también hace que el mantenimiento de la seguridad del conocimiento sea todo un desafío.

5. Actividades de Gestión del Conocimiento

Aunque es ampliamente reconocido que la Gestión del Conocimiento está rodeada de vaguedad y ambigüedad parece existir cierta consistencia en relación a las actividades y componentes involucrados (Patel y Hlupic, 2002). Como se observa en la Tabla III.2, las primeras representan las actividades primarias que pueden tomar lugar durante el desarrollo de la Gestión del Conocimiento y

consisten, a grandes rasgos, en la creación, la representación y la distribución del conocimiento pudiendo ser consideradas ya sea individualmente o en varias combinaciones (Van Zolingen y col., 2001). Los segundos reflejan los aspectos de una organización que deben tenerse en cuenta para que un esfuerzo de Gestión del Conocimiento sea exitoso. Los componentes consisten de aspectos humanos, organizacionales y tecnológicos, y a diferencia de las actividades, no pueden existir separadamente. De hecho, los tres deben recibir la misma atención (Duffy, 2001).

| Gestión del Conocimiento | | | | | |
|--------------------------|----------------|--------------|------------------|---------------------------|-----------------------|
| Actividades | | | Componentes | | |
| Creación | Representación | Distribución | Aspectos Humanos | Aspectos Organizacionales | Aspectos Tecnológicos |

Tabla III.2. Actividades y Componentes de la Gestión de Conocimiento

Basados en una perspectiva orientada al conocimiento, muchos teóricos han sugerido varios tipos de capacidades organizacionales que serían esenciales para la Gestión del Conocimiento tal como se muestra en la Tabla III.3 (Lee y col., 2001). De dicha literatura pueden deducirse las siguientes consideraciones generales. En primer lugar, las organizaciones necesitan adquirir conocimiento crítico externamente como también crearlo internamente. Cohen y Levinthal enfatizan lo que denominan la *capacidad de absorción* para reconocer el valor del conocimiento nuevo y externo, asimilarlo y aplicarlo a fines comerciales para actividades de innovación organizacionales (Cohen y Levinthal, 1990). Kogut y Zander también definieron como una capacidad organizacional necesaria la capacidad de combinar el aprendizaje interno y externo (Kogut y Zander, 1996).

En segundo lugar, puede decirse que el objetivo final de la Gestión del Conocimiento es ganar ventajas competitivas y sostenerlas produciendo nuevos productos o servicios o mejorando los procesos organizacionales en términos de velocidad, calidad y costos (Junarkar, 1997; Quinn y col., 1996). En este sentido, Grant argumenta que, ya que la producción requiere la aplicación de varios tipos de conocimiento especializado, el rol primario de una organización es la integración del conocimiento (Grant, 1996).

Finalmente, cabe mencionar que el plan estratégico de una organización debería reflejar una visión dinámica de estas capacidades organizacionales ya que la Gestión de Conocimiento es una más de las actividades de gestión que se adaptan continuamente a los cambios en las necesidades del mercado.

| Autores | Vistas | Capacidades |
|---------------------------|---|------------------------------|
| (Nonaka, 1994) | La capacidad de creación del conocimiento introduciendo el modelo de creación del conocimiento y el modelo en espiral | Creación del Conocimiento |
| (Grant, 1996) | Capacidad organizacional para integrar conocimiento junto con la habilidad para realizar repetidamente una tarea productiva para crear valor en sus salidas | Integración del Conocimiento |
| (Kogut y Zander, 1996) | Habilidad organizacional para aprender nuevas habilidades a partir de la combinación de conocimiento interno y externo | Combinación de conocimiento |
| (Cohen y Levinthal, 1990) | Capacidad de absorción como una habilidad organizacional para reconocer, asimilar y aplicar conocimiento | Absorción de Conocimiento |
| (Quinn y col., 1996) | Desarrollo de la capacidad de gestionar el conocimiento organizacional de acuerdo a los cambios del entorno | Desarrollo de conocimiento |
| (Badaracco, 1991) | Habilidad organizacional para aprender o adquirir el conocimiento necesario de otras organizaciones | Enlaces de conocimiento |

Tabla III.3. Capacidades Organizacionales para la Gestión del Conocimiento

Basados en las definiciones de la Tabla III.3 y el análisis de los modelos descriptos, esta tesis sugiere que las capacidades organizacionales fundamentales

en la Gestión del Conocimiento se forman a partir de la acumulación, gestión y utilización del conocimiento organizacional como una ventaja competitiva sustentable. La acumulación de conocimiento organizacional puede alcanzarse mediante la adquisición de conocimiento desde fuentes externas y mediante creación interna. De todo lo anterior se desprende que las principales actividades de la Gestión de Conocimiento organizacional son la creación, la distribución, la búsqueda y recuperación de conocimiento.

5.1 Crear Conocimiento

Nonaka afirma que puede crearse nuevo conocimiento organizacional mediante cuatro procesos de conversión que involucran conocimiento tácito y explícito: socialización, externalización, combinación e internalización (Nonaka, 1994). La externalización del conocimiento se refiere a la conversión de conocimiento tácito en conocimiento explícito. Esta actividad puede verse como la transferencia de la experticia en la resolución de problemas desde una fuente de conocimiento, por ejemplo un experto de dominio, a un formalismo computacional (Abidi y col., 2005).

El segundo tipo de conversión de conocimiento, la socialización, se refiere a la creación de nuevo conocimiento tácito a partir de conocimiento tácito compartido. Se transfiere el conocimiento tácito de una persona a otra mediante un proceso activo y experimental donde el conocimiento es capturado a través de interacción directa, observando el comportamiento de otros y copiando sus comportamientos y creencias. Kimble y col. afirman que las comunidades de conocimiento son centrales para el mantenimiento del conocimiento tácito (Kimble y col., 2001).

La combinación de conocimiento se refiere a la creación de nuevo conocimiento mediante el intercambio de conocimiento explícito en la organización. El cuarto tipo de conversión de conocimiento, la internalización, toma lugar cuando el conocimiento explícito se convierte en tácito. En el

Capítulo VI se propone relacionar estos procesos con el aprendizaje, el entrenamiento y la tutoría.

5.2 Compartir Conocimiento

Alavi sugiere que una de las razones más importantes para enfocarse en la capacidad de compartir conocimiento es que la creación del conocimiento, en sí misma, no conduce a un mejor desempeño (Alavi, 2000). Antes bien, las compañías tienen que crear valor usando el conocimiento, y el conocimiento sólo puede ser utilizado si se comparte exitosamente. Por lo tanto, las organizaciones tienen que gestionar efectivamente el proceso de transferencia de conocimiento para utilizarlo como ventaja competitiva.

5.2.1. Contexto de Transferencia

De acuerdo a las opiniones de una serie de autores puede decirse que el proceso de transferencia de conocimiento está delimitado colectivamente por cinco componentes: contexto fuente, contexto receptor, contexto de conocimiento, contexto relacional y contexto situacional (Bresman y col., 1999; Cumming y Teng, 2003; Simonin, 1999; Szulanski, 1996; Zander y Kogut, 1994). Como se observa en la Figura III.12, los contextos fuente y receptor se refieren a los atributos de la fuente y receptor de conocimiento que pueden facilitar o impedir la transferencia de conocimiento. El contexto relacional se refiere a los atributos que caracterizan la relación entre la fuente y el receptor del conocimiento. El contexto de conocimiento se refiere a la naturaleza y caracterización del tipo de conocimiento que está siendo transferido. Finalmente, el contexto situacional se refiere a las características del entorno que rodean el proceso de transferencia de conocimiento (por ejemplo, modalidad de comunicación, proximidad física).

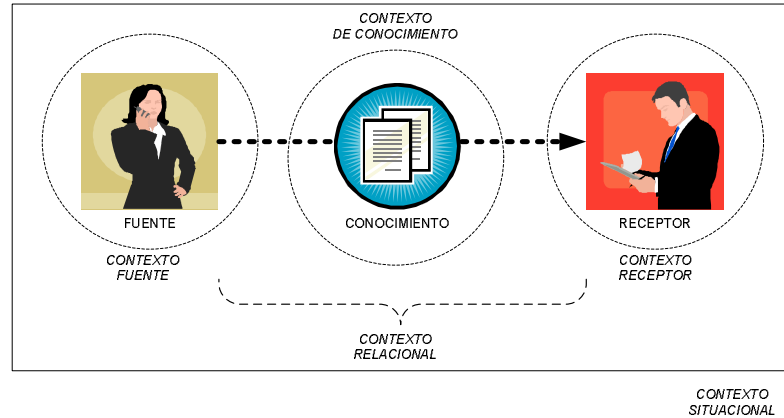


Figura III.12. Contextos de Transferencia de Conocimiento

Es importante tener en cuenta cada uno de estos contextos ya que la naturaleza de cada uno de ellos puede convertirlos en facilitadores o impedimentos en el proceso de compartir conocimiento.

5.2.2. Comunidades de Práctica y de Conocimiento

Una Comunidad de Práctica es un grupo de personas que tienen objetivos e intereses similares. Emplean prácticas comunes, trabajan con las mismas herramientas y se expresan con un vocabulario compartido. A través de tal actividad común, mantienen creencias y sistemas de valores similares. Las comunidades de práctica son mayormente informales y diferentes de las redes, equipos y unidades organizacionales. Una comunidad de práctica se define a través de tres dimensiones: sus objetivos, actividades y las capacidades (rutinas, artefactos, vocabulario, etc.) que desarrollan a través del tiempo. Las comunidades están en todas partes, a veces no reconocidas, a veces legitimadas y a veces reconocidas como un bien a ser fomentado. Este tipo de comunidades pueden reducir los tiempos de aprendizaje, incrementar la sensibilidad ante las demandas de los clientes, reducir el re-trabajo y evitar la “re-invencción de la rueda” (Voss y Schäfer, 2003).

En una Comunidad de Práctica tradicional, los miembros se encuentran personalmente en una base más o menos regular y usan el teléfono o e-mail para comunicaciones ad-hoc. Las comunidades que traspasan las fronteras organizacionales se apoyan en herramientas basadas en tecnologías de información y telecomunicación, tanto para comunicación sincrónica y asincrónica como para compartir documentos y repositorios en general. En estas comunidades, un rol crucial es el de moderador, que facilita las discusiones y la alimentación de los repositorios de conocimiento. Las comunidades de práctica pueden incluso desarrollar ontologías para su vocabulario compartido para categorizar el contenido de la base de conocimiento (Voss y Schäfer, 2003).

Una Comunidad de Práctica puede comenzar con una colección de personas que realizan actividades laborales similares. A medida que madura, puede tomar atributos que marcan la transformación de un grupo en una comunidad, una estructura auto-organizada con un propósito colectivo mantenido en su lugar por relaciones sociales. Últimamente, las comunidades de práctica han recibido justificadamente una gran atención cuando se trata de identificar el capital intelectual de una organización.

Algunas de las definiciones anteriores de Comunidad de Práctica enfatizan en diferente grado la centralidad de las actividades relacionadas a las tareas. Wenger por su parte, la define como un grupo de personas informalmente unidas por experticia compartida y una pasión por un proyecto común (Wenger y Zinder, 2000). Menville y Foot la definen como un grupo de profesionales informalmente unidos mediante la exposición a una clase de problemas y una persecución de soluciones comunes, y quienes por lo tanto, encarnan un almacén de conocimiento tácito (Menville y Foot, 1996).

También ha sido definida como (i) un grupo cuyos miembros se involucran regularmente en tareas de aprendizaje y distribución de conocimiento basadas en su interés común (Lesser y Storck, 2001) y (ii) una red informal e institucionalizada de profesionales gestionando dominios de conocimiento (Gongla y Rizzuto, 2001). Para enfatizar la importancia de las actividades relacionadas a las tareas de distribución de conocimiento en relación a su

existencia, en esta tesis se define a las comunidades de práctica como un grupo de personas que se unen para compartir información, conocimiento, entendimiento, experiencia y herramientas dentro de un área de práctica común.

Dichas comunidades se desarrollan externamente a través de medios tales como las acciones de organizaciones *sponsors* o internamente como resultado de acciones de sus miembros ya que con el incremento de la fuerza de trabajo móvil de la actualidad, los individuos están generalmente más alineados con su identidad profesional que con su pertenencia a una organización (Allee, 2000).

Este tipo de comunidades puede ayudar a reducir los costos de entrenamiento para los individuos que trabajan en el dominio de la comunidad. Lesser y Storck encontraron que las comunidades de práctica eran de gran ayuda para los recién llegados para identificar a expertos en diferentes materias a los cuales pueden consultar y quienes los pueden guiar a recursos en la compañía (Lesser y Storck, 2001). Aún para los miembros estables de la comunidad, el hecho de compartir conocimiento es una actividad clave.

Al identificar a un grupo como una Comunidad de Práctica una organización hace visible un área crítica de habilidad. Sin esta conciencia, existen grandes probabilidades de que parte del talento corporativo pudiera desaparecer silenciosamente. Es posible que los miembros de una comunidad que abandonan la organización puedan estar dispersos a través de varias divisiones y por lo tanto no causar ninguna alerta. Pero esto, a largo plazo, puede resultar en una crisis corporativa porque todos los miembros son expertos en un mismo campo. Reconociendo a este grupo como una Comunidad de Práctica una organización se asegura de mantener y alimentar una base de habilidades clave.

Con la importancia de las comunidades de práctica establecida, las organizaciones pueden enfocar su atención en que su sistema de Gestión de Conocimiento permita a las mismas florecer. A medida que las organizaciones deciden dar soporte a estas comunidades, deberían saber que las comunidades de práctica no son de ninguna manera entidades uniformes. Existe una amplia variación entre las comunidades de prácticas y comprender estas diferencias

puede facilitar en gran medida que el soporte ofrecido realmente concuerde con las necesidades de cada una.

Dos observaciones son claves para entender las variedades de comunidades de práctica que son candidatas para el soporte organizacional: el número potencial de comunidades de prácticas es grande y hay muchos tipos de grupos y comunidades que no son comunidades de prácticas. Reconocer finas distinciones entre varias comunidades puede ayudar a una organización a entender si merecen soporte especializado, y si es así, qué tipo de soporte. La Tabla III.4 presenta los diferentes tipos de comunidades identificadas en esta tesis (Agresti, 2003):

| Tipo de Comunidad | Características |
|----------------------------|--|
| Comunidades de Práctica | Incluyen personas que realizan actividades de trabajo similares |
| Comunidades de Experticia | Los individuos poseen altos niveles de conocimiento en la misma área temática |
| Comunidades de Interés | Incluye aquellos que comparten un interés sobre un área temática |
| Comunidades de Aprendizaje | Estas personas se auto-organizan para aprender y crecer personal y profesionalmente |
| Equipos de Proyecto | Estos individuos se unen como grupo por un periodo de tiempo específico para hacer una tarea en particular y luego se separan |
| Fuerzas de Tareas | Este grupo tiene atributos similares a un grupo de proyecto, pero las personas trabajan de una forma totalmente dedicada a un único objetivo, por un período de tiempo muy corto y bajo una gran presión de agenda |
| Equipos de Alto Desempeño | Este grupo posee atributos asociados más cercanamente con un verdadero equipo que el grupo de trabajo o de proyecto típico, convirtiéndose a través del tiempo en una unidad altamente efectiva que se mantiene junta a través de un extenso período de tiempo realizando asignaciones de trabajo sucesivas y creciendo en madurez y efectividad |
| Unidad Organizacional | Estas personas comparten una membresía en una entidad definida como parte de una estructura organizacional |

Tabla III.4. Tipos de Comunidades

Un tipo similar de comunidad pero a su vez con características propias son las comunidades de conocimiento. Aunque tanto las comunidades de práctica

como las comunidades de conocimiento son comunidades formadas dentro de la organización, existen dos características principales que las distinguen. La primera es que la compañía toma la responsabilidad de identificar las posibles áreas de interés para establecer las comunidades de conocimiento. Contrariamente, las comunidades de práctica son orgánicas por naturaleza, esto es, se forman espontáneamente en respuesta a intereses profesionales que subyacen en la firma. La segunda característica es que mientras la organización activamente establece y da soporte a las comunidades de conocimiento, típicamente, no tienen un objetivo bien definido, más que expandir el conocimiento en áreas de interés común. La dirección y los objetivos de una Comunidad de Conocimiento tienden a ser menos definidos (Malone, 2002). Finalmente, a diferencia de estas últimas, las cuales son organizadas por la compañía con el propósito de filtrar conocimiento de valor potencial, una Comunidad de Práctica tiene por propósito principal el aprendizaje. En el Capítulo VI se propone hacer un uso efectivo de estas comunidades mediante su conexión en red.

5.3 Representar y Recuperar Conocimiento

Una gran cantidad de esfuerzo se ha puesto en la creación de los medios electrónicos necesarios para capturar y almacenar conocimiento y mejorar la comunicación. Sin embargo, esto no es suficiente para una implementación de una Gestión del Conocimiento efectiva. La experiencia muestra que pocos trabajadores contribuyen a los repositorios de conocimiento (base de casos, bases de conocimiento, etc.) o buscan conocimientos en ellos, y de esta forma, el conocimiento generado en el transcurso de sus actividades diarias se pierde (Vasconcelos y col., 2002).

Pueden mencionarse tres factores que contribuyen a la falta de uso de este tipo de repositorios (Kwan y Balasubramanian, 2003). Por un lado, la

contribución de conocimiento en los repositorios requiere de un esfuerzo de documentación extra para los trabajadores quienes, a menos que perciban un beneficio inmediato, no justifican el trabajo adicional. Asegurar que las personas que contribuyen en estos repositorios vean los beneficios de su contribución es crucial. Por otro lado, compartir conocimiento requiere de un marco mental común entre la fuente y el receptor, pero la gente con diferentes trasfondos tienen diferentes estructuras y perspectivas de conocimiento. Más aún, el diseño de repositorios se enfoca en los contenidos y tiende a proveer poco contexto del conocimiento que contiene. El conocimiento es, por definición, altamente dependiente del contexto mientras que, generalmente, una representación explícita tiende a causar una eliminación del contexto. Sin la información contextual, los trabajadores no entienden ni confían completamente en la fuente de conocimiento y por lo tanto no la adoptan (Ackerman, 1994). Además, el conocimiento tiene un ciclo de vida. En algún punto cesa de ser relevante y debemos ser capaces de clasificar el conocimiento y recuperar sólo aquel que es importante. Finalmente, en la mayoría de los casos, no existe dentro de la organización una cultura que fomente el intercambio de conocimiento.

Para enfrentar estos inconvenientes es necesario desarrollar sistemas de información facilitadores de conocimiento que provean un marco común para capturar, incrementar, almacenar, organizar, analizar y compartir no sólo información y datos sino también conocimiento (Nemati y col., 2002). Actualmente, las Memorias Organizacionales se proponen como soporte para un uso y preservación efectiva del conocimiento a través del tiempo y el espacio (tanto como sea posible) sin intervención humana. Desde la perspectiva organizacional, una Memoria Organizacional puede actuar como una herramienta para la Gestión del Conocimiento y facilita los tres tipos de aprendizaje organizacional: aprendizaje individual, aprendizaje a través de comunicación directa, y aprendizaje usando un repositorio de conocimiento (Van Heijst y col., 1996). En el Capítulo VI se analiza con detalle este tipo de repositorio como medio para la representación y recuperación del conocimiento.

6. Conclusiones

Existe una gran y creciente cantidad de literatura que enfatiza la importancia del conocimiento y el aprendizaje organizacional. En este capítulo se destacó la conveniencia de estructurar las iniciativas de Gestión del Conocimiento como ventaja competitiva a la hora de considerar al conocimiento dentro de la estrategia de la organización.

Se ha realizado un gran esfuerzo tendiente mayormente a la conversión del conocimiento tácito a explícito bajo la suposición de que en esta forma es más fácilmente compartido a través de toda la organización, desestimando, que no es posible ni conveniente explicitar todo el conocimiento que posee una persona.

El foco en las tecnologías de información provoca que se trate al conocimiento como un objeto que puede ser manipulado por un sistema de información. Sin embargo, las actividades relacionadas con el conocimiento tienen un alto componente social que no puede desconocerse si se quiere gestionar el conocimiento. Las organizaciones deben preocuparse por crear una cultura organizacional que facilite la creación, la distribución y uso del conocimiento.

En el campo académico, se han propuesto una variedad de modelos de Gestión del Conocimiento que pueden clasificarse en dos tipos principales: descriptivos y prescriptivos. Los modelos descriptivos tratan de caracterizar la naturaleza de la Gestión del Conocimiento, mientras que los modelos prescriptivos presentan metodologías a seguir en la conducción del proceso de Gestión del Conocimiento.

Se presentaron dos enfoques relacionados a las tendencias en la implementación de dichos modelos: el de repositorio y el de red. El enfoque de repositorio apunta a la codificación del conocimiento, es decir, la creación y mantenimiento de almacenes de conocimiento explícito. Por el contrario, el enfoque de red, apunta a usar el poder de las tecnologías de información y

comunicación para dar soporte al flujo de conocimiento en el ámbito organizacional y entre redes de dominios de conocimiento.

Otro aspecto considerado es que la acumulación de conocimiento organizacional puede alcanzarse mediante la adquisición de conocimiento desde fuentes externas y mediante creación interna. Esto implica la necesidad de realizar diferentes actividades para la Gestión de Conocimiento organizacional. Las principales son la creación, la distribución, la búsqueda y recuperación de conocimiento.

Finalmente, previo a la formulación concreta de la problemática que pretende resolver esta tesis, en el capítulo siguiente, se describen dos tecnologías utilizables en la implementación de sistemas de Gestión de Conocimiento.

Memorias Organizacionales y Ontologías

1. Introducción

Como se puntualizó en el capítulo anterior, una de las actividades claves en la Gestión del Conocimiento es la representación y recuperación del conocimiento. En este capítulo se discuten las principales herramientas utilizables para estas tareas. En la sección 2 se presentan los diferentes tipos de Memorias Organizacionales enfatizando su rol como medio para la recuperación del conocimiento. En la sección 3 se describen las ontologías de dominio como herramientas para la representación del conocimiento. Finalmente, en la sección 4, se presentan las conclusiones del capítulo.

2. Tipología y Facetas de las Memorias Organizacionales

Para Ackerman y col. es necesario considerar una Memoria Organizacional desde dos aspectos diferentes: como un objeto y como un proceso. Es decir, mantiene un estado y además está embebida en muchos procesos organizacionales e individuales (Ackerman y Halverson, 1998). Es importante enfatizar la diferencia entre el aspecto estático (el conocimiento capturado) y el aspecto dinámico de la memoria (la habilidad de memorizar y

recordar) ya que ambos aspectos deben estar presentes en una solución completa de Gestión del Conocimiento.

De hecho, pueden mencionarse tres aspectos que pueden utilizarse para definir lo que es una Memoria Organizacional: el contenido de la memoria (el *qué* o naturaleza del conocimiento), la forma de la memoria (el *dónde* o soporte de almacenamiento), y el trabajo de la memoria (el *cómo* o sistema de gestión del conocimiento). Aquí se presentan algunos elementos de definición encontrados en la literatura para cada una de estas facetas:

- ✓ **Contenido:** en este sentido puede decirse que una Memoria Organizacional contiene la experiencia organizacional adquirida por los empleados relacionada con el trabajo que realizan. Es un repositorio de conocimiento y conocimiento laboral (know-how) de un conjunto de individuos que trabajan en una organización en particular y que captura los bienes de conocimiento acumulados de la misma (Euzenat, 1996; Kuhn y Abecker, 1997). Puntualmente, consiste en la suma total de los recursos de información y conocimiento dentro de una organización. Tales recursos están típicamente distribuidos y están caracterizados por la multiplicidad y la diversidad: bases de datos de la empresa, textos legibles por computadoras, reportes y recursos de documentación, requerimientos de productos, análisis de diseño, etc. (Nagendra Prasad y Plaza, 1996). Constituye una representación explícita, corpórea y persistente del conocimiento corporativo y la información de una organización (sus procesos de negocios, sus procedimientos, sus políticas - misión, reglas y normas - y sus datos), preservando el razonamiento y el conocimiento junto con su diversidad y contradicciones para poder reutilizarlo más adelante (Van Heijst y col., 1996).

- ✓ **Forma:** Kuhn y col. caracterizan una Memoria Organizacional como un sistema de información más amplio o abarcativo (Kuhn y Abecker, 1997). En otros enfoques las Memorias Organizacionales toman la forma

de soluciones de biblioteca eficientes basados en sistemas de gestión de documentos. Por otro lado, una Memoria Organizacional puede basarse en políticas de gestión de recursos humanos viendo a las personas como “contenedores de conocimiento”.

- ✓ **Comportamiento:** la principal tarea de una Memoria Organizacional es hacer disponibles los bienes de conocimiento de una empresa mejorando la eficiencia y efectividad de los procesos intensivos en conocimiento. Además, es un sistema que permite la integración de conocimiento organizacional no estructurado y disperso mejorando su acceso, difusión y reutilización entre los miembros de una organización y los sistemas de información (Von Krogh, 1998). Finalmente, una de las funciones principales de una Memoria Organizacional es mejorar la capacidad de aprendizaje de una organización.

En esta tesis se define una Memoria Organizacional como:

Una representación persistente, corpórea y explícita del conocimiento y la información de una organización, para facilitar su acceso y reutilización por parte de los miembros de la organización en sus tareas.

Sin embargo, no siempre es posible representar todo el conocimiento en forma corpórea y explícita. A veces es mejor referenciar o indexar un recurso externo más que duplicarlo para poder consultarlo cuando sea necesario, pero no tener que memorizarlo y mantenerlo. Las razones tienen que ver con cuestiones de costo, de factibilidad (copyright, cantidad de datos, dificultad de formalización, etc.), de cambio y volatilidad, falta de la experticia necesaria para el mantenimiento, etc. Por lo tanto, una solución integral de Memoria Organizacional (memoria y sistema de gestión) debería permitir el indexado de fuentes externas, incorporadas en una persona o en otra memoria. Teniendo en cuenta esto último, una definición extendida podría ser:

Una Memoria Organizacional es una representación e indexado persistente, corpóreo y explícito del conocimiento e información o sus fuentes en una organización, para facilitar el acceso, intercambio y reutilización por parte de los miembros de la organización, para sus tareas individuales y colectivas.

Cabe enfatizar en este punto que el deseo de controlar y gestionar el conocimiento contrasta con su naturaleza fluida, dispersa, intangible, subjetiva y algunas veces tácita. Se describirán a continuación las diferentes tipologías y facetas de las Memorias Organizacionales.

Frecuentemente, al no encontrar un modelo de Gestión del Conocimiento completo o que se adecue a una situación en particular, las organizaciones desarrollan soluciones propias que dan origen a una multiplicidad de tipos de Memorias Organizacionales y sistemas de gestión.

Desde un punto de vista externo, Barthés presenta varias facetas de las Memorias Organizacionales que pueden ser consideradas: socio-organizacional, económica, financiera, técnica, humana y legal (Barthes, 1996). Desde un punto de vista interno, Van Heijst y col. presentan una clasificación basada en los procesos de gestión que se pretenden implementar tal como se observa en la Tabla IV.1 (Van Heijst y col., 1996):

| | Colección Pasiva | Colección Activa |
|---------------------|--|--|
| Distribución Pasiva | <p>Desván de conocimiento: es una Memoria Organizacional usada como un archivo que puede ser consultado y actualizado cuando sea necesario. No es intrusivo, pero requiere de una gran disciplina por parte de los miembros de la organización para no quedar obsoleto.</p> | <p>Esponja de conocimiento: es una Memoria Organizacional que se alimenta activamente para mantenerla más o menos completa. Su uso queda como responsabilidad de los miembros de la organización.</p> |
| Distribución Activa | <p>Editor de conocimiento: es una Memoria Organizacional donde la contribución se deja a los trabajadores individuales mientras que los mantenedores de la memoria analizan el conocimiento entrante y lo combinan con el conocimiento ya</p> | <p>Bomba/drenaje de conocimiento: es una Memoria Organizacional que asegura que el conocimiento desarrollado en la organización es efectivamente capturado y usado desde y por los miembros de la organización.</p> |

| | | |
|--|--|--|
| | almacenado y reenvían las noticias relevantes a miembros potencialmente interesados. | |
|--|--|--|

Tabla IV.1. Clasificación de las Memorias Organizacionales en función de los Procesos de Gestión que Implementan

Otros investigadores clasifican las memorias dependiendo del conocimiento contenido en ellas (Dieng y col., 1999):

- ✓ **Memoria técnica/profesional/de negocios:** compuesta por los métodos, herramientas y documentos de referencia usados en una profesión dada. Este es conocimiento sobre un dominio, los resultados de búsquedas e investigaciones, es decir, el conocimiento usado diariamente dentro de la organización por sus miembros para realizar sus actividades cotidianas.
- ✓ **Memoria de gestión:** relacionada a la organización, sus actividades, productos y actividades; sobre la organización en sí misma, su estructura, su configuración, sus principios de gestión, sus políticas, su historia. Captura las estructuras organizacionales pasadas y presentes (recursos humanos, gestión, etc.). Esta memoria está extremadamente cercana al modelado organizacional.
- ✓ **Memoria individual:** caracteriza el estado, competencias, conocimiento laboral (know-how) y actividades de un miembro dado de la organización
- ✓ **Memoria de proyecto:** se adquiere en el contexto de un proyecto que debe ser almacenado junto con el conocimiento para preservar su significado. Comprende la definición del proyecto, las actividades,

historia y resultados. Se usa para capitalizar lecciones y experiencias de proyectos anteriores. La memoria de un proyecto preserva la memoria técnica y la memoria de gestión movilizadas para el proyecto. Puede ser la memoria de un proyecto en curso o de un proyecto pasado. En cualquier caso, es importante capturar el curso/progreso, la base lógica y el contexto.

Finalmente, las Memorias Organizacionales pueden clasificarse de acuerdo a algunas características adicionales:

- ✓ ***Memoria no computacional:*** está compuesta por artefactos físicos (documentos en papel, cintas de video, etc.) que capturan conocimiento que no ha sido elicitado previamente. Dieng y col. distinguen dos objetivos para la construcción de tal tipo de memoria: para elaborar documentos que sinteticen conocimiento que no ha sido explicitado en reportes o documentación técnica, y está más relacionado con el conocimiento laboral (know-how) de los expertos organizacionales; o para mejorar la producción de la empresa a través de la proposiciones de expertos sobre sus tareas en el proceso de diseño (Dieng y col., 1999).

- ✓ ***Data warehouses y Data marts:*** en muchas organizaciones la primera herramienta de Gestión del Conocimiento es un data warehouse, es decir, un área de almacenamiento central para los datos de transacción de una organización (O'Leary, 1998). Generalmente replica o por lo menos accede al contenido de varias bases de datos de la organización. A partir de ellas, ofrece una amplia variedad de datos y trata de presentar una imagen coherente de la condición de los negocios en un momento dado. Los sistemas de generación de reportes automáticos basados en consultas y vistas de las bases de datos como así también técnicas de recuperación de conocimiento y minería de datos más complejas pueden usarse para

permitir a los trabajadores del conocimiento recolectar información que dé soporte a la toma de decisiones. Los data marts son como el data warehouse pero usualmente más pequeños, se enfocan en un tema o departamento en particular y pueden ser subconjuntos de data warehouses más grandes. En general están relacionados con una comunidad de trabajadores de conocimiento.

- ✓ ***Memorias internas vs. memorias externas:*** una Memoria Organizacional no necesariamente se restringe a una única organización (Rabarijoana y col., 1999). Una Memoria Organizacional interna puede basarse en recursos internos mientras que una Memoria Organizacional externa más bien incluye información y conocimiento obtenido del mundo exterior, pero útil para las actividades de la organización. La recuperación e integración de la información disponible en la Web es una característica interesante y podría pensarse en la Memoria Organizacional como un filtro para vencer la heterogeneidad y sobrecarga de información de la Web.

- ✓ ***Memoria basada en documentos o warehouses de conocimiento:*** se basa en documentos existentes para construir la memoria. La construcción de tal tipo de memoria comienza con la recolección de diferentes documentos y requiere de una interfaz para gestionarlos (agregado de documentos, recuperación de documentos, etc.) (Dieng y col., 1999). Un buen sistema de documentación es probablemente la solución de Gestión del Conocimiento más barata y factible. Este tipo de Memoria Organizacional está más cerca de los warehouses de conocimiento que propone O'Leary que apuntan a datos cualitativos y contienen manuales y reglas de diseño, especificaciones y requerimientos (O'Leary, 2002). Rabarijoana y col. identifica varios tipos de documentos

que pueden ser explotados en una Memoria Organizacional basada en documentos (Rabarijoana y col., 1999):

- i.* documentos relacionados con proyectos: especificaciones del producto a ser diseñado o manufacturado, documentos de diseño, documentos de pruebas, reportes técnicos contractuales,
 - ii.* libros de referencia en una profesión dada,
 - iii.* documentos visuales como fotos, planos escaneados, documentos iconográficos,
 - iv.* reportes técnicos, artículos técnicos y científicos,
 - v.* libros, tesis, normas, documentos archivados, guías, expedientes de inteligencia tecnológica,
 - vi.* documentación en línea, manuales del usuario, expedientes de negocios, etc.
- ✓ ***Memoria Organizacional basada en conocimiento:*** esta memoria está basada en la elicitación y modelado explícito del conocimiento de los expertos. Puede ser mezclada con la memoria basada en documentos previo indexado de los documentos a través de una representación formal de conocimiento subyacente. Sin embargo, el objetivo de este enfoque es proveer asistencia a los usuarios, proveyéndolos con conocimiento corporativo relevante, pero dejándole la responsabilidad de la interpretación contextual y la evaluación de esta información (Kuhn y Abecker, 1997).
- ✓ ***Memorias basadas en casos:*** las organizaciones tienen una colección de experiencias pasadas (éxitos o fracasos) que pueden ser explícitamente representadas en el mismo formalismo para poder compararlas; estas

experiencias formalizadas son llamadas casos y su gestión puede explotar el razonamiento basado en casos (Simon, 1996; Simon y Grandbastien, 1995). Dieng y col. distinguen dos objetivos: evitar la dispersión de la experticia concentrando el conocimiento de todos los expertos en casos dedicados; y permitir la continua evolución de la memoria gracias a la adición progresiva de nuevos casos (Dieng y col., 1999). El razonamiento basado en casos permite capitalizar sobre casos ya encontrados para resolver casos nuevos. El mecanismo de recuperación se construye sobre una medida de similitud para encontrar casos pasados lo suficientemente cercanos como para sugerir una solución. Este enfoque es muy útil para memorias de proyecto.

- ✓ ***Memoria basada en las personas:*** los individuos representan la locación primaria donde se localizan los recursos intelectuales de una organización (Dzbor y col., 2000). Por lo tanto, un tipo de aplicación corporativa que generalmente se usa es una base de conocimiento de recursos humanos basada en las capacidades y habilidades de los empleados, su educación, especialidad, experiencia previa, etc. Aunque es importante formalizar y materializar el conocimiento para hacerlo perenne y perpetuarlo, también es claro que, hasta el momento, no todo el conocimiento puede ser capturado en un sistema simbólico para ser memorizado. En ese caso es importante capturar la identidad de las fuentes (por ejemplo, un experto) e indexar este sistema de recursos externos para incluirlo en la memoria general y también para saber qué parte de la memoria actual no está formalizada y almacenada en forma segura. Druker afirma que como el conocimiento es el concepto clave de la sociedad futura, las personas serían el eje central ya que, aunque parcialmente almacenado en base de datos y libros, el conocimiento más valioso está incorporado en las personas (Drucker, 1988). Liao y col. proponen un sistema de Memoria Organizacional basado en conocimiento de competencias que facilita la localización de la persona de contacto apropiada (Liao y col., 1999). Otro

tipo de sistema de localización de expertos es el propuesto por Sangüesa Sol y col. que construye una representación de redes de “conocidos” minando las referencias de las páginas web (Sangüesa Sol y Pujol Serra, 2001).

Aunque tradicionalmente las empresas han acometido la Gestión del Conocimiento ya sea desde un punto de vista de gestión o desde un punto de vista puramente tecnológico, el establecimiento de una Memoria Organizacional requiere de un enfoque balanceado e integrado. Más aún, de acuerdo a Ackerman y col. no existe una única Memoria Organizacional, sino más bien una supramemoria individual que engloba varias personas y artefactos (Ackerman y Halverson, 1998). Esta visión de una red de artefactos y personas, de memoria y de procesamiento limitada por convenciones sociales, provee una comprensión más profunda y a final de cuentas más útil de la vida organizacional. De lo anterior se desprende que la gestión efectiva de conocimiento requiere de soluciones híbridas que involucren tanto personas como tecnología (Davenport, 1996). El rol de las personas en esta sociedad es la de resolver problemas y derivar soluciones innovadoras y creativas usando el conocimiento preservado y provisto por la memoria. Esta idea reemplaza a los sistemas expertos tradicionales que actuaban en forma más autónoma, pero que carecían de muchas de las capacidades esenciales de las personas. Las Memorias Organizacionales son ejemplos típicos de herramientas basadas en computadoras de asistencia inteligente (Dzbor y col., 2000).

A pesar de la heterogeneidad inherente al conocimiento, en el curso de sus actividades, las personas básicamente reutilizan documentos existentes y crean nuevos, o los utilizan para describir otros tipos de objetos de conocimiento (imágenes, audio, video, perfiles profesionales, etc.). A medida que los individuos emplean estos documentos y producen nuevos, una considerable cantidad de conocimiento se construye alrededor de estos documentos. Este conocimiento puede ser tanto de cómo están relacionados unos con otros, como también una comprensión más profunda sobre su contenido. Este conocimiento

típicamente reside en la mente de las personas y nunca es almacenado. Otra característica interesante relacionada es que a medida que pasa el tiempo las personas olvidan este conocimiento. Para poder colaborar efectivamente y evitar la duplicación de esfuerzos necesitan comunicar y mantener inventarios abarcativos para asegurar que el mismo trabajo no se haya hecho en otro lado (Slade y Bokma, 2001).

La gestión de documentos no sólo es cuestión de una correcta clasificación. A pesar de que las técnicas de recuperación y gestión de la información basada en búsquedas tienen sus usos, generalmente depende de la habilidad del usuario para localizar y recuperar correctamente los documentos. Las herramientas de búsqueda en web son un ejemplo típico de este problema y generalmente proveen malos y dispares resultados dejando que los usuarios naveguen el sitio manualmente para obtener la información buscada.

Los sistemas de gestión de documentos tienen un número de beneficios para mantener una gran colección de documentos donde éstos son compartidos por una fuerza de trabajo y ayudan a cubrir una variedad de necesidades tales como archivo, rastreo de documentos, protección, etc. para que los documentos sean protegidos y accedidos de una forma organizada y que no exista una proliferación de copias del mismo documento que puede llevar a la confusión. Al mismo tiempo estos sistemas no dan soporte adecuado a las necesidades de la Gestión del Conocimiento de proveer caminos alternativos ni proveen formas intuitivas para registrar las interrelaciones de los documentos y el conocimiento asociado. Para gestionar los documentos en una forma más intuitiva, existe la necesidad de ir más allá de considerar los documentos en forma aislada y representar el contexto donde los documentos encajan. Esto permitiría tanto el acceso directo a cualquier elemento asociado como también hacer explícitas cualquier relación entre ellos. Este enfoque también sería capaz de comunicar el contexto en el cual encajan los documentos y registros de conocimiento lo cual es vital para la correcta interpretación en un entorno compartido.

Esta tendencia es cada vez más evidente en la actualidad donde se propone que las estrategias de Gestión del Conocimiento deberían seguir

principios conceptuales y permitir a los usuarios catalogar y acceder a los objetos de conocimiento en una forma más natural e intuitiva. Esto requiere un modelo de la organización y del dominio respectivo y la asociación de los documentos con elementos de este modelo. De esta forma el contexto en el cual encajan los documentos y cómo se relacionan con otros documentos puede explicitarse.

Las ontologías han surgido como la herramienta ideal para desarrollar un modelo detallado de un dominio seleccionado. Al mismo tiempo, los enfoques basados en ontologías son en principio apropiados a los efectos del desarrollo de clasificaciones específicas de la organización o específicos del dominio y pueden usarse para categorizar y asociar ítems de información y conocimiento.

3. Ontologías – Modelado de Conocimiento

Con el creciente acceso a repositorios de conocimiento independiente y heterogéneo, el tratamiento de las diferencias en la estructura y semántica de los objetos de conocimiento juega un rol principal en los sistemas de información.

Desde los primeros estudios en sistemas de información interoperacionales se han hecho progresos relacionados a la heterogeneidad sintáctica (tipos de datos y formatos) y estructural (integración de esquemas, lenguajes de consulta e interfaces) (Aparicio y col., 2005). A medida que los sistemas de información interoperacionales fueron confrontados con tareas de Gestión de Conocimiento cada vez más complejas, la tecnología necesaria para tratar con éxito estos temas debió enfocarse en la semántica subyacente a los datos utilizados por estos sistemas.

Investigaciones recientes en recuperación de información e integración de datos han enfatizado el uso de ontologías y funciones de similitud semántica como mecanismo de comparación de objetos que pueden ser recuperados o integrados a través de repositorios heterogéneos (Guarino y col., 1999) (Rodríguez y Egenhofer, 2003). En lo que resta de este capítulo se analizan

diferentes definiciones de ontología y su uso como herramienta de modelado del conocimiento heterogéneo dentro de una Memoria Organizacional.

En este contexto, una ontología es un tipo de base de conocimiento que describe conceptos a través de definiciones que son lo suficientemente detalladas para capturar la semántica de un dominio. Una ontología captura una cierta visión del mundo, da soporte para las consultas intencionales relacionadas con el contenido de la base de datos y refleja la relevancia de los datos proveyendo una descripción declarativa de la información semántica independiente de la representación de los datos (Rodríguez y Egenhofer, 2003).

Existe una gran variación entre el nivel de detalle y lógica de diferentes representaciones ontológicas. Por ejemplo, una ontología terminológica es una colección de categorías organizadas por un orden parcial que es inducido por inclusión. Una ontología diferente y más detallada es una ontología axiomatizada, esto es, una ontología terminológica cuyas categorías están distinguidas por axiomas y definiciones formuladas en lógica o en algún lenguaje que pueda ser traducido en lógica (Bartsch, 2004).

Recientemente, se ha asistido a una explosión de interés en las ontologías como artefactos para representar el conocimiento humano y como un componente crítico de la gestión del conocimiento, la Red Semántica, las aplicaciones B2B (Business to Business) y muchas otras áreas de aplicación. Varias comunidades de investigadores asumen que las ontologías son la estructura de modelado apropiada para la representación de conocimiento. Sin embargo, poca discusión se ha realizado en relación al verdadero rango de conocimiento que una ontología puede representar con éxito (Staab, 2004).

Es obvio que no puede representarse fácilmente ciertos tipos de conocimiento (por ejemplo, habilidades de un empleado) o que no pueden transformarse ciertos tipos de representaciones en formatos ontológicos apropiados (por ejemplo, conocimiento contenido en diagramas y esquemas). Por otro lado, hay ciertos tipos de conocimiento que son extremadamente adecuados para una representación ontológica, como por ejemplo la información

taxonómica. La mayoría de las definiciones de ontología, aunque no todas, insisten en que una ontología específicamente representa estructuras conceptuales compartidas y comunes. El interrogante aquí es si este proclamado consenso es garantía de adecuación.

Otro tema de análisis es hasta qué punto las ontologías son diferentes de las redes, los gráficos, los tesauros, diccionarios y taxonomías. Específicamente, sobre el solapamiento de la definición de ontología (interpretada como una clasificación jerárquica de entidades o cosas) y la de tesoro o taxonomía (clasificaciones jerárquicas de palabras o sentidos léxicos) existe una creencia generalizada de que son construcciones diferentes, tan diferentes como, en otra dimensión, las enciclopedias y los diccionarios. Existen también casos particulares en los cuales estas definiciones parecen mezclarse. Por ejemplo, WordNet¹ es denominado una ontología, lo cual parece ser en ocasiones, pero esto no es importante en relación a su función como principal recurso para el procesamiento de lenguaje natural como tampoco es importante que un diccionario contenga hechos del mundo real del tipo “un crisantemo es una flor procedente de China”.

Ontología es, generalmente, un término genérico y rara vez detalladamente definido. Algunas de las estructuras denominadas ontologías son jerarquías estrictas de nombres de categorías y nada más. Otras son taxonomías o gráficos donde las categorías pueden repetirse y se permiten bucles en la estructura de datos. Otras dan soporte a vocabularios corporativos de términos preferidos y sus sinónimos y pueden ser multilingües. Otras requieren de una estructura de metadatos compleja que puede usarse para priorizar la recuperación, para indicar políticas de retención o conformidad con las leyes, o para soporte técnico de aplicaciones relacionadas. Todo esto parece indicar que hay una unidad o foco en base al propósito más que del tipo estructural.

Las ontologías corporativas, cualquiera sea el nivel de sofisticación, tienen como objetivo dar soporte a la sistematización de grandes volúmenes de información usando abstracciones. Esto es diametralmente opuesto a las

¹ Disponible en <http://wordnet.princeton.edu/index.shtml>

ontologías de Inteligencia Artificial de mediados de los 70 y principio de los 80, donde el objetivo era generalmente representar un dominio pequeño con un alto grado de detalle.

Las ontologías corporativas enfrentan varios problemas que generalmente no son considerados en la investigación académica (Brewster y col., 2004). Estos incluyen seguridad y pertenencia, confianza, la audiencia deseada, y el medio usado para ver la ontología. Parecería obvio que las ontologías usadas en las empresas reflejen el tamaño de esas empresas. Como tales, es probable que sean grandes y particionadas de acuerdo a los intereses de la empresa. Si la ontología es interna, probablemente reflejará las divisiones dentro de la organización, tales como Marketing, I&D, Recursos Humanos, etc. Si la ontología es para uso externo, probablemente representará categorías de productos, divisiones de ventas y soporte técnico, etc. En ambos casos, una parte separada de la organización será *dueña* de cada sección de la ontología y será responsable por su desarrollo y mantenimiento. Por lo tanto, se requieren mecanismos de seguridad de manera tal que sólo aquellos autorizados puedan editar o incluso ver las secciones de la ontología de las cuales son responsables.

No obstante lo anterior, después de más de una década de discusión, la comunidad científica todavía no ha alcanzado un consenso completo en relación a lo que precisamente es una ontología. La definición de Gruber, que establece que una ontología es una especificación explícita de una conceptualización, todavía es verdadera pero tiene algunos puntos discutibles (Gruber, 1993a). El problema principal es que no existe un consenso en lo que es una especificación. Ciertamente, una jerarquía taxonómica de conceptos es una especificación apropiada, aunque las taxonomías por sí mismas son más bien pobres en lo que pueden representar sin mencionar la incapacidad de representar atributos de los conceptos.

Existen varias definiciones de la palabra *ontología* las cuales han evolucionado a través del tiempo. Una de las primeras definiciones es la de Neches quien establece que “*una ontología define los términos y relaciones básicos que comprenden el vocabulario de un área como también las reglas*

para combinar los términos y las relaciones para definir extensiones de ese vocabulario” (Neches y col., 1991). Esta definición descriptiva dice qué hacer para construir una ontología y da algunas guías vagas: la definición identifica términos básicos y relaciones entre términos, identifica reglas para combinar los términos, y provee definiciones de tales términos y relaciones. Hay que destacar, de acuerdo a la definición de Neches, que una ontología incluye no sólo los términos explícitamente definidos en ella sino también el conocimiento que puede ser inferido de ella.

Unos años más tarde, Gruber definió una ontología como *“una especificación explícita de una conceptualización”* (Gruber, 1993a). Esta definición se convirtió en la más citada en la literatura y la comunidad de ontologistas. Borst modificó ligeramente la definición de Gruber diciendo que *“las ontologías son especificaciones formales de conceptualizaciones compartidas”* (Borst, 1997). Estas dos últimas definiciones han sido mezcladas y explicadas por Studer y col. (Studer y col., 1998). Conceptualización se refiere a un modelo abstracto de algún fenómeno en el mundo habiendo identificado los conceptos relevantes de ese fenómeno. Explícito significa que el tipo de conceptos usados y las restricciones en su uso están definidos explícitamente. Formal se refiere al hecho de que la ontología debería ser entendible por una máquina. Compartida refleja la noción de que una ontología captura conocimiento consensual, esto es, no es privativo de un individuo sino aceptado por el grupo.

En 1995, Guarino y col. recopilaron y analizaron siete definiciones de ontología y proporcionaron sus propias definiciones sintácticas y semánticas (Guarino y col., 1995). En ese trabajo en particular, estos autores propusieron considerar que una ontología es *“una teoría lógica que da una explicación explícita y parcial de una conceptualización”*, donde una conceptualización es básicamente una idea del mundo que una persona o grupo de personas puede tener. Aunque en la superficie la noción de conceptualización es similar a la de Studer y col., puede decirse que Guarino y col. fueron un paso más adelante debido a que establecieron cómo construir una ontología mediante la

construcción de una teoría lógica (Studer y col., 1998) (Guarino y col., 1995). Por lo tanto, estrictamente hablando, esta definición sería sólo aplicable a ontologías desarrolladas en lógica.

Existe otro grupo de definiciones basadas en el proceso seguido para construir la ontología. Estas definiciones también incluyen algunos puntos de interés sobre la relación entre las ontologías y las bases de conocimiento. Por ejemplo, la definición dada por Bernaras y col. dentro del marco del proyecto KACTUS (Schreiber y col., 1995): *“una ontología provee los medios para describir explícitamente la conceptualización detrás del conocimiento representado en una base de conocimiento”* (Bernaras y col., 1996). Esta definición propone “extraer” la ontología a partir de una base de conocimiento, lo cual refleja el enfoque que tienen los autores para construir ontologías. En este enfoque, la ontología se construye, siguiendo una estrategia de abajo hacia arriba, basándose en una base de conocimiento de una aplicación mediante un proceso de abstracción. A medida que se construyen más aplicaciones, la ontología se vuelve más general, y, por lo tanto, se alejaría de lo que sería una base de conocimiento.

Otra estrategia para construir ontologías es reutilizar ontologías más generales para construir ontologías específicas de dominio y bases de conocimiento, en este sentido: *“una ontología es un conjunto estructurado jerárquicamente de términos para describir un dominio que puede ser usado como base estructural para una base de conocimiento”*. De acuerdo a esta definición, la misma ontología puede ser usada para construir varias bases de conocimiento, las cuales compartirían el mismo esqueleto. Las extensiones del esqueleto deberían ser posibles a bajo nivel agregando subconceptos específicos de dominio, o a un alto nivel agregando conceptos de nivel intermedio y alto que cubren nuevas áreas (Swartout y col., 1997).

A veces, la noción de ontología se diluye, en el sentido de que las taxonomías se consideran ontologías completas. Por ejemplo, UNSPSC², e-

² <http://www.unspsc.org>

cl@ss³, y RosettaNet⁴, que son estándares en el dominio de comercio electrónico, y el Directorio de Yahoo que es una taxonomía para buscar en la web, a veces se consideran ontologías porque proporcionan una conceptualización consensual de un dominio dado. Otra distinción posible es entre las ontologías que son principalmente taxonomías de las ontologías que modelan el dominio de una forma más profunda y proveen más restricciones sobre la semántica de dominio. Las mismas se denominan ontologías livianas (lightweight) y pesadas (heavyweight) respectivamente. Por un lado, las ontologías livianas incluyen conceptos, taxonomías de conceptos, relaciones entre conceptos y propiedades que describen los conceptos. Por otro lado, las ontologías pesadas agregan axiomas y restricciones a las ontologías livianas (Corcho y col., 2003).

Debido a que las ontologías son ampliamente usadas para diferentes propósitos (procesamiento de lenguaje natural, comercio electrónico, integración de información inteligente, la red semántica, etc.) en diferentes comunidades (ingeniería del conocimiento, ingeniería de software y base de datos, etc.), Uschold y Jasper proveen una nueva definición de la palabra ontología para popularizarla en otras disciplinas (Uschold y Jasper, 1999). Hay que destacar que la comunidad de base de datos, como la comunidad orientada a objetos también construye modelos de dominio usando conceptos, relaciones, propiedades, etc., pero la mayoría de las veces impone menos restricciones semánticas que aquellas impuestas en las ontologías pesadas. Uschold y Jasper definen una ontología como sigue: *“una ontología puede tomar una variedad de formas, pero incluirá necesariamente un vocabulario de términos y algunas especificaciones de su significado. Esto incluye definiciones y una indicación de cómo los conceptos están interrelacionados lo cual colectivamente impone una estructura en el dominio y restringe las posibles interpretaciones de los términos”*.

Independientemente de lo anterior, puede decirse que existe un consenso a niveles básicos dentro de la comunidad de ontologistas y por lo tanto no existe

³ <http://www.eclass.org>

⁴ <http://www.rosettanel.org>

confusión con respecto al uso de una ontología. Las diferentes definiciones proveen puntos de vista complementarios y diferentes de la misma realidad. Algunos autores proveen definiciones que son independientes de los procesos seguidos para construir una ontología, mientras que otras definiciones son influenciadas por el proceso de construcción. Como conclusión general puede decirse que el objetivo de las ontologías es capturar conocimiento consensual de una forma genérica y formal, y pueden ser reutilizadas y compartidas a través de diferentes aplicaciones de software y por grupos de personas.

Interrogantes relacionados a cuánto conocimiento puede representar una ontología y cuán adecuada es una ontología para la representación de conocimiento son en su mayoría incontestables ya que los lenguajes de especificación de ontologías varían grandemente en su poder de expresión. Es mucho más útil concentrarse en qué es una conceptualización en primer lugar más que profundizar en la expresividad de vehículos particulares para la codificación de ontologías.

Aunque no existe un simple predicado que nos diga sin ambigüedades si una especificación particular es una ontología hay ciertas cosas en las que existe un acuerdo. Puede acordarse que las ontologías enumeran conceptos representativos en un área de aplicación y que típicamente definen propiedades de conceptos y relaciones entre conceptos y generalmente indican restricciones sobre esas propiedades y relaciones. Una ontología provee un dominio de discurso para discutir sobre un área de aplicación, pero no representa (ni puede hacerlo) todo el conocimiento de un agente.

En este sentido, y en relación a las limitaciones de las ontologías como medio de representación de conocimiento, Alan Newell caracteriza el conocimiento como un fenómeno de comportamiento, viendo el conocimiento en términos de los objetivos de los agentes, las acciones de las cuales el agente puede ser capaz, y los medios por los cuales el agente selecciona las acciones para alcanzar sus objetivos (Newell, 1982). Una observación clave es que el conocimiento le permite a un agente implementar procedimientos para alcanzar

sus objetivos, y que atribuimos conocimiento a un agente porque lo observamos comportándose en el mundo de una manera aparentemente racional.

Esta noción de conocimiento va mucho más allá de la especificación de una conceptualización, de una enumeración de conceptos y relaciones. Desde la perspectiva de Newell, el conocimiento es más que la simple descripción de lo que existe en el mundo, directamente enlaza objetivos con acciones. En este sentido, el conocimiento tiene un alto contenido procedural que no puede ser directamente instanciado con una ontología.

Muchos sistemas inteligentes están diseñados primariamente para contestar preguntas sobre grandes cuerpos de conocimiento. En estos casos, las ontologías proveen la base necesaria para la representación en la construcción de sistemas completos. Para construir sistemas que resuelvan tareas del mundo real, sin embargo, no es suficiente con la especificación de nuestras conceptualizaciones, es necesario enlazar las ontologías a las fuentes de conocimiento organizacionales (Staab, 2004).

Desde la perspectiva de la ingeniería de conocimiento, el modelado de conocimiento, consiste en representarlo para poder almacenarlo, comunicarlo y manipularlo externamente. Automatizar su manipulación externa conduce al diseño de sistemas basados en conocimiento, esto es, sistemas en los cuales el comportamiento se basa en la manipulación simbólica de modelos formales de piezas de conocimiento para realizar operaciones significativas que simulen capacidades inteligentes.

El paso de representación trae aparejado el problema de la forma, es decir, la elección de un formalismo de representación que permite capturar la semántica en juego en las piezas de conocimiento consideradas. Un enfoque analizado por la ingeniería de conocimiento que ha emergido en la última década está basado precisamente en el modelado ontológico. En este ámbito, como una primera aproximación, puede decirse que una ontología es la parte del modelo de conocimiento que captura la semántica de las primitivas usadas para realizar

afirmaciones formales sobre el dominio de aplicación de la solución basada en conocimiento.

La utilidad de las ontologías en el área de Gestión del Conocimiento es evidente. Las personas, organizaciones y sistemas de software deben comunicarse entre ellos. Sin embargo, debido a las diferentes necesidades y contextos de fondo, puede haber puntos de vista y presunciones ampliamente variados relacionados esencialmente con un mismo tema. Cada uno utiliza diferentes jergas, cada uno puede tener conceptos, estructuras y métodos diferentes, superpuestos y/o mal emparejados (Uschold y Gruninger, 1996).

En el ámbito organizacional, algunas de las consecuencias de la falta de una comprensión compartida son: una pobre comunicación, dificultades en identificar requerimientos y por lo tanto, en especificar un sistema, interoperabilidad limitada, potencial de reutilización y distribución limitada y por lo tanto, esfuerzos gastados en lo que comúnmente se conoce como la *reinención de la rueda*. Existe una obvia necesidad de reducir o eliminar la confusión conceptual o terminológica y llegar a un entendimiento compartido. El desarrollo e implementación de un relato explícito de una comprensión compartida (esto es una *ontología*) en un área dada, puede mejorar la comunicación, lo cual a su vez puede traer aparejado una mayor reutilización, distribución, e interoperabilidad del conocimiento disponible. Desde esta óptica, una ontología es un marco unificador para diferentes puntos de vista y sirve como la base para la comunicación entre las personas, entre personas y sistemas, y entre sistemas: este marco unificador conceptual tiene como propósito funcionar como una lengua franca.

Las ontologías son consideradas como poderosas herramientas para solucionar la ambigüedad inherente a la heterogeneidad del conocimiento organizacional ya que proveen una base semántica y un vocabulario conceptual consensuado, sobre el cual uno puede construir descripciones y actos de comunicación. A medida que se desarrollan sistemas más inteligentes, aparece como más importante el rol esencial que juega el conocimiento general sobre las cosas y sus categorías en la formulación de inferencias. Por lo tanto, es necesario

proveer este conocimiento a las máquinas para que se comporten de forma inteligente e inteligible.

En este punto es importante notar que nada en la definición original de una ontología usada en la ingeniería de conocimiento (esto es, una ontología es una especificación de una conceptualización (Gruber, 1993b)) obliga a los ontologistas a usar un lenguaje formal para hacer explícita una ontología. La representación final de las intenciones y la estructura ontológica puede hacer uso de lenguajes más o menos formales, dependiendo del uso que se le intenta dar a la ontología. Un aprovechamiento automatizado de una ontología por parte de un sistema artificial implicará probablemente una formalización de algunos aspectos particulares de la ontología para permitir una manipulación automática de estos aspectos. La expresión formal de una intención provee una representación no ambigua y precisa del significado de un concepto, es decir, permite a un software manipularlo y usarlo. Desde el punto de vista de la formalización, Sowa distingue entre una ontología terminológica y una ontología formal. Son los dos extremos de un mismo continuo: mientras más axiomas se agregan a una ontología terminológica la misma puede evolucionar a una ontología formal o axiomatizada (Sowa, 1984).

En Gestión del Conocimiento, las ontologías proveen una buena metáfora para la estructuración de las bases de conocimiento. Actualmente el enfoque más prometedor para la gestión de grandes corpus de documentos descansa en el empleo de tecnologías que involucran ontologías y redes semánticas para la estructuración y gestión del conocimiento documentado (Ahlers y Weimer, 2002). Naturalmente, cuando se construye una ontología, se está restringido a modelar una pequeña ventana del mundo real, vista desde una perspectiva. Por ejemplo, dependiendo del contexto, podemos entender los conceptos de “*hombre y mujer*” como “*esposo y esposa*”, “*padre y madre*”, “*empleado y empleador*”, “*cliente y vendedor*” y así sucesivamente. No todos estos significados pueden anticiparse, y mucho menos modelarse, cuando se construye una ontología.

Puntualmente, en esta tesis las ontologías se consideran como la herramienta ideal de modelado de las fuentes de conocimiento heterogéneas debido a que:

- ✓ Proveen estructuras de navegación y vistas para la visualización manual de los conceptos incluidos en el modelo (McGuiness, 1998; O'Leary, 1998)
- ✓ Facilitan el acceso en lenguaje natural (Guarino y col., 1999)
- ✓ Proveen conocimiento de apoyo para la expansión o reescritura de una consulta (Bodner y Song, 1996; Sintek y col., 2000)
- ✓ Permiten la gestión de medios no textuales (Khan y McLeod, 2000) y,
- ✓ Dan soporte para la integración y recuperación de información desde fuentes distintas y distribuidas (Heflin y Hendler, 2000; Staab y col., 2000).

Virtualmente en todos estos escenarios las ontologías son la base para la articulación de las demandas de información por parte de los consumidores de información, o para la caracterización de las ofertas de información por parte de los proveedores de conocimiento. Cuando se pretende facilitar el acceso en lenguaje natural, aparece la necesidad de tratamiento del lenguaje natural, lo cual trae aparejada la necesidad de enfrentar problemas asociados con fenómenos del lenguaje como por ejemplo: la polisemia y la sinonimia. Como se observa en la Figura IV.1, la polisemia ocurre cuando la misma palabra denota más de un significado y la sinonimia ocurre cuando diferentes palabras denotan el mismo concepto. Por ejemplo, mientras la palabra banco puede denotar más de un significado (por ejemplo, una institución financiera, un edificio de una institución financiera o un asiento) el conjunto de sinónimos constituido por: banco, compañía bancaria e institución financiera de depósito denotan el mismo concepto (Rodríguez y Egenhofer, 2003).

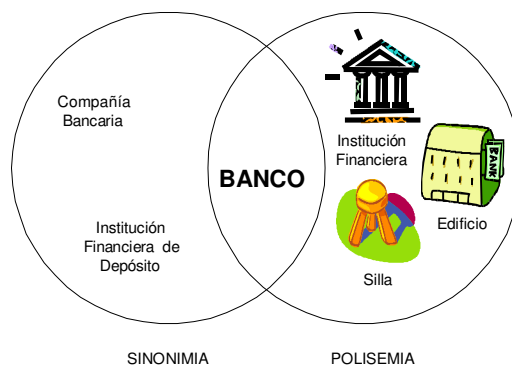


Figura IV.1. Polisemia y Sinonimia

Otras dos relaciones importantes en la especificación de ontologías son la hiponimia y la meronimia. Hiponimia, también llamada relación es-un es la relación más común en una ontología. Esta relación va desde un concepto más específico a uno más general. La relación es-un es transitiva y asimétrica y define una estructura jerárquica donde los términos heredan todas las características de sus términos padres. Meronimia es un orden parcial de tipos de conceptos mediante la relación parte-todo.

4. Conclusiones

Una de las cuestiones claves para la Gestión del Conocimiento es la representación y recuperación del conocimiento y en este capítulo se discutieron dos herramientas fundamentales en este sentido: las Memorias Organizacionales y las ontologías.

Existen numerosas definiciones de Memoria Organizacional generadas fundamentalmente en función del tipo de conocimiento organizacional que tratan de preservar. Independientemente de esta diversidad de enfoques, en esta tesis se enfatiza el papel de la Memoria Organizacional como medio de preservar el

conocimiento organizacional y tratar de vencer la heterogeneidad propia de los objetos de conocimiento pertenecientes a una organización.

La misma multiplicidad de definiciones y enfoques se encuentra con respecto a las ontologías, sin embargo, existe un consenso en relación a su utilidad como medio de modelado y representación del conocimiento organizacional. Particularmente en esta tesis se hará uso de ontologías cuyos componentes de especificación incluyen entidades (clases), relaciones semánticas entre clases y características distintivas de estas clases, mientras que se deja para trabajos futuros el tratamiento de ontologías axiomatizadas más complejas.

Requisitos para un Modelo Conceptual de Gestión del Conocimiento

1. Introducción

A partir del análisis de la situación actual de la Gestión del Conocimiento realizado a lo largo de los capítulos previos, se ha identificado, como aporte de esta tesis, un conjunto de requisitos que un modelo conceptual de gestión del conocimiento organizacional debería satisfacer para constituirse en marco de referencia para una implementación de Gestión de Conocimiento y para el desarrollo de tecnologías de información. Los mismos se discuten en las secciones siguientes.

2. Requisito I: Alineación de las iniciativas de Gestión del Conocimiento con la Estrategia Organizacional

La necesidad de lograr esta alineación se ha convertido en un factor clave para que la organización pueda lograr el éxito tanto en su gestión interna (actividades diarias) como en sus iniciativas de Gestión del Conocimiento. Los trabajadores del conocimiento deben comprender la naturaleza de esta relación para que sus esfuerzos diarios estén direccionados hacia el objetivo estratégico

organizacional. Este requisito se infiere principalmente del análisis realizado en la sección 3 del Capítulo II.

3. Requisito II: Identificación del Conocimiento Organizacional (toma de conciencia)

Es común que los miembros de una organización generalmente no estén concientes de los recursos críticos que permanecen ocultos en la misma. Esta es una de las razones por las cuales mucho del conocimiento es olvidado al corto tiempo de ser desarrollado. La brecha entre los recursos de conocimiento e información colectivos continuamente evolucionando y cambiando de una organización y la conciencia de los empleados de la existencia de tales recursos y de sus cambios pueden conducir a una pérdida en la productividad. Muchas veces una parte de la organización repite el trabajo de otra parte simplemente porque es imposible llevar el rastro y hacer uso del conocimiento generado. Las organizaciones necesitan saber cuáles son sus bienes de conocimiento corporativos y cómo gestionarlos y usarlos para capitalizarlos. Los gerentes deben darse cuenta de lo importante que es saber cuánto saben para hacer un uso óptimo de su conocimiento. Este requisito se infiere principalmente del análisis realizado en la sección 2 del Capítulo II.

4. Requisito III: Estructuración de las Actividades de Gestión de Conocimiento

La Gestión del conocimiento debe ser entendida como la instancia de gestión mediante la cual se obtiene, despliega o utiliza una variedad de recursos básicos para apoyar el desarrollo del conocimiento dentro de la organización. Es por ello que entender cómo estructurar las iniciativas de Gestión del Conocimiento generará una ventaja a la hora de considerar al conocimiento dentro de la estrategia de la organización. Este requisito se infiere principalmente del análisis realizado en la sección 1 del Capítulo III.

5. Requisito IV: Consideración de las Principales Actividades Relacionadas con la Gestión del Conocimiento

Este requisito se infiere principalmente del análisis realizado en la sección 5 del Capítulo III, en el cual las actividades de Gestión del conocimiento se clasifican en tres actividades principales: creación, distribución, y representación y recuperación del conocimiento. En relación a cada una de ellas se identifica un requisito que se describe a continuación.

5.1. Requisito IV.1: En Relación a la Creación del Conocimiento

Sustento de la actividad mediante la vinculación con procesos sociales: la creación del conocimiento es esencialmente un proceso de naturaleza social que conlleva transformaciones entre las formas tácitas y explícitas del conocimiento. Dichas transformaciones (combinación, externalización, internalización y socialización) deben ser sustentadas dentro de la organización para lograr una efectiva creación del conocimiento.

5.2. Requisito IV.2: En Relación a la Distribución del Conocimiento

Contextualización del conocimiento: la idea de contexto está asociada al conocimiento desde su misma definición, siendo básicamente lo que lo diferencia de los datos y la información. Es crucial que al distribuir el conocimiento a través de la organización el mismo esté acompañado de la situación contextual que le dio origen.

Identificación de comunidades de práctica y comunidades de conocimiento: La importancia de estas comunidades radica en el hecho de que el conocimiento, como se mencionó anteriormente, no puede ser separado de su contexto. En todo tipo de actividad de conocimiento, aún donde la tecnología es muy útil, tanto los que contribuyen como los que buscan conocimiento

requieren de una comunidad común para compartir experiencias generales con personas que hacen lo mismo. Mientras que están dentro de una comunidad, los trabajadores están informal y contextualmente unidos por un interés común en compartir conocimiento y aplicar prácticas comunes.

Implementación de Redes de Conocimiento: el establecimiento de redes de conocimiento facilita la transferencia de conocimiento entre los dominios organizacionales y ayuda a canalizar los esfuerzos de los trabajadores. Las redes de conocimiento son, básicamente, los vehículos por excelencia mediante los cuales el conocimiento puede ser comunicado y compartido dentro de una organización.

5.3. Requisito IV.3: En Relación a la Representación y Recuperación del Conocimiento

Contextualización del conocimiento: compartir conocimiento requiere de un marco mental común entre la fuente y el receptor, pero la gente con diferentes trasfondos tienen diferentes estructuras y perspectivas de conocimiento. Más aún, el diseño de repositorios se enfoca en los contenidos y tiende a proveer poco contexto del conocimiento que contiene. El conocimiento es, por definición, altamente dependiente del contexto mientras que, generalmente, una representación explícita tiende a causar una eliminación del contexto. Sin la información contextual, los trabajadores no entienden ni confían completamente en la fuente de conocimiento y por lo tanto no la adoptan

Transparencia de los mecanismos de representación y recuperación de conocimiento (sobrecarga de trabajo): la contribución de conocimiento en los repositorios requiere de un esfuerzo de documentación extra para los trabajadores quienes, a menos que perciban un beneficio inmediato, no justifican el trabajo adicional. Es crucial que los mecanismos utilizados para alimentar y recuperar conocimiento de los repositorios sean lo más transparente posible para los trabajadores del conocimiento.

6. Requisito V: Gestión Distribuida del Conocimiento

Existen varias razones que fundamentan la necesidad de gestionar el conocimiento en forma distribuida (Bonifacio y col., 2002). En primer lugar, están las razones asociadas a las dificultades del desarrollo de un único repositorio centralizado en una organización (dificultad para reducir a un mismo formalismo todo el conocimiento de una organización, cuestiones de duplicación, mantenimiento de versiones, permisos, etc.). Por otro lado, existen ventajas en relación a la gestión distribuida que tienen que ver con la posibilidad de manejar múltiples perspectivas locales dentro de la organización las cuales fomentan la comprensión, el aprendizaje y la innovación. Un análisis más detallado de la necesidad de gestión distribuida del conocimiento se presenta en la sección 2 del Capítulo VII.

7. Requisito VI: Balance entre los Aspectos Sociales y Tecnológicos de la Gestión del Conocimiento

En la mayoría de los casos se observa un excesivo énfasis en el uso de las tecnologías de información para la captura, codificación y distribución del conocimiento existente en una organización. Este esfuerzo apunta mayormente a la conversión del conocimiento tácito a explícito bajo la suposición de que en esta forma es más fácilmente compartido a través de toda la organización. En otras palabras, el foco está en el suministro de información. Sin embargo, como ya se ha mencionado previamente en esta tesis, no es posible ni conveniente explicitar todo el conocimiento que posee una persona. El foco en las tecnologías de información también provoca que se trate al conocimiento como un objeto que puede ser manipulado por un sistema de información. Esta visión de la Gestión del Conocimiento es, como mínimo, una simplificación importante que provoca que se desechen recursos organizacionales que no pueden ser traducidos en un conjunto de categorías y habilidades al ceñirse sólo a enfoques de codificación o representación del conocimiento. Además, como se mencionó

anteriormente, las actividades relacionadas con el conocimiento tienen un alto componente social que no puede desconocerse si se quiere gestionar el conocimiento. Este requisito se infiere principalmente del análisis realizado en las secciones 4 y 5 del Capítulo III.

8. Requisito VII: Cambio en la Cultura Organizacional

En esta nueva era económica, donde las decisiones deben tomarse en un entorno muy cambiante y los errores tienen consecuencias cada vez más difíciles de anticipar, antes de introducir iniciativas de Gestión del Conocimiento, las organizaciones deben preocuparse por crear una cultura organizacional que facilite la creación, la distribución y uso del conocimiento. Una vez lograda, es el momento de hacer énfasis en las tecnologías de información como uno de los mejores medios para lograr la Gestión del Conocimiento. Este requisito se infiere principalmente del análisis realizado en la sección 4.2 del Capítulo III.

9. Satisfacción de Requisitos por parte de los Modelos de Gestión del Conocimiento

En la siguiente tabla se presenta un resumen de los modelos de Gestión del Conocimiento descritos en las secciones 2 y 3 del Capítulo III, y los requisitos descritos en las secciones previas de este capítulo, indicando aquellos que cada modelo verifica.

| Requisito | Wiig | Leonard-Barton | Arthur Andersen | Choo | Van der Spek - Spijkervet | Sveiby | Petrash | Nonaka | Szulanski | Alavi |
|----------------|------|----------------|-----------------|------|---------------------------|--------|---------|--------|-----------|-------|
| Requisito I | x | x | x | x | ✓ | x | x | x | x | x |
| Requisito II | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | x | ✓ | ✓ |
| Requisito III | ✓ | x | ✓ | ✓ | ✓ | x | x | x | x | ✓ |
| Requisito IV.1 | x | x | x | x | x | x | x | ✓ | x | ✓ |
| Requisito IV.2 | x | x | x | x | x | x | x | x | x | x |
| Requisito IV.3 | x | x | x | x | x | x | x | x | x | x |
| Requisito V | x | x | x | x | x | x | x | x | x | x |
| Requisito VI | x | x | x | x | x | x | x | x | x | x |
| Requisito VII | x | x | ✓ | x | ✓ | x | x | x | x | x |

Tabla V.1. Adecuación del los Modelos a los Requisitos

Como se puede observar en la Tabla V.1, ninguno de los modelos conceptuales de gestión del conocimiento propuestos en la bibliografía satisface todos los requisitos identificados en este capítulo.

10. Problemática a Resolver

Los modelos conceptuales de Gestión del Conocimiento propuestos en la bibliografía que, como se mostró en la sección anterior, ninguno satisface todos los requisitos identificados en este capítulo, corresponden en su mayoría a la década del 90 en la que se produjo el auge de la Gestión del Conocimiento. Con posterioridad, se comenzaron a desarrollar adaptaciones de estos modelos para áreas específicas (educación, consultoría, etc.) y se comenzó a hacer foco en las soluciones tecnológicas asociadas a dichos modelos (Despres y Chauvel, 2002) (Xie y col., 2006), (Lei y col., 2006), (Lin y col., 2007), (Li y col., 2008), (Qu y Ren, 2008), (Massa y Testa, 2009), (Mansingh y col., 2009). Los resultados se sintetizan a continuación.

La Gestión del Conocimiento ha visto una evolución significativa a través del periodo de tiempo en el que fue el centro de la atención de la gestión. Durante este periodo, aunque hubo mucho debate sobre la naturaleza del

conocimiento, y el rol de la Gestión del Conocimiento, también ha habido una evaluación crítica relacionada con sus prácticas y puntos de vista comúnmente sostenidos. En años recientes, la Gestión del Conocimiento ha caído bajo escrutinio crítico que cuestiona el éxito de los intentos de gestionar conocimiento, especialmente de aquellos con un intenso foco en las tecnologías de información (Grant y Grant, 2008).

La revisión realizada de la literatura asociada a la Gestión del Conocimiento revela que la mayoría de los trabajos fueron escritos por y para profesionales en tecnologías de información o sistemas de información, lo que sugiere que esta comunidad se ha convertido en una especie de patrón para la Gestión del Conocimiento (Swan y Scarborough, 2002). Este patrocinio tiene efectos contrapuestos, por un lado ofreciendo herramientas que hacen posible la implementación concreta de las actividades básicas de la Gestión del Conocimiento, pero por otro, encasillándola en un área de foco estrecho limitando, por lo tanto, su efectividad.

En algunos casos se observa un excesivo énfasis en el uso de las tecnologías de información para la captura, codificación y distribución del conocimiento existente en una organización (Tuomi, 2002), (Yang y Ho, 2007), (Raghu y Vinze, 2007), (Du y col., 2009), (Nunes y col., 2009). Este esfuerzo apunta mayormente a la conversión del conocimiento tácito a explícito bajo la suposición de que en esta forma es más fácilmente compartido a través de toda la organización. En otras palabras, el foco está en el suministro de información. Sin embargo, como ya se ha mencionado previamente en esta tesis, no es posible ni conveniente explicitar todo el conocimiento que posee una persona.

El foco en las tecnologías de información también provoca que se trate al conocimiento como un objeto que puede ser manipulado por un sistema de información. Esta visión de la Gestión del Conocimiento es, como mínimo, una simplificación importante que provoca que se desechen recursos organizacionales que no pueden ser traducidos en un conjunto de categorías y habilidades al ceñirse sólo a enfoques de codificación o representación del conocimiento. Una consecuencia directa de esta visión es la necesaria

descontextualización del conocimiento para ser gestionado como cualquier otro bien organizacional.

En resumen, el foco en las tecnologías de información y la visión del conocimiento como un objeto que puede ser gestionado como cualquier otro bien organizacional, si bien ha tenido cierto éxito, no ha generado los resultados esperados. Las taxonomías de conocimiento y los modelos orientados a su gestión (discutidos en los Capítulos I y III respectivamente) han sido útiles para entender la naturaleza del conocimiento en las organizaciones pero en sí mismos, no muestran cómo hacer uso efectivo del conocimiento.

Surge entonces una problemática concreta que es la falta de un modelo conceptual unificado y más abarcativo que satisfaga todos los requisitos identificados en esta tesis que sirva como marco de referencia para las iniciativas de Gestión de Conocimiento y de desarrollo de tecnologías de información para su implementación concreta.

11. Conclusiones

En este capítulo, se identificó un conjunto de requisitos que un modelo conceptual de gestión del conocimiento organizacional debería satisfacer para constituirse en marco de referencia para una implementación de Gestión de Conocimiento y para el desarrollo de tecnologías de información; y se ha mostrado que ninguno de los modelos conceptuales de gestión del conocimiento propuestos en la bibliografía satisface todos estos requisitos.

Se identificó entonces como problemática concreta la falta de un modelo conceptual unificado y más abarcativo que satisfaga todos los requisitos identificados en esta tesis que sirva como marco de referencia para las iniciativas de Gestión de Conocimiento y de desarrollo de tecnologías de información para su implementación concreta.

Con el propósito de contribuir a la solución de esta problemática, en esta tesis se plantea como objetivo proponer un modelo conceptual de gestión del

conocimiento organizacional que, a diferencia de los modelos descriptos, satisfaga los requisitos identificados y se constituya en marco de referencia para una implementación de Gestión de Conocimiento y para el desarrollo de tecnologías de información. La esencia de este desafío se presenta como aporte de esta tesis en el Capítulo VI como una serie de factores claves combinados en una propuesta de modelo conceptual unificado y más abarcativo para describir la naturaleza de la Gestión del Conocimiento distribuida.

Modelo Conceptual de Gestión del Conocimiento Distribuido

1. Introducción

En el Capítulo V, los diferentes aspectos del proceso de Gestión del Conocimiento fueron resumidos en un conjunto de requisitos que un modelo conceptual de gestión del conocimiento organizacional debería satisfacer para constituirse en marco de referencia para una implementación de Gestión de Conocimiento y para el desarrollo de tecnologías de información. En dicho capítulo se mostró que los diversos modelos para la Gestión del Conocimiento presentes en la literatura, si bien hacen foco en determinados aspectos del proceso, ninguno los resume a todos. Esto dio lugar al surgimiento de implementaciones específicas que se alejan de las ideas que dan fundamento a la Gestión del Conocimiento.

Para resolver la problemática planteada, en este capítulo se propone como aporte original de esta tesis, un modelo conceptual de gestión del conocimiento organizacional que, a diferencia de los modelos descriptos, satisfaga los requisitos identificados y se constituya en marco de referencia para una implementación de Gestión de Conocimiento y para el desarrollo de tecnologías de información. Dicho modelo está constituido por tres procesos que dan soporte a la creación, representación y distribución del conocimiento.

En la sección 2 se presentan los tres procesos que constituyen los bloques constructores del modelo conceptual propuesto. Dichos procesos ejecutan las actividades involucradas en la Gestión del Conocimiento descritas en la sección 5 del Capítulo III: creación, representación, recuperación e intercambio de conocimiento, desde una perspectiva que excede la sola aplicación de las tecnologías de información.

Se propone, ir más allá de la simple integración de estas tareas relacionándolas con otros procesos claves para la Gestión del Conocimiento vinculando la creación del conocimiento con los procesos de enseñanza, aprendizaje, entrenamiento (*coaching*) y tutoría (*mentoring*) y la generación de conocimiento individual y colectivo. En relación con el intercambio de conocimiento se sostiene que este proceso debe ser fomentado entre los dominios de conocimiento de la organización y que debería ser soportado por una red de conocimiento. Para la representación y recuperación de conocimiento se propone una Memoria Organizacional Distribuida basada en ontologías de dominio y descrita con detalle en el Capítulo VII.

En la sección 3 se presenta el modelo conceptual, y en la sección 4 se realiza un análisis de la adecuación del modelo conceptual propuesto a los requisitos planteados. Finalmente, en la sección 5 se presentan las conclusiones del presente capítulo.

2. Procesos que Componen el Modelo Conceptual

2.1. Proceso de Creación de Conocimiento

Existen dos dimensiones principales del conocimiento las cuales son decisivas para su creación. La primera dimensión describe los niveles de conocimiento distinguiendo quién lo posee: nivel individual, grupal, organizacional e inter-organizacional. La segunda dimensión, y tal vez la más

importante, es el tipo de conocimiento: tácito y explícito (Nonaka y Takeuchi, 1995). Estos tipos de conocimiento no existen independientemente pero pueden ser convertidos uno en el otro como se describió en la sección 5.1 del Capítulo III. Se propone aquí una expansión del modelo de Nonaka sosteniendo que estos procesos de conversión pueden ser efectivamente sostenidos ligándolos a las capacidades de aprendizaje, enseñanza, entrenamiento y tutoría dentro de la organización.

Como se muestra en la Figura VI.1, la *externalización* del conocimiento se refiere a la conversión de conocimiento tácito en conocimiento explícito. Consiste en transformar el conocimiento tácito en una forma explícita de forma tal que pueda ser visto, manipulado y comunicado. Puede ser una articulación individual del conocimiento de una persona que busca el reconocimiento y la expresión de sus ideas, imágenes, modelos mentales, metáforas, analogías, etc., o también puede consistir en la elicitación y expresión del conocimiento tácito de otros convirtiéndolo en explícito.

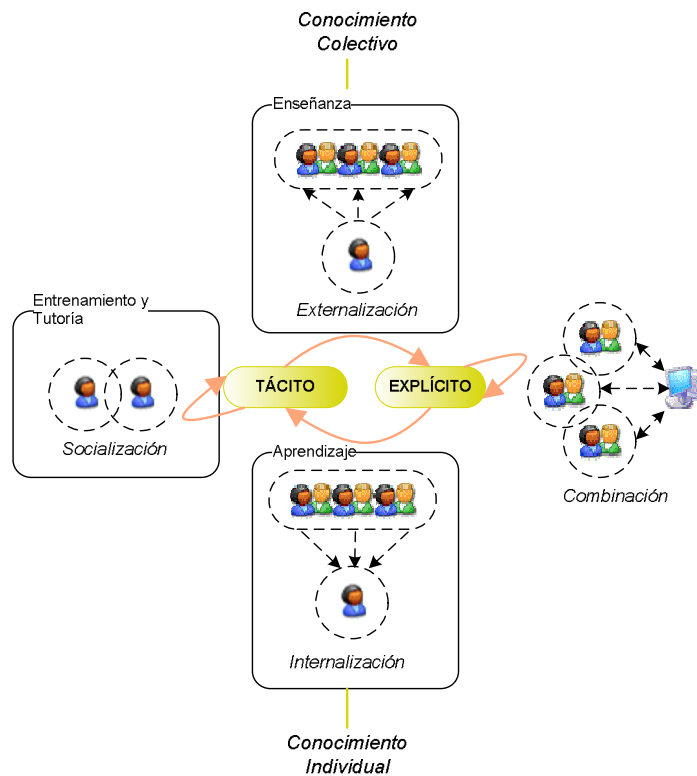


Figura VI.1. Proceso de Creación de Conocimiento

En este proceso, los individuos tratan de articular su conocimiento tácito elicitando sus experiencias y creencias. La externalización describe un proceso de transformación que, por un lado implica la conversión de tácito a explícito, pero además, implica el intercambio de conocimiento entre individuos y un grupo. Debido a que el conocimiento tácito es difícil de expresar, el proceso de conversión es soportado generalmente por el uso de metáforas, analogías y un lenguaje rico en imágenes como por ejemplo modelos, diagramas o prototipos. Es por esta razón que en esta tesis se presenta la externalización desde una perspectiva de proceso de enseñanza que desencadena un círculo virtuoso que favorece la creación del conocimiento (Harris, 2008) como se observa en la Figura VI.2. Entendiendo la externalización como un proceso de enseñanza sostenido no sólo se garantiza una parte primordial del proceso de creación del conocimiento sino que se incrementa el conocimiento colectivo de la organización.

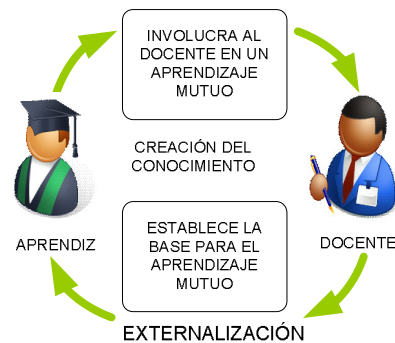


Figura VI.2. Círculo Virtuoso de Enseñanza

El segundo tipo de conversión de conocimiento, la *socialización*, se refiere a la creación de nuevo conocimiento tácito a partir de conocimiento tácito compartido. La visión presentada en esta tesis establece que los individuos dentro de la organización incrementan efectivamente su conocimiento tácito a

través de procesos de entrenamiento y tutoría donde los trabajadores expertos guían a los principiantes en su proceso de aprendizaje. Básicamente, la socialización comprende el intercambio de conocimiento tácito entre individuos para transmitir conocimiento y experiencias personales. La experiencia conjunta resulta en un nuevo conocimiento tácito compartido, como por ejemplo, una nueva habilidad técnica. En la práctica, esto puede significar ganar conocimiento intuitivo y personal a través de proximidad física y alcanzar una comunicación directa con clientes y proveedores.

El entrenamiento es una de las áreas de avanzada en el desarrollo personal, profesional y de negocios. Básicamente, el entrenamiento consiste en ayudar a otros a identificar y definir sus objetivos específicos y conducirlos en la obtención de dichas metas. Esta actividad trata con la construcción de habilidades individuales en las personas, desde la formación de objetivos hasta la toma de decisiones y la resolución de problemas. Los entrenadores recurren a la creatividad, los recursos y el conocimiento de los aprendices para ayudarlos a ser más efectivos.

Los beneficios de las actividades de entrenamiento son numerosos, permitiendo la entrega de resultados extraordinarios, el alcance de objetivos individuales y organizacionales y la ejecución de planes y proyectos. Lo que es más, debido al énfasis en el aprendizaje y el desarrollo de la confianza en el aprendiz, el incremento en el desempeño es sostenido generalmente por un largo periodo de tiempo y tiene un impacto en áreas que no están directamente relacionadas con aquella bajo entrenamiento.

Con respecto a la tutoría los beneficios se relacionan con el desarrollo de profesionales con un alto potencial de liderazgo, se acorta la curva de aprendizaje e incrementa la retención y el compromiso con la organización aumentando al mismo tiempo su capital intelectual (Murray, 2001). Como puede observarse en la Tabla VI.1, existen diferencias entre el entrenamiento y la tutoría. La tutoría es una situación de aprendizaje en ambos sentidos, mutuamente beneficiosa, donde el tutor provee consejo, comparte conocimiento

y experiencias y enseña usando un enfoque sin presiones y de auto-descubrimiento.

| | Tutor | Entrenador |
|---------------------------------|--|---|
| Foco | Individual: el foco está en la persona. Puede dar consejos pero el aprendiz puede elegir qué hacer. El tutor tiene un profundo interés y se involucra personalmente | Desempeño: enfocado en la tarea y orientado al desempeño. El entrenador trata de guiar a la persona hacia un resultado final monitoreando continuamente su progreso. El entrenador es imparcial, enfocado exclusivamente en el mejoramiento del desempeño. |
| Rol | Facilitador sin agenda: los tutores son facilitadores y maestros que le permiten al aprendiz descubrir su propia dirección. | Agenda específica: el entrenador tiene una agenda específica orientada a reforzar o cambiar las habilidades y comportamientos. |
| Relación | Selección propia: la relación surge por mutua elección entre el tutor y el aprendiz, pudiendo decidir la duración, la temática, etc. | Asociado al trabajo: en una organización, el entrenamiento viene dado por una tarea en particular que el aprendiz debe dominar. |
| Fuente de la influencia | Valor percibido: la influencia del tutor viene dada por el valor que se percibe puede aportar el mismo a la relación. | Posición: la influencia o autoridad del entrenador viene dada por su posición. |
| Retorno a nivel personal | Afirmación/aprendizaje: existe un proceso de aprendizaje por parte del tutor que surge de la retroalimentación y razonamientos del aprendiz. | Trabajo en equipo/desempeño: el beneficio para el entrenador viene dado por un mejor trabajo en equipo y un aumento en el desempeño de sus aprendices. |
| Ámbito | La vida: la figura del tutor no es exclusiva del ámbito organizacional. | Relacionada a una tarea: el entrenamiento está relacionado a una tarea en particular. |

Tabla VI.1. Comparación entre Tutores y Entrenadores (adaptado de <http://www.coachingandmentoring.com/Articles/mentoring.html>)

La *combinación* de conocimiento se refiere a la creación de nuevo conocimiento mediante el intercambio de conocimiento explícito en la organización. Consiste en tomar conocimiento explícito y explícito, combinarlo con otro conocimiento explícito y desarrollar nuevo conocimiento explícito. Aquí es donde las tecnologías de información son útiles, porque el conocimiento explícito es incorporado en artefactos (por ejemplo, documentos) que pueden ser coleccionados, manipulados y diseminados permitiendo la transferencia de conocimiento a través de la organización.

Se propone en esta tesis que este proceso debería ser soportado por un Sistema de Memoria Organizacional Distribuida, descrito en el Capítulo VII, que permite a los trabajadores ordenar, reutilizar, agregar y re-contextualizar conocimiento explícito. En este punto, parte del conocimiento tácito adquirido ha atravesado un proceso de cristalización durante el cual fue evaluado y validado por expertos y finalmente es puesto a disposición para su distribución y utilización.

El cuarto tipo de conversión de conocimiento, la *internalización*, toma lugar cuando el conocimiento explícito se convierte en tácito. Consiste en el entendimiento y absorción del conocimiento explícito compartido colectivamente en conocimiento tácito individual procesable por el poseedor. Una vez que una persona ha aprendido un proceso profundamente, se convierte en algo interno y puede ser usado de forma automática o refleja. La internalización es en su mayoría experimental a través de la ejecución efectiva de las tareas, en situaciones reales o simulaciones.

En esta tesis se presenta la internalización como un proceso concreto de aprendizaje donde los individuos incorporan nuevo conocimiento actualizando sus modelos mentales. La internalización comprende la conversión del conocimiento organizacional explícito en conocimiento tácito individual. Este conocimiento tácito y las experiencias ganadas a nivel individual pueden ser compartidos nuevamente con otros (socializados) para convertirse en conocimiento organizacional, poniéndose, de esta forma, nuevamente en movimiento una vez más el ciclo de creación del conocimiento.

La internalización mediante el aprendizaje permite a los individuos y a las organizaciones cambiar rutinas de comportamiento condicionadas haciéndolos más adaptativos para generar y evaluar un amplio rango de hipótesis sobre el futuro y seleccionar las herramientas mentales y entornos apropiados para mejorar sus capacidades de aprendizaje, reflexión e implementación.

El aprendizaje organizacional requiere de aprendizaje individual, y al mismo tiempo, el aprendizaje individual tiene que interactuar en un entorno

social dinámico para poder contribuir al aprendizaje organizacional. Esta relación entre aprendizaje individual y organizacional puede ser conceptualizada como una espiral de creación del conocimiento donde las compañías pueden convertirse en *organizaciones que aprenden*, permitiendo y gestionando los procesos de conversión de conocimiento dinámicos entre los individuos y la organización, y entre el conocimiento explícito y tácito.

Pueden identificarse dos tipos de aprendizaje organizacional, el aprendizaje incremental y el aprendizaje radical. El aprendizaje incremental se caracteriza por la resolución de problemas simples y rutinarios y que no requiere un cambio en el sistema de pensamiento. El aprendizaje radical es un tipo de aprendizaje innovativo que desafía directamente el modelo mental prevaleciente sobre el cual se apoya el sistema de razonamiento. Es este tipo de aprendizaje el que constituye la base de la creación del conocimiento organizacional y que debe ser sustentado por los procesos de entrenamiento y tutoría.

La Gestión del Conocimiento debería propiciar la construcción de un entorno de aprendizaje donde las personas puedan recurrir a un conjunto de recursos para comprender determinadas situaciones y construir soluciones significativas. En este ambiente, debería enfatizarse la importancia de actividades que ayuden al aprendiz a construir entendimiento y a desarrollar herramientas relacionadas con la resolución de problemas.

2.2. Proceso para Compartir Conocimiento

Uno de los prerequisites para permitir la colaboración entre individuos con diferentes trasfondos en términos de dominios y niveles de experticia es la habilidad del grupo para crear un sentido de mutualidad y por lo tanto un marco de referencia compartido.

En esta tesis, se establece que el proceso de compartir conocimiento debe ser fomentado entre los dominios de conocimiento de la organización. El filtrado de conocimiento desde estos dominios de conocimiento a los procesos clave de

la organización es realizado a través de comunidades de práctica y comunidades de conocimiento. En la sección 5.2.2. del Capítulo III se describió la naturaleza de estas comunidades y su rol en la transferencia de conocimiento.

La transferencia de conocimiento ocurre cuando el conocimiento es difundido desde una entidad (por ejemplo, un individuo) a otras. Esta transferencia puede desarrollarse a través de un proceso de socialización, educación y/o aprendizaje y puede ser un proceso intencional o puede ocurrir como resultado de otra actividad.

Se estableció anteriormente que los mecanismos por excelencia para la transferencia de conocimiento son las comunidades de práctica y las comunidades de conocimiento. La importancia de estas comunidades radica en el hecho de que el conocimiento no puede ser separado de su contexto. En todo tipo de actividad de conocimiento, aún donde la tecnología es muy útil, tanto los que contribuyen como los que buscan conocimiento requieren de una comunidad común para compartir experiencias generales con personas que hacen lo mismo. Mientras que están dentro de una comunidad, los trabajadores están informal y contextualmente unidos por un interés común en compartir conocimiento y aplicar prácticas comunes.

Un desafío encontrado en las iniciativas de Gestión del Conocimiento es cómo conectar estas comunidades (comunidades de conocimiento y comunidades de práctica) para permitir compartir conocimiento no sólo dentro sino entre comunidades. La solución propuesta en esta tesis, como se observa en la Figura VI.3, es establecer una red de conocimiento entre ellas. Una red de conocimiento es una red informal de comunidades entre dominios de conocimiento que facilita la transferencia de conocimiento entre los dominios y ayuda a canalizar los esfuerzos de los trabajadores.

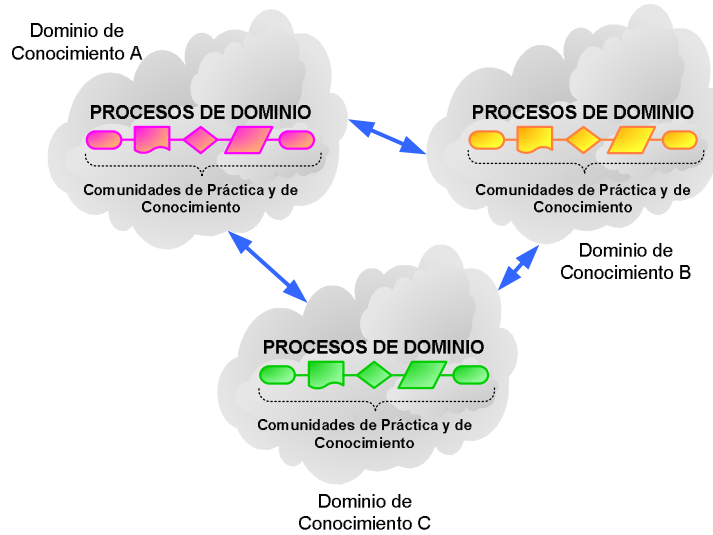


Figura VI.3. Red de Conocimiento

Las redes de conocimiento son vehículos mediante los cuales el conocimiento puede ser comunicado y compartido y particularmente en esta tesis, se propone dar soporte a una red de conocimiento a través de una Memoria Organizacional Distribuida como se describe en la siguiente sección.

2.3. Proceso de Representación y Recuperación de Conocimiento

Las técnicas de adquisición del conocimiento van de la mano con las cuestiones de representación del conocimiento ya que la manipulación del conocimiento adquirido depende largamente de cómo éste se representa para reflejar el modelo mental de un experto. Recordar lo que una organización ha aprendido y reutilizar el conocimiento relevante (generado internamente o adquirido externamente) es otro de los aspectos importantes para la Gestión del Conocimiento. Al igual que los individuos, las organizaciones pueden perder el rastro de lo que saben y olvidar. Por lo tanto las organizaciones codifican su conocimiento para poder preservarlo y reutilizarlo. Debido a que las organizaciones están continuamente involucradas en el proceso de generación y

aplicación de conocimiento, sería inútil intentar codificar y almacenar todo el conocimiento organizacional y, como se mencionó antes, el esfuerzo de codificación del conocimiento dependerá de su tipo (tácito o explícito).

Para enfrentar estos inconvenientes surge la necesidad de desarrollar sistemas de información facilitadores de conocimiento que provean un marco común para capturar, incrementar, almacenar, organizar, analizar y compartir no sólo información y datos sino también conocimiento. Actualmente, según se describió en la sección 2 del Capítulo IV, las Memorias Organizacionales se proponen como soporte para un uso y preservación efectiva del conocimiento a través del tiempo y el espacio (tanto como sea posible) sin intervención humana. Desde la perspectiva organizacional, una Memoria Organizacional puede actuar como una herramienta para la Gestión del Conocimiento facilitando la transmisión del conocimiento, los tres tipos de aprendizaje organizacional (aprendizaje individual, aprendizaje a través de comunicación directa, y aprendizaje usando un repositorio de conocimiento) y en última instancia, la creación del conocimiento.

Una Memoria Organizacional comprende una variedad de fuentes de conocimiento donde están disponibles elementos de información de diferentes clases, estructuras, contenidos y tipos de medio y debería ser capaz de controlar y acceder estas fuentes de conocimiento heterogéneas de acuerdo a las necesidades de información de los usuarios. A pesar de que la definición previa sugiere un enfoque centralizado, la centralización de una Memoria Organizacional presenta algunas desventajas relacionadas con la naturaleza distribuida del conocimiento organizacional y al alto costo de mantenimiento de una estructura centralizada. Estas razones conducen a proponer en esta tesis un enfoque de Memoria Organizacional Distribuida que apunta a resolver las problemáticas comunes en el uso de repositorios en las organizaciones (falta de contribución por parte de los trabajadores, descontextualización del conocimiento, etc., discutidos en la sección 5.3 del Capítulo III). Este enfoque es descrito con detalle en el capítulo siguiente (Ale y col., 2004a, 2004b).

3. Modelo Conceptual de Gestión del Conocimiento Distribuido

La problemática identificada en la sección 10 del Capítulo V, planteó la necesidad de un nuevo modelo conceptual que resuelva las limitaciones de los enfoques anteriores. El modelo propuesto en esta tesis apunta a cubrir esta necesidad teniendo en cuenta una nueva tendencia de pensamiento sobre el uso del conocimiento en las organizaciones la cual reconoce una mayor complejidad en los desafíos relacionados al conocimiento que enfrentan las organizaciones. Estos desafíos surgen por la naturaleza altamente personal del conocimiento y la dificultad de institucionalizar este conocimiento sin tomar en cuenta una variedad de grupos y temas sociales.

En contraposición el excesivo énfasis en el suministro de conocimiento, en el modelo conceptual propuesto se contemplan los procesos de creación e integración de conocimiento desde dos puntos de vista (individual y colectivo). Se hace énfasis en la demanda de conocimiento y no en su suministro, ya que se entiende que los procesos de conocimiento son, en esencia, sociales y por lo tanto, asociados a procesos de enseñanza, aprendizaje, tutoría y entrenamiento dentro de comunidades en la organización.

En contraposición a la idea del conocimiento como un objeto que puede ser manipulado por un sistema de información, esta tesis considera factor clave para la Gestión del Conocimiento la creación de un contexto compartido. Para poder compartir efectivamente conocimiento debe existir un correcto balance entre contexto, contenido y, a su vez, el nivel de abstracción utilizado. En el modelo conceptual propuesto esto se aborda mediante el uso de una ontología de dominio (Brusa y col., 2008) como artefacto para representar tanto contenido (conocimiento) como contexto y permite seleccionar el nivel de abstracción adecuado para cada comunidad dentro de la organización.

En contraposición al foco de atención en la tecnología, esta tesis propone mover el foco de atención a las personas que integran la organización. Para lograr esto es necesario identificar dentro de la organización los dominios de

conocimiento existentes como así también las comunidades de práctica y de conocimiento que los integran, conectándolos en una red de conocimiento que potencie su papel en las actividades de creación y distribución de conocimiento. La identificación de los dominios de conocimiento se basará en la inspección y discusión de las actividades de la compañía junto con gerentes y expertos de la organización. En esta tesis se establece, que los enfoques de repositorio y de red (descritos en la sección 4 del Capítulo III) no son mutuamente excluyentes y que ambos son necesarios para una implementación de Gestión del Conocimiento.

En la Figura VI.4 se esquematiza el Modelo Conceptual de Gestión del Conocimiento Distribuido propuesto que abarca los tres procesos que dan soporte a la creación, representación y distribución del conocimiento, descritos en las secciones previas. Este modelo consiste de un Sistema de Memoria Organizacional Distribuida y un Sistema de Gestión del Conocimiento, los cuales se describen a continuación:

Sistema de Memoria Organizacional Distribuida: consiste de los procesos y componentes informáticos usados para capturar, almacenar, buscar, recuperar, visualizar y manipular objetos de conocimiento. En el modelo presentado en esta tesis correspondería en un principio a la implementación de un modelo de repositorio o enfoque tecnológico. Sin embargo, se salvan las cuestiones de la actualización mediante una estrategia automática, las debilidades de la centralización mediante un enfoque distribuido y la descontextualización del conocimiento a través del uso de una ontología de dominio.

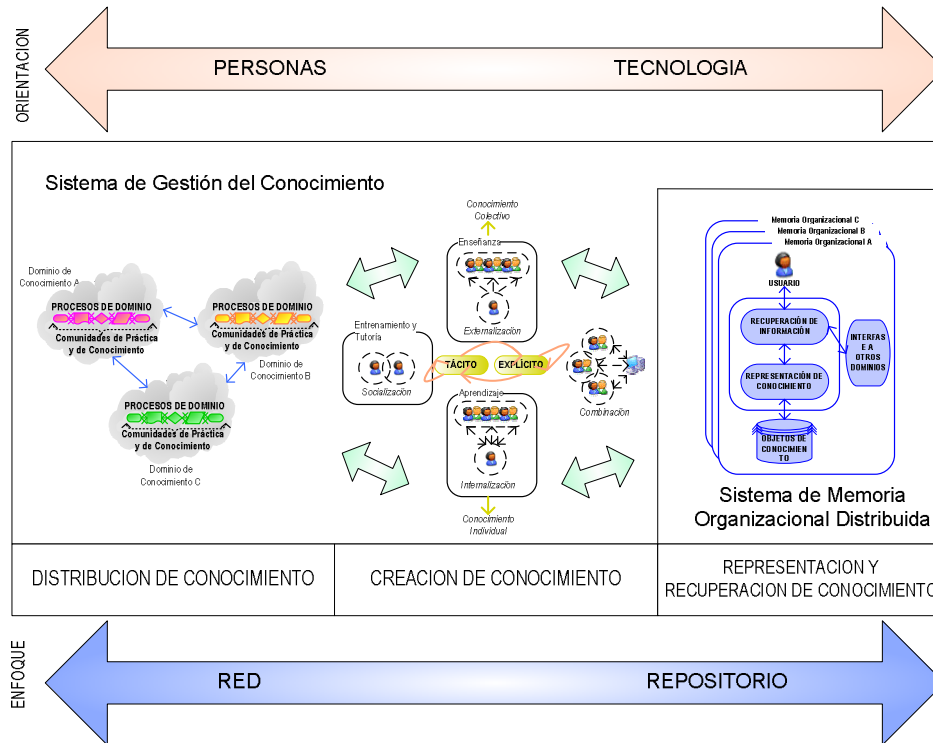


Figura VI.4. Modelo Conceptual de Gestión del Conocimiento Distribuido

Sistema de Gestión del Conocimiento: consiste de las herramientas y procesos usados por los trabajadores para identificar y transmitir conocimiento a los repositorios contenidos en una Memoria Organizacional. En este punto, el Sistema de Memoria Organizacional Distribuida ofrece el soporte necesario para la implementación de comunidades de conocimiento, comunidades de práctica y en definitiva, redes de conocimiento que sirven como base para el desarrollo de un modelo de red de conocimiento.

4. Adecuación del Modelo Propuesto a los Requisitos

No existen estándares para la evaluación de modelos conceptuales por lo que los mismos se evalúan de forma ad hoc, basados en el sentido común, opiniones subjetivas o la experiencia (Moody, 2005). Un enfoque no empírico comúnmente utilizado es la evaluación del modelo con respecto al cumplimiento

de un conjunto de requisitos y la comparación con otros modelos, en cuanto al cumplimiento de dichos requisitos (Recker, 2005). Esta es la estrategia utilizada en esta tesis para evaluar el modelo propuesto. En la Tabla VI.2 se presenta el análisis de cumplimiento del conjunto de requisitos, definidos en el Capítulo V, por parte del Modelo Conceptual de Gestión del Conocimiento Distribuido propuesto en esta tesis.

| Requisito | Adecuación del Modelo Propuesto |
|---|--|
| Requisito I: Alineación de las Iniciativas de Gestión del Conocimiento con la Estrategia Organizacional | En el modelo propuesto el origen de la aplicación de dicho modelo parte de un concienzudo análisis de los dominios de conocimiento presentes en la organización, seleccionándose aquellos que se consideren prioritarios para la actividad organizacional. Dicho criterio de selección pone de manifiesto aquellas áreas que son claves para el desarrollo de la estrategia organizacional. |
| Requisito II: Identificación del Conocimiento Organizacional (toma de conciencia) | En el modelo propuesto la identificación del conocimiento organizacional se da en varios niveles. A nivel organizacional se deben identificar los dominios de conocimiento prioritarios dentro de la organización junto con los principales referentes dentro de las comunidades que los componen. A nivel de cada uno de los dominios de conocimiento identificados se realiza un modelado explícito del conocimiento relevante a través de una ontología de dominio. |
| Requisito III: Estructuración de las Actividades de Gestión del Conocimiento | El modelo propuesto ofrece una guía de los principales aspectos a identificar dentro de la organización (comunidades, dominios, redes, etc.) y ofrece una herramienta concreta para la implementación de las actividades principales de Gestión del Conocimiento (memoria organizacional distribuida) |
| Requisito IV.1: En Relación a la Creación del Conocimiento ✓ sustento de la actividad mediante la vinculación con procesos sociales Requisito IV.2: En Relación a la Distribución del Conocimiento ✓ contextualización del conocimiento ✓ identificación de las | <p>En el modelo propuesto en esta tesis, se propone sustentar dichas transformaciones mediante el incentivo de los procesos de aprendizaje, enseñanza, tutoría y entrenamiento que se desarrollan diariamente en las organizaciones y que son más fácilmente comprendidas por los trabajadores del conocimiento.</p> <p>A nivel organizacional, la contextualización del conocimiento se presenta a través de la identificación de los dominios de conocimiento y las comunidades que los generan. De esta forma, la fuente, perspectiva y situación que dio origen a un objeto de conocimiento es preservada.</p> |

| | |
|--|---|
| <p>Comunidades de Práctica y Comunidades de Conocimiento</p> <p>✓ implementación de Redes de Conocimiento</p> <p>Requisito IV.3: En relación a la Representación y a la Recuperación</p> <p>✓ contextualización del conocimiento</p> <p>✓ transparencia de los mecanismos de representación y recuperación de conocimiento (sobrecarga de trabajo)</p> | <p>La identificación de los dominios de conocimiento se basa en la inspección y discusión de las actividades de la compañía junto con gerentes y expertos de la organización. Típicamente, el conocimiento declarativo (saber qué) y el conocimiento funcional principal (saber cómo) relacionados con las actividades primordiales de la compañía delimitan los dominios de conocimiento más representativos de la organización. Dentro de cada dominio se identifican aquellas comunidades representativas que ayudarán a la definición de la ontología de dominio.</p> <p>Particularmente en esta tesis, se propone dar soporte a una red de conocimiento a través de una memoria organizacional distribuida</p> <p>A un nivel de dominio, las actividades de representación y recuperación de conocimiento tienen en cuenta el contexto del conocimiento a través de la utilización de una ontología de dominio.</p> <p>En este sentido, los mecanismos de representación del conocimiento dentro de la memoria organizacional distribuida (proceso de anotado semántico) están ocultos al usuario y en cuanto a la recuperación, se le ofrece al usuario la posibilidad de plantear consultas en lenguaje natural evitando el aprendizaje y uso de cualquier formalismo.</p> |
| <p>Requisito V: Gestión Distribuida del Conocimiento</p> | <p>Con respecto a la gestión distribuida del conocimiento, el modelo ofrece el establecimiento de una red de conocimiento mediante la implementación de una memoria organizacional en cada uno de los dominios identificados.</p> |
| <p>Requisito VI: Balance entre los Aspectos Sociales y Tecnológicos de la Gestión del Conocimiento</p> | <p>En el Modelo Conceptual propuesto se contemplan los procesos de creación e integración de conocimiento desde varios puntos de vista (individual y colectivo). Se hace además, énfasis en la demanda de conocimiento y no en su suministro, ya que se entiende que los procesos de conocimiento son, en esencia, sociales y por lo tanto, asociados a procesos de enseñanza, aprendizaje, tutoría y entrenamiento dentro de comunidades en la organización.</p> |
| <p>Requisito VII: Cambio en la Cultura Organizacional</p> | <p>La implementación del modelo propuesto excede la mera implantación de una solución tecnológica, conduciendo a la organización a través de un proceso que permite no solo la identificación del conocimiento organizacional sino también de los recursos humanos claves como así también de las relaciones planteadas entre ellos. Esta toma de conciencia de la importancia del conocimiento y sus recursos asociados conlleva un cambio cultural dentro de la organización.</p> |

Tabla VI.2. Adecuación del Modelo Conceptual Propuesto a los Requisitos

El análisis sintetizado en la Tabla VI.2, muestra que el modelo conceptual propuesto, satisface los requisitos identificados en el Capítulo V. De modo que si se lo compara con los modelos conceptuales de gestión de conocimiento propuestos en la bibliografía, cuyo análisis de satisfacción de dichos requisitos se presentó en la sección 9 del Capítulo V, es posible concluir que el modelo conceptual propuesto es una alternativa superadora, pudiendo servir como marco de referencia para una implementación de Gestión de Conocimiento y para el desarrollo de tecnologías de información.

5. Conclusiones

Se ha propuesto un Modelo Conceptual de Gestión del Conocimiento organizacional que verifica todos los requisitos identificados, constituyéndose de este modo en una alternativa de solución a la problemática planteada en la sección 10 del Capítulo V. El mismo sirve de guía para implementar una iniciativa de Gestión de Conocimiento en una organización; para evaluar la situación actual, con relación al tema, en un momento dado de la vida organizacional; o como marco de referencia para el desarrollo de tecnologías de información. Este modelo fue generado tomando en consideración las limitaciones de otros modelos y las críticas más comunes asociadas al fracaso de las iniciativas de Gestión del Conocimiento sintetizadas en la sección 10 del Capítulo V.

El proceso de creación de conocimiento, del modelo conceptual propuesto, se lleva a cabo a través de cuatro procesos de conversión entre comunidades configuradas en una red de conocimiento que constituyen el *Sistema de Gestión del Conocimiento*. Esta red de conocimiento se implementa en función de un *Sistema de Memoria Organizacional Distribuida* que ayuda a canalizar los esfuerzos de los trabajadores. Este último sistema constituye la parte implementable del modelo.

En lo que sigue, esta tesis se centra en una propuesta de una arquitectura para el *Sistema de Memoria Organizacional Distribuida*, la cual se presenta en detalle en el Capítulo VII, su implementación prototipo en el Capítulo VIII, y su validación empírica en el Capítulo IX. Queda fuera del alcance de esta tesis la implementación del *Sistema de Gestión del Conocimiento*.

Arquitectura del Sistema de Memoria Organizacional Distribuida

1. Introducción

Las organizaciones son sociedades de seres: si estos seres tienen memorias individuales, la organización naturalmente desarrollará una memoria colectiva que será como mínimo la sumatoria de todas estas memorias individuales pero generalmente representa mucho más. En capítulos anteriores se ha establecido la relación que tiene la Gestión del Conocimiento con las Memorias Organizacionales y en el presente capítulo se desarrolla un Sistema de Memoria Organizacional Distribuido basado en un modelo propio de arquitectura tendiente a sobrellevar las limitaciones asociadas comúnmente a este tipo de repositorios. Se tomó como marco de referencia para el desarrollo de dicho sistema el Modelo Conceptual de Gestión del Conocimiento organizacional propuesto en el Capítulo VI.

En la sección 2 se analiza con detalle la necesidad de realizar una Gestión del Conocimiento Distribuida, planteada como Requisito V en la sección 6 del Capítulo V.

En la sección 3 se propone la arquitectura para un Sistema de Memoria Organizacional Distribuida describiendo las estrategias propuestas para cada uno

de los componentes. Finalmente, en la sección 4 se presentan las conclusiones del presente capítulo.

2. Necesidad de una Gestión Distribuida

Según se discutió en la sección 10 del Capítulo V, una falla sustancial en la mayoría de los proyectos de Gestión de Conocimiento se origina, no sólo por problemas tecnológicos, sino en una visión objetivista, en la cual los aspectos sociales, contextuales y subjetivos del conocimiento son considerados como “ruido”, características del conocimiento en su estado “bruto” que tienen que ser eliminadas vía abstracción.

El éxito de una vista epistemológica objetivista en la Gestión del Conocimiento se explica al notar que encuadra muy bien con el paradigma centralizado y tradicional de control gerencial. Sin embargo, un gran número de investigadores, trabajando en diferentes disciplinas han argumentado convincentemente en contra de esta vista objetivista. El argumento básico es que lo que sabemos no puede ser visto simplemente como una fotografía del mundo, ya que siempre presupone algún grado de interpretación. Más aún, no existe un lenguaje ideal que simplemente describa el mundo tal cual es. Ciertamente, dependiendo de diferentes esquemas de interpretación, las personas pueden usar la misma palabra con diferentes significados o diferentes palabras para significar lo mismo; dos personas pueden observar el mismo fenómeno pero ver diferentes problemas, diferentes oportunidades y diferentes desafíos (Bonifacio y col., 2002).

Esta vista epistemológica, en la cual la parte explícita de lo que sabemos obtiene su significado (típicamente implícito o sobreentendido) de un esquema de interpretación, conduce a algunas consecuencias importantes relacionadas al uso de clasificaciones y otras conceptualizaciones en los sistemas de Gestión del Conocimiento. Ciertamente, se desprende de lo anterior, que una clasificación no es una organización neutral de una colección de ítems (por ejemplo, documentos

en una base de datos), sino es la emergencia de un esquema de interpretación, de acuerdo al cual tiene sentido clasificar las cosas de esa manera. Resumiendo, una clasificación es siempre el resultado de un proceso de racionalización y representa el punto de vista de aquellos que tomaron parte del proceso.

Si esto es verdad, entonces hay al menos tres buenas razones para permitir múltiples clasificaciones dentro de un sistema de Gestión del Conocimiento de una organización compleja:

- ✓ en primer lugar, trabajando en su clasificación local, cada unidad organizacional hace explícita y más fuerte una perspectiva común, y ayuda a entender las prácticas diarias y el conocimiento laboral (know how) local;
- ✓ en segundo lugar, la resultante de ese esfuerzo de comprensión es una importante fuente de valor para la organización, ya que permite a cada unidad acceder a diferentes puntos de vista en la organización, y obtener una intuición de cómo luciría el mundo desde una perspectiva diferente;
- ✓ finalmente, la continua interacción de múltiples perspectivas locales es el factor clave para desencadenar la innovación. En analogía a lo que pasa en la investigación científica, la exposición a diferentes perspectivas es tan importante para la innovación como el hecho de que un investigador pertenezca a una fuerte comunidad científica, y es tal vez el disparador más importante de sucesos intelectuales originales.

Por lo tanto, es necesario repensar la forma en que se diseñan los sistemas de Gestión del Conocimiento ya que los mismos deben dar soporte a dos procesos cualitativamente diferentes: la gestión autónoma de clasificaciones locales, y su coordinación para su interacción y la distribución del conocimiento.

Este es uno de los objetivos principales del Sistema de Memoria Organizacional Distribuida propuesto en esta tesis, cuya arquitectura se describe a continuación desarrollando las implicaciones tecnológicas con detalle en el capítulo siguiente.

3. Arquitectura de un Sistema de Memoria Organizacional Distribuida

Lo que se señaló en la sección anterior tiene consecuencias relevantes en la forma en que se deben diseñar los sistemas de Gestión del Conocimiento para organizaciones complejas. Los requerimientos para una gestión distribuida del conocimiento pueden resumirse en los siguientes principios:

- ✓ *Principio de autonomía:* cada unidad organizacional debería tener un alto grado de autonomía para gestionar el conocimiento local. La autonomía puede darse a varios niveles (por ejemplo, tecnológico, de formato, proceso de contribución, etc.). En el caso particular de la arquitectura presentada en esta tesis, la autonomía tiene relación con la semántica, viéndola como la posibilidad de elegir la perspectiva más apropiada de lo que se conoce localmente, más específicamente mediante el desarrollo de una ontología de dominio.

- ✓ *Principio de coordinación:* diferentes unidades deberían ser capaces de intercambiar conocimiento a través de mecanismos de interoperabilidad entre perspectivas autónomas mediante la posibilidad de acceso a otras perspectivas no locales.

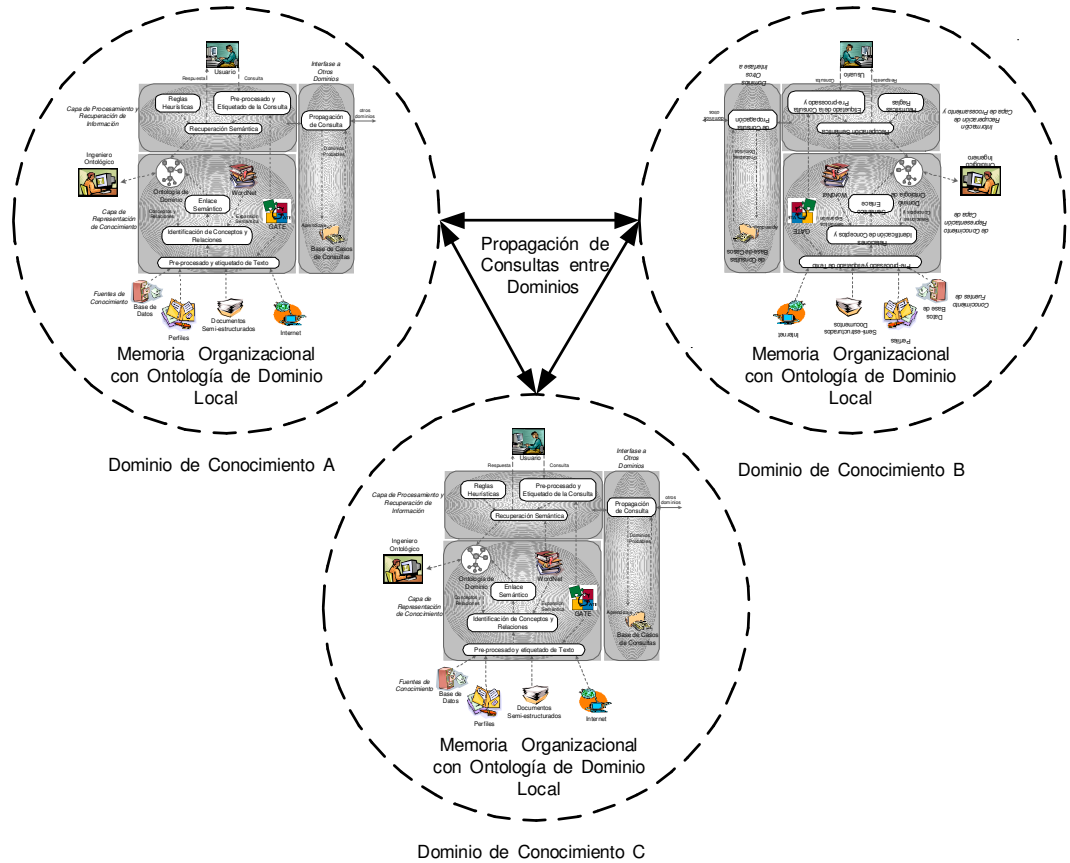


Figura VII.1. Arquitectura de Gestión de Conocimiento Distribuida a través de un Esquema de Memorias Organizacionales

Los sistemas de Gestión del Conocimiento actuales no satisfacen los dos principios expuestos ya que tienden a la creación de una única clasificación compartida (contraria al primer principio) y a la implementación de la coordinación mediante la eliminación de perspectivas (más que su interoperación). En esta tesis en cambio se propone un Sistema de Memoria Organizacional Distribuida cuya arquitectura, representada esquemáticamente en la Figura VII.1, consiste de los siguientes bloques principales:

- ✓ *Dominios de Conocimiento*: se piensa a la organización como una constelación de dominios de conocimiento, donde un dominio de conocimiento es una unidad, ya sea formalmente definida en el organigrama de la compañía (por ejemplo una división organizacional) o informal (por ejemplo un grupo de interés) la cual detenta alguna forma de autonomía. En el modelo conceptual de gestión de conocimiento distribuido propuesto en la sección 3 del Capítulo VI, la identificación de los dominios de conocimiento se basa en la inspección y discusión de las actividades de la compañía junto con gerentes y expertos de la organización. Típicamente, el conocimiento declarativo (saber qué) y el conocimiento funcional principal (saber cómo) relacionados con las actividades primordiales de la compañía delimitan los dominios de conocimiento más representativos de la organización.

- ✓ *Memoria Organizacional con Ontología de Dominio Local*: permite el establecimiento de un contexto en el cual es posible definir una perspectiva local (a través de la ontología de dominio) de las diferentes fuentes de conocimiento del dominio.

- ✓ *Propagación de Consulta entre Dominios*: cada Memoria Organizacional posee una interfaz que le permite la propagación de las consultas a otros dominios en caso de ser necesario.

En este tipo particular de Memoria Organizacional, como se observa en la Figura VII.2, las características, atributos y semántica de los objetos de conocimiento, como también las relaciones entre ellos, se representan a través de una ontología de dominio (Ale y col., 2005a, 2005b, 2005c).

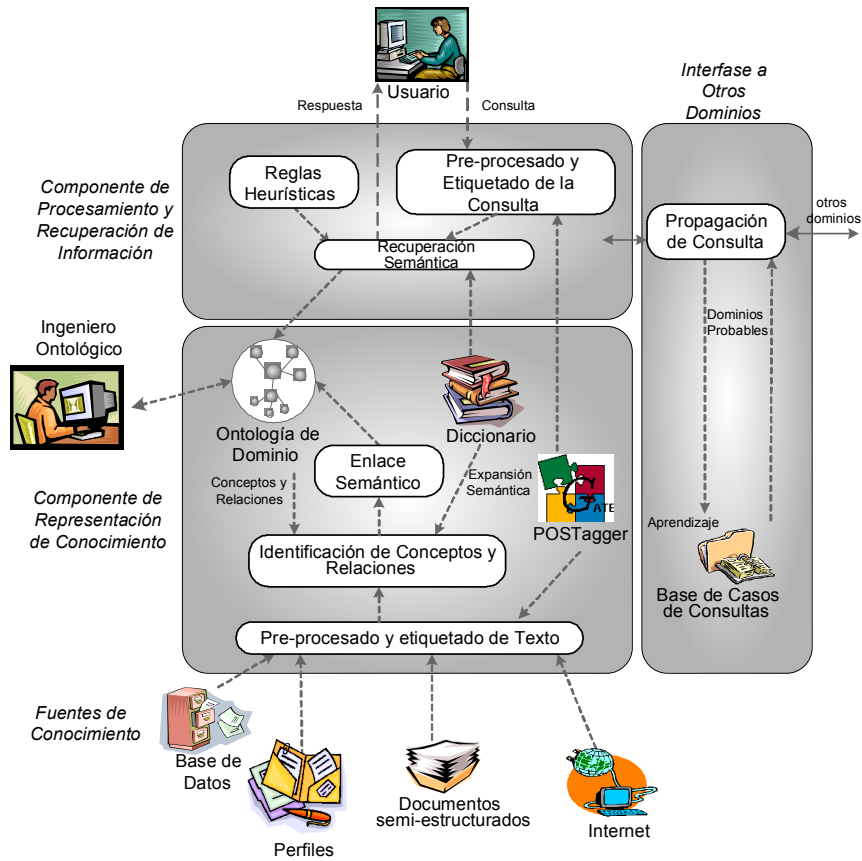


Figura VII.2. Representación del Conocimiento y Recuperación de Información a través de una Memoria Organizacional en un Dominio de Conocimiento

Como se mencionó en la sección 3 del Capítulo IV, las ontologías tienen como meta capturar el conocimiento de dominio en una forma genérica y proveen un entendimiento comúnmente consensuado de un dominio, que puede ser reutilizado, compartido y aplicado en aplicaciones y grupos. Un beneficio adicional ofrecido por las ontologías y explotado en esta arquitectura es la representación de contexto. Las ontologías proveen un modelo de dominio que permite ver los objetos de conocimiento en su contexto facilitando una posterior reutilización e interpretación.

La arquitectura propuesta tiene tres componentes principales que implementan los tres procesos necesarios para la Gestión del Conocimiento descriptos en la sección 2 del Capítulo VI.

3.1. Componente de Representación del Conocimiento

Este componente es responsable de la extracción y representación del conocimiento a partir de fuentes de conocimiento heterogéneas. Implementa una estrategia de clasificación automática basada en ontologías de las fuentes de conocimiento, superando el problema de sobrecarga de documentación (Ale y col., 2006a, 2006b). La adquisición y representación del conocimiento son cruciales para la implementación de un Sistema de Información de Memoria Organizacional como para cualquier otro sistema basado en conocimiento. En un caso ideal, un Sistema de Información de Memoria Organizacional sería auto-adaptativo y auto-organizado, recolectando dentro de las operaciones de negocios habituales los datos y la información relevantes para una reutilización posterior sin perturbar el flujo normal de trabajo de los empleados. Aunque esto sin duda minimizaría los costos de adquisición, no siempre es posible. Siempre es necesaria algún tipo de adquisición manual y en el caso del esquema de Memoria Organizacional Distribuida propuesta en esta tesis, esta tarea está a cargo de un ingeniero ontológico cuya función primordial es el desarrollo de las ontologías de dominio.

En relación a la representación, los documentos de texto, correos, gráficos, etc. están ampliamente disponibles y son muchos más confortables y expresivos para las personas que cualquier otra representación formal de conocimiento. Por otro lado, solamente nociones formalizadas permiten inferencias sofisticadas y recuperaciones precisas. Es por esta razón, y buscando una combinación significativa de lo formal y lo informal, que el enfoque adoptado en esta tesis es no modificar la representación de los objetos de conocimiento con un formalismo en particular, sino utilizar una representación homogénea que funcione como capa intermedia, transparente para el usuario y que facilite la búsqueda y el acceso a formas heterogéneas de conocimiento.

Como se mencionó anteriormente, el objetivo es enlazar, desde un punto de vista semántico, documentos organizacionales relevantes con una ontología de dominio que es el corazón de la Memoria Organizacional Distribuida. Más

específicamente, se propone una estrategia de anotado semántico donde la ontología de dominio se usa como la principal estructura de clasificación. La propuesta se basa en la hipótesis de que las ontologías de dominio contienen todos los conceptos y relaciones relevantes en el dominio. La necesidad de desarrollar una estrategia de anotado propia adecuada a la arquitectura de Memoria Organizacional propuesta se justifica en la sección 3 del Anexo E.

Para alcanzar este objetivo se desarrolló, como aporte de esta tesis, una estrategia de anotado que comprende los siguientes pasos:

- i.* Pre-procesado del texto (separación en “tokens”, etiquetación, etc.)
- ii.* Aplicación de diferentes técnicas para identificar la presencia de conceptos ontológicos en el texto (búsqueda de ocurrencias exactas, expansión semántica, inclusión de superclases)
- iii.* Persistencia (almacenamiento) de las relaciones entre los documentos y la ontología de dominio.

La estrategia propuesta para este componente se presenta en el Algoritmo VII.1 y se describe con detalle a continuación usando un ejemplo que involucra una versión extendida de la ontología Travel⁵ que contiene conceptos del área de turismo. Los detalles técnicos y de implementación se presentan en el Capítulo VIII.

⁵ Disponible, en su versión original, en <http://protege.stanford.edu/plugins/pwl/owl-library/index.html>

INICIO

Selección del Corpus a ser tratado
{ Etiquetado del Corpus (postagging)
Selección de Sustantivos } Proceso de Etiquetado

Proceso de Búsqueda de Ocurrencias Exactas

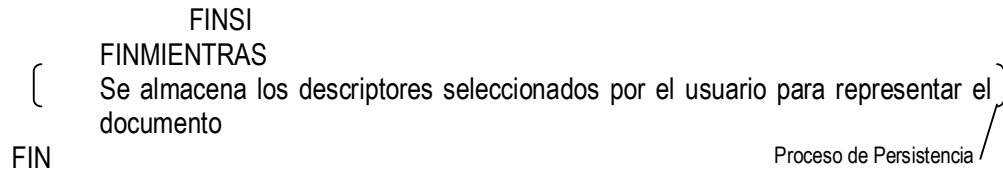
MIENTRAS existan sustantivos en la lista
SI (el sustantivo coincide con un concepto ontológico) ENTONCES
El sustantivo se marca como posible descriptor y se lo enlaza al concepto ontológico
Las Profundidades Ontológica y Semántica permanecen en cero /* ya que no fue necesaria una expansión semántica */
SINO
El sustantivo se pasa a una lista de descriptores no enlazados
FINSI
FINMIENTRAS

Proceso de Expansión Semántica

MIENTRAS existan descriptores no enlazados en la lista
Buscar el sustantivo no enlazado en un diccionario
SI (se encontró algún synset) /*synset=conjunto de sinónimos*/
ENTONCES
SI (el sinónimo coincide con algún concepto ontológico)
ENTONCES
El sustantivo se marca como posible descriptor y se lo enlaza al concepto ontológico
Se incrementa la Profundidad Semántica
SINO
SI (existe un hiperonimo) /*hiperonimo=concepto de jerarquía superior o padre*/ ENTONCES
Volver a búsqueda en el diccionario
SINO
Continuar con otro sustantivo no enlazado de la lista
FINSI
FINSI
SINO
Se descarta el sustantivo
FINSI
FINMIENTRAS

Proceso de Inclusión de Superclases

MIENTRAS existan conceptos ontológicos enlazados
SI (el concepto tiene un padre en la ontología) ENTONCES
Se marca el padre como posible descriptor
Se incrementa Profundidad Ontológica
/* El proceso se repite hasta llegar a los conceptos sin padre*/
SINO
Se continúa con otro concepto ontológico



Algoritmo VII.1. Lógica del Componente de Representación del Conocimiento

El proceso comienza con la *selección del corpus* (conjunto) de documentos a ser tratados. Cada documento de dicho corpus pasa por un proceso de *etiquetado* o “POSTagging” (Part-Of-Speech Tagging) en el cual se le asigna a cada “token” o elemento del documento una etiqueta con su naturaleza sintáctica (JJ-adjetivo, NN-Sustantivo singular común, VBZ-Verbo, tiempo presente, tercera persona singular, CC-Conjunción coordinante, IN-Preposición o conjunción subordinante, DT-Determinante, NNS-Sustantivo, plural común, CD-Número Cardinal, VBP-Verbo, tiempo presente, VB-Verbo, forma básica, NNP-Sustantivo singular propio, RB-Adverbio, etc.). En la Figura VII.3 puede observarse el resultado del proceso de etiquetado de una porción de texto.

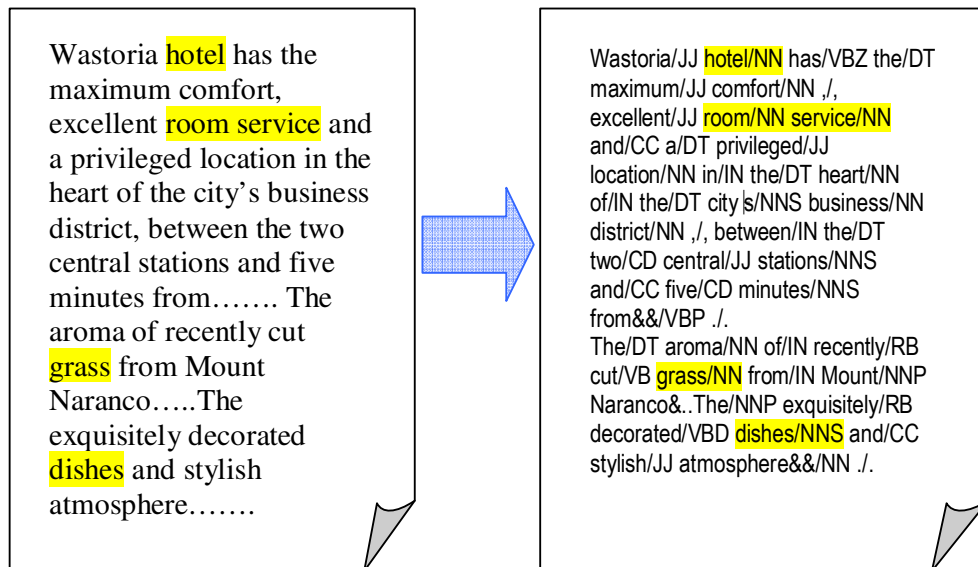


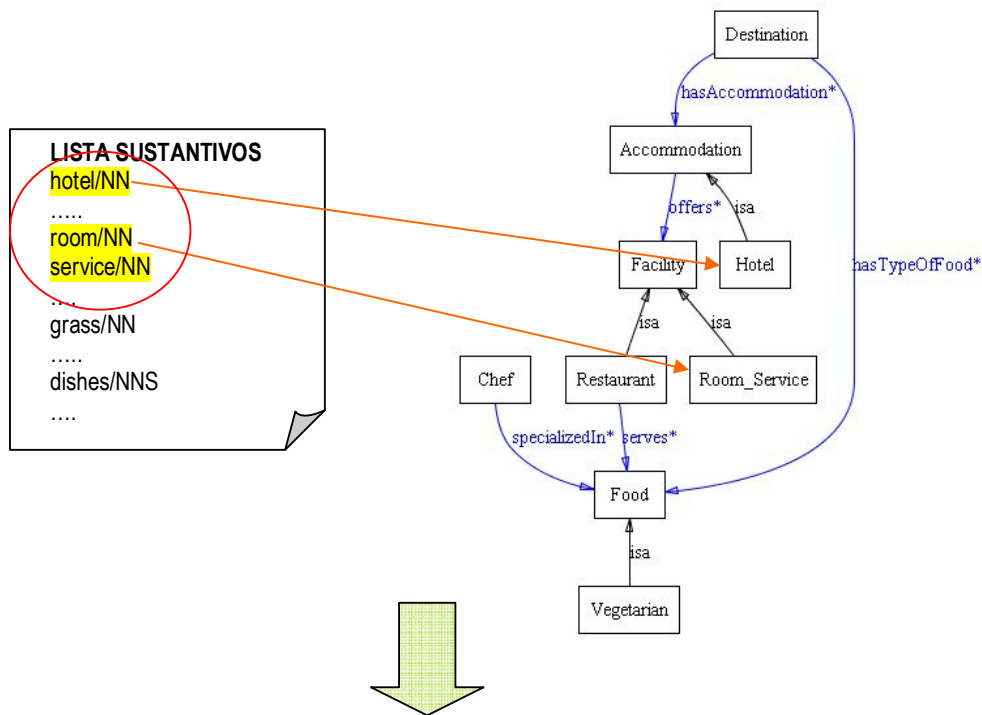
Figura VII.3. Proceso de Etiquetado (NN o NNS = Sustantivo)

Una vez finalizado el etiquetado se *seleccionan los sustantivos* para su posterior tratamiento. Usualmente la decisión de si una palabra en particular puede ser o no utilizada como término representativo del documento está relacionada a la naturaleza sintáctica de la palabra. De hecho, los sustantivos generalmente contienen más semántica que los adjetivos, adverbios y verbos (Baeza-Yates y Ribeiro-Neto, 1999). Además, una convención ampliamente aceptada es que los nombres de los conceptos ontológicos sean sustantivos y los nombres de las relaciones entre dichos conceptos sean verbos. Finalmente, como en nuestro caso, los términos representativos serán determinados en función de los conceptos ontológicos de cada una de las ontologías de dominio, la estrategia se enfoca en esta categoría dentro del texto etiquetado.

A partir de los sustantivos seleccionados del texto etiquetado, comienza un proceso de *búsqueda de ocurrencias exactas*. En dicho proceso se compara cada uno de los sustantivos con los conceptos de la ontología de dominio que representa el conocimiento considerado como representativo de cada dominio. Esta comparación también tiene en cuenta aquellos conceptos ontológicos que son sustantivos compuestos, es decir, que contienen más de una palabra. Si se encuentra una coincidencia, *se marca ese sustantivo como descriptor* del documento y se lo enlaza al concepto ontológico.

Los parámetros asociados a las profundidades de búsqueda tanto ontológica como semántica permanecen en cero ya que no fue necesario recorrer la ontología ni realizar expansiones semánticas. Si no se encuentra ninguna coincidencia para el sustantivo, el mismo se pasa a una *lista de descriptores no enlazados* para su posterior tratamiento.

Como se observa en la Figura VII.4 se han detectado dos conceptos ontológicos directamente *hotel* y *room service* siendo este último un sustantivo compuesto. El resto de los sustantivos, que no pudieron ser directamente relacionados con un concepto ontológico, pasan a integrar una lista de *descriptores no enlazados*.



| DESCRIPTORES DEL DOCUMENTO | PROFUNDIDAD SEMANTICA | PROFUNDIDAD ONTOLOGICA |
|----------------------------|-----------------------|------------------------|
| Room_Service | 0 | 0 |
| Hotel | 0 | 0 |
| DESCRIPTORES NO ENLAZADOS | | |
| Grass | | |
| Dishes | | |

Figura VII.4. Proceso de Búsqueda de Ocurrencias Exactas

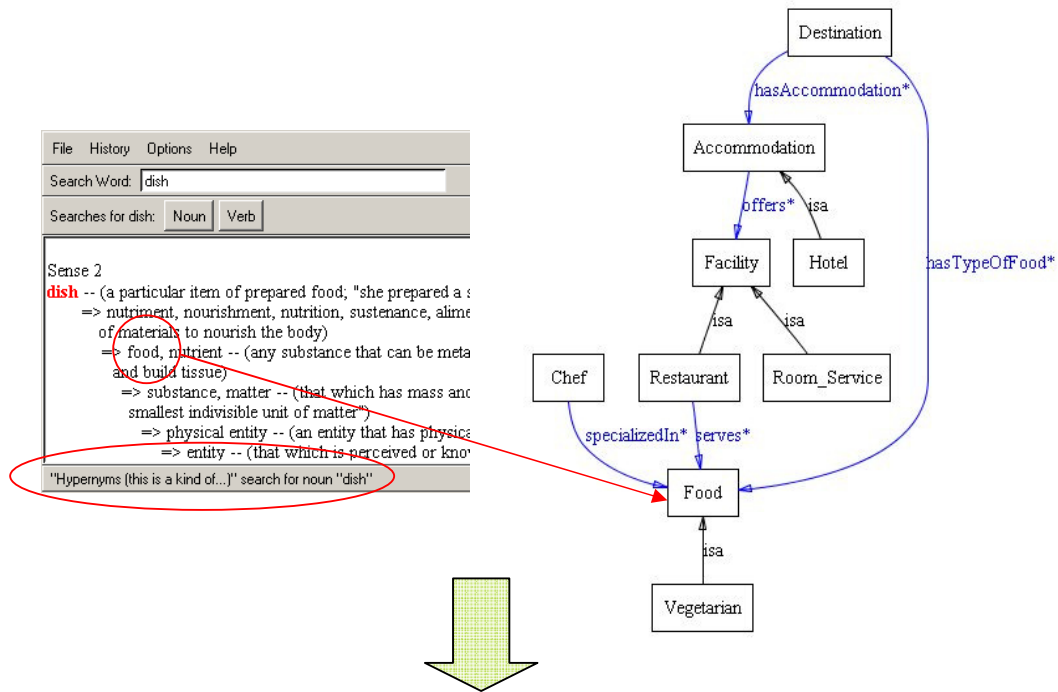
Con la lista de *descriptores no enlazados* se procede a un *proceso de expansión semántica* para ampliar el rango de coincidencias, mejorando de esta forma la representación del documento. Para cada sustantivo no enlazado se *busca su definición en un diccionario*. Si se encuentra algún conjunto de sinónimos que definen el sustantivo, se comparan dichos sinónimos con los conceptos de la ontología. Si se encuentra una coincidencia (entre uno de los sinónimos que define al sustantivo y un concepto ontológico) se marca al sustantivo como descriptor del documento, se lo enlaza al concepto ontológico

que mostró la coincidencia y se incrementa el parámetro de profundidad semántica ya que fue necesaria una expansión.

Los parámetros de profundidad semántica y ontológica serán de ayuda para el ingeniero ontológico al momento de tomar la decisión de cuáles descriptores serán definitivamente seleccionados para representar cada documento. Valores altos en estos parámetros, indicarían que fue necesario moverse a través de la estructura del diccionario (profundidad semántica) o de la ontología (profundidad ontológica) en numerosas ocasiones para encontrar una coincidencia degradando, de esta forma, la representatividad del sustantivo.

Si dentro del conjunto de sinónimos que define al sustantivo todavía no enlazado no se encuentran coincidencias con algún concepto ontológico, se *busca su hiperónimo* (concepto de jerarquía superior o padre, respuesta a la pregunta ¿qué es?) en el diccionario y se repite el proceso de búsqueda en el diccionario con el hiperónimo.

El proceso termina cuando se agotan las instancias de búsqueda y si no se encontró ninguna coincidencia, se descarta el sustantivo como posible descriptor. En la Figura VII.5 puede verse que al recorrerse la estructura de hiperónimos del diccionario (relación *es del tipo de...*) ingresando por la palabra *dish* se obtiene el sustantivo *food* el cual es un concepto ontológico. De esta forma, tanto *dishes* como *food*, pasan a ser posibles descriptores del documento.

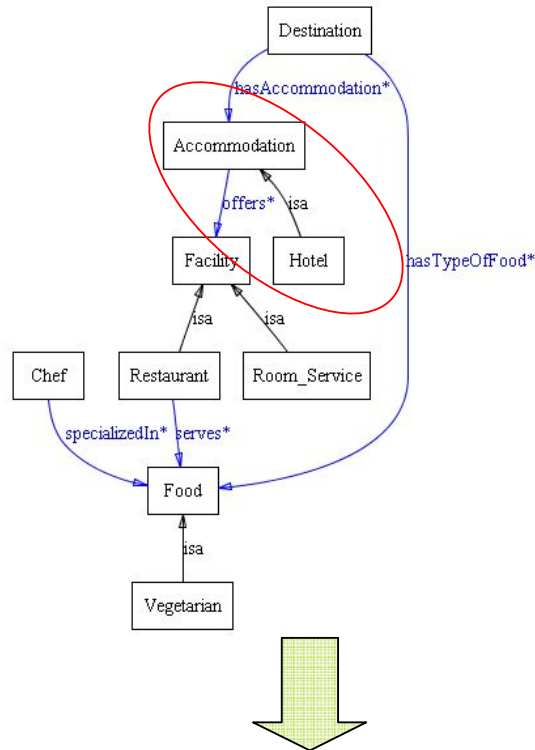


| DESCRITORES DEL DOCUMENTO | PROFUNDIDAD SEMANTICA | PROFUNDIDAD ONTOLOGICA |
|---------------------------|-----------------------|------------------------|
| Room_Service | 0 | 0 |
| Hotel | 0 | 0 |
| Food/Dishes | 1 | 0 |
| DESCRITORES NO ENLAZADOS | | |
| Grass | | |

Figura VII.5. Proceso de Expansión Semántica

Con el listado de descriptores enlazados se procede a un *proceso de inclusión de superclases*. En este proceso se navega por la estructura ontológica, hacia arriba, a partir de cada concepto ontológico enlazado previamente para recuperar los hiperónimos o padres, ampliando de esta forma, la cantidad de descriptores representativos del documento. Con cada nivel que se sube en la ontología se incrementa el parámetro de profundidad ontológica. El crecimiento de este parámetro indica la generalidad del concepto ontológico como descriptor, ya que a medida que se sube en la jerarquía ontológica, los conceptos son más generales y por ende, con menor potencialidad representativa. Como se observa

en la Figura VII.6 al finalizar este proceso se ha agregado un nuevo descriptor (*Accommodation*) aumentando la profundidad ontológica asociada.



| DESCRIPTORES DEL DOCUMENTO | PROFUNDIDAD SEMANTICA | PROFUNDIDAD ONTOLOGICA |
|----------------------------------|-----------------------|------------------------|
| Room_Service | 0 | 0 |
| Hotel | 0 | 0 |
| Food/Dishes | 1 | 0 |
| Accommodation | 0 | 1 |
| DESCRIPTORES NO ENLAZADOS | | |
| Grass | | |

Figura VII.6. Proceso de Inclusión de Superclases

Finalmente, se procede a realizar un *proceso de persistencia*, donde se almacena la relación existente entre el documento y los conceptos que el ingeniero ontológico en definitiva seleccionó como descriptores del documento.

3.2. Componente de Procesamiento y Recuperación de Información

Desde el punto de vista de la utilización del conocimiento, pueden mencionarse dos características principales que puede distinguir este modelo de Memoria Organizacional Distribuida de un sistema de información y conocimiento convencional:

- i.* el suministro activo y dependiente del contexto del conocimiento e información útil para el usuario en una situación dada y,
- ii.* la independencia de la aplicación de la base de conocimiento del Sistema de Información de Memoria Organizacional que debería ser útil para múltiples sistemas de soporte a tareas específicas

Por supuesto, el enfoque activo y la independencia de las tareas son de alguna manera objetivos conflictivos. Los enfoques realmente inteligentes que ofrecen soporte activo requieren tanto conocimiento de dominio y de las tareas que difícilmente pueden ser construidos de una forma independiente de la aplicación. Es por ello que este modelo concentra el conocimiento de dominio requerido para una efectiva representación y recuperación del conocimiento en una ontología de dominio que es transparente para los usuarios.

Concretamente, el componente de Procesamiento y Recuperación de Información es el responsable por el análisis de la consulta del usuario, la transformación de la consulta a un formato que permita su tratamiento y la recuperación de la información requerida.

La estrategia propuesta para este componente se presenta en el Algoritmo VII.2 y se describe con detalle a continuación. Los detalles técnicos y de implementación se presentan en el Capítulo VIII.

INICIO

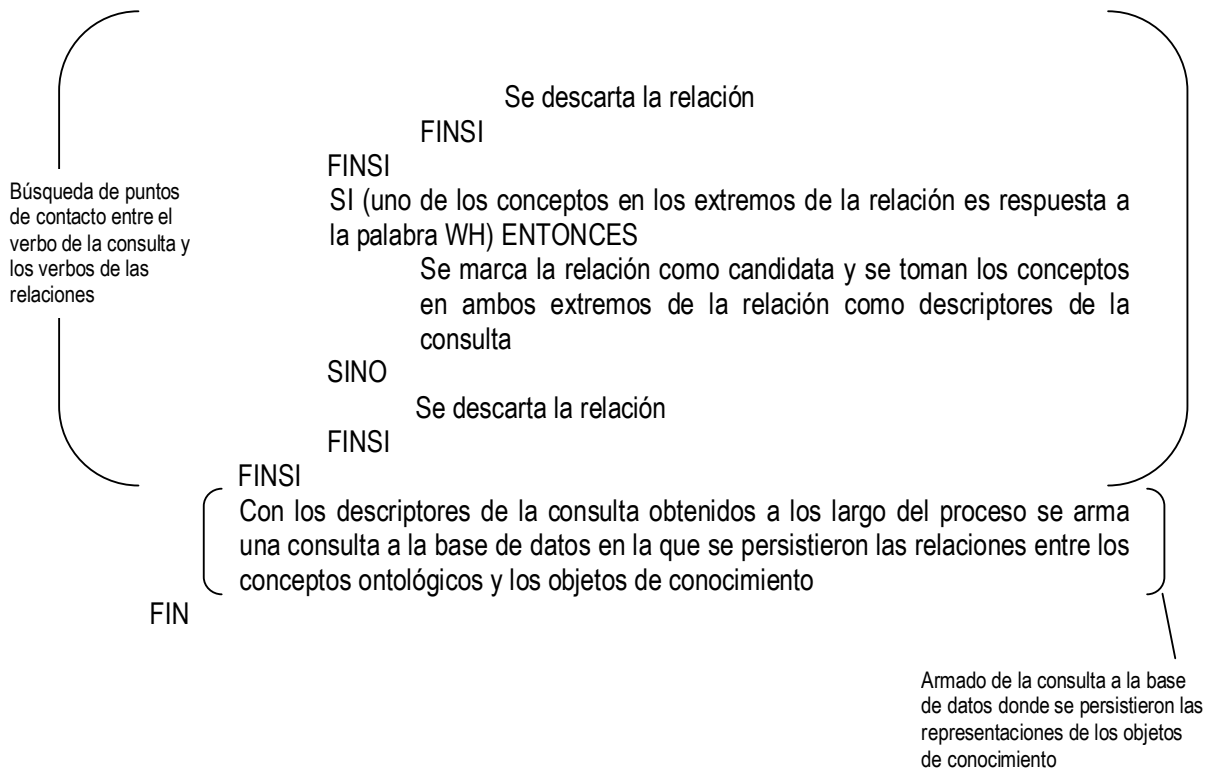
/* Dada la escasa información que proporciona el texto de una consulta en lenguaje natural debido a su brevedad, se trata de obtener más información a partir del uso de los verbos y las palabra pregunta (where, who, when, etc.) que contenga.*/
Recepción de la consulta en lenguaje natural
Etiquetado de la consulta (postagging)
Selección de sustantivos, verbos y palabra pregunta /*palabra pregunta=WH words, where, when, who, etc.*/
SI (el vocablo es un verbo) ENTONCES
Busca el verbo en un diccionario
Etiquetado de las definiciones de los verbos (postagging)
Selección de sustantivos
SINO SI (el vocablo es un sustantivo) ENTONCES
Se aplica la estrategia de anotado descripta anteriormente para los sustantivos /*VER CAPA DE REPRESENTACION DEL CONOCIMIENTO*/
Se recuperan las relaciones ontológicas asociadas a los conceptos ontológicos identificados a partir de los sustantivos de la consulta.
Se separan los verbos del nombre de las relaciones
Busca el verbo en un diccionario
Etiquetado de las definiciones de los verbos (postagging)
Selección de sustantivos
FINSI
FINSI
Se comparan los conjuntos de sustantivos que describen el verbo de la consulta y los de las relaciones
SI (existe una coincidencia) ENTONCES
Se verifica que el sustantivo esté asociado a la ontología /*esto descarta coincidencias triviales que no pertenezcan al dominio*/
Se desencadenan los procesos de Búsqueda de Ocurrencias Exactas y de Expansión Semántica ya descriptos para la Capa de Representación de Conocimiento.
SI (el sustantivo en común en las descripciones de los verbos puede ser relacionado semánticamente a la ontología) ENTONCES
Se marca la relación como candidata y se toman los conceptos en ambos extremos de la relación como descriptores de la consulta
SINO
SI (uno de los conceptos en los extremos de la relación es respuesta a la palabra WH) ENTONCES
Se marca la relación como candidata y se toman los conceptos en ambos extremos de la relación como descriptores de la consulta
SINO

Análisis Lingüístico de la Consulta

Análisis del verbo de la Consulta

Análisis del verbo de las relaciones ontológicas entre los descriptores de la Consulta

Búsqueda de puntos de contacto entre el verbo de la consulta y los verbos de las relaciones



Algoritmo VII.2. Lógica del Componente de Procesamiento y Recuperación de Información

La estrategia implementada en este componente tiene como objetivo el tratamiento de la consulta del usuario, expresada en lenguaje natural. Debido a la escasa información que proporciona el texto de una consulta en lenguaje natural debido mayormente a su brevedad, se propone obtener más información a partir del o los verbos que contenga y de las palabra-pregunta (o palabras *Wh – when, where, what, who*, etc.).

El proceso comienza con la recepción de la consulta en lenguaje natural, realizándose un *etiquetado (postagging)* similar al del componente anterior. En este caso, además de los *sustantivos*, se seleccionan para su posterior tratamiento los *verbos* y la *palabra-pregunta (Wh)*. Por ejemplo, en la Figura VII.7 puede verse el resultado del etiquetado de la pregunta “*Where can I eat vegetarian dishes?*”.

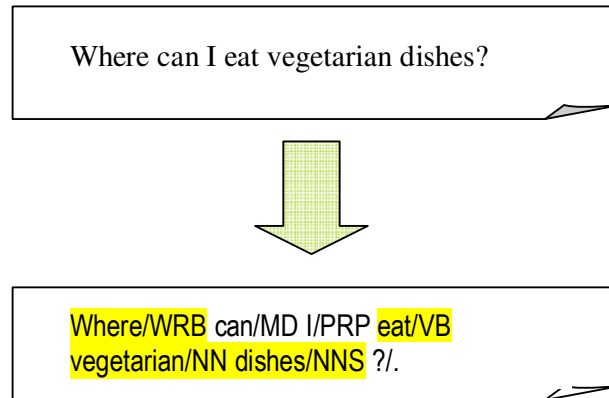


Figura VII.7. Proceso de Etiquetado de la Consulta (WRB = palabra pregunta; VB = verbo; NN o NNS = sustantivo)

Luego, se comienza con el tratamiento de los términos rescatados de la consulta. Si el término es un *verbo* se busca su significado en el diccionario. Se analizan las *definiciones* mediante un etiquetado (*postagging*) seleccionándose los *sustantivos* presentes en dichas definiciones. Este conjunto de *sustantivos asociados al verbo de la consulta* se reserva para su posterior uso.

Si por el contrario, el término es un *sustantivo*, se aplica la *estrategia de anotado* descrita en el componente anterior al final de la cual se obtienen los descriptores ontológicos asociados a la consulta.

A partir de estos descriptores ontológicos asociados a la consulta, se recuperan las *relaciones ontológicas* que existan entre estos conceptos separando los verbos en dichas relaciones. Para cada *verbo* se busca su significado en el diccionario. Se analizan las *definiciones* mediante un etiquetado (*postagging*) seleccionándose los *sustantivos* presentes en dichas definiciones. Este conjunto de *sustantivos asociados a los verbos de las relaciones* se reserva para su posterior uso.

En la Figura VII.8 puede observarse el resultado del proceso hasta el momento. Se han identificado dos conceptos: *vegetarian* y *food* y a través de las relaciones hay dos candidatos posibles: *restaurant* y *chef*.

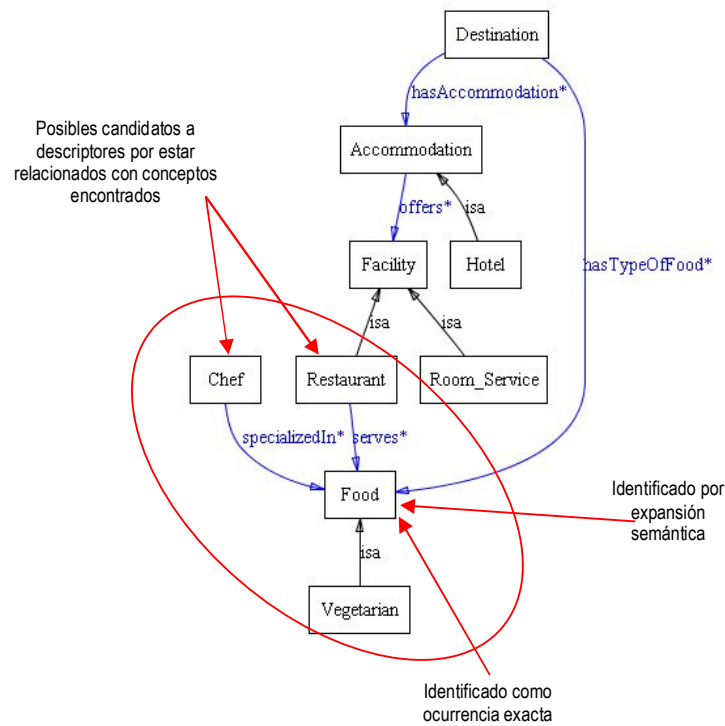


Figura VII.8. Posible Representación de la Consulta

El proceso desarrollado a continuación, que realiza la *búsqueda de puntos de contacto* entre el verbo de la consulta y los verbos de las relaciones ontológicas seleccionadas, tiene como objetivo ampliar la cantidad de descriptores de la consulta, y por ende, la comprensión de la misma, haciendo uso de las relaciones definidas en la ontología de dominio. Se *comparan* los conjuntos de sustantivos asociados al verbo de la consulta y a los verbos de las relaciones. Si existe una coincidencia, se verifica que el *sustantivo esté asociado a la ontología* de dominio (se desencadenan los procesos de Búsqueda Ocurrencias Exactas y Expansión Semántica ya descritos en el Componente de Representación de Conocimiento). Esta verificación tiene como objetivo descartar coincidencias triviales, es decir, aquellas que no sean de interés en el dominio de trabajo.

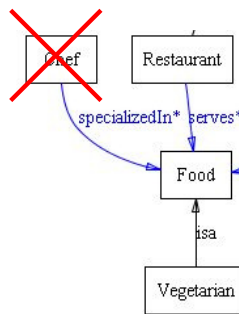
Siguiendo el ejemplo anterior, se recuperan los conceptos ontológicos identificados en la consulta junto con sus vecinos, *Restaurant* y *Chef*. Para decidir si alguno de los vecinos es útil para representar la consulta (y no buscar solamente por *Food* y *Vegetarian*) se evalúa la similitud entre el verbo de la consulta (*eat*) y los verbos de las relaciones asociadas a los conceptos identificados (*serve*, *specialize*) usando los conjuntos de sinónimos y términos relacionados del diccionario. En la Figura VII.9 puede verse un resumen de los significados asociados a cada verbo (en las relaciones y la consulta) y los puntos de contacto detectados. De acuerdo a este análisis, *serve* tiene una similitud semántica mayor con *eat* que *specialize*.

| eat | serve | specialize |
|---|--|--|
| <ul style="list-style-type: none"> - take in solid food - eat a meal - take in food - worry or cause anxiety in a persistent way - cause to deteriorate due to the action of water, air, or an acid - | <ul style="list-style-type: none"> - serve a purpose, role, or function - do duty or hold offices - help to some food - provide (usually but not necessarily food) - spend time in prison or in a labor camp - | <ul style="list-style-type: none"> - become more special - be specific about - suit to a special purpose - devote oneself to a special area of work - evolve so as to lead to a new species or develop in a way most suited to the environment - |

Figura VII.9. Puntos de Contacto entre los Verbos

Si el sustantivo en común puede ser relacionado con la ontología de dominio se marca la relación como candidata y se toman los *conceptos ontológicos en ambos extremos de la relación* como descriptores de la consulta. Ahora bien, si el sustantivo en común entre los asociados al verbo de la consulta y a los verbos de las relaciones no puede ser asociado a la ontología de dominio, se verifica si el concepto que todavía no es un descriptor de la consulta y que está en el otro extremo de la relación bajo análisis es *respuesta a la palabra-pregunta de la consulta*. Esto se realiza verificando su tipo, por ejemplo: si es de tipo *LOCATION* será respuesta a una pregunta del tipo *WHERE*, si es del tipo *PERSON* será respuesta a una pregunta del tipo *WHO*, etc. La tipología del concepto ontológico analizado será determinada mediante su búsqueda en el diccionario.

Si no hay puntos de contacto entre el verbo de la consulta y el verbo de la relación, o el concepto en el extremo de la relación no marcado previamente como descriptor de la consulta no es respuesta a la palabra-pregunta, la relación se descarta. Siguiendo con el ejemplo, y como se observa en la Figura VII.10, *Restaurant* es del tipo *LOCATION* (tipo esperado de acuerdo a la palabra-pregunta *WHERE*) y *Chef* es del tipo *PERSON* confirmando que la porción de la ontología de dominio que mejor representa la consulta contiene los conceptos: *Food*, *Vegetarian* y *Restaurant*.



Restaurant => Lugar (respuesta a Where)

Chef => Persona (respuesta a Who)

| DESCRIPTORES DE LA CONSULTA | METODO OBTENCION |
|-----------------------------|--------------------------------------|
| Vegetarian | Ocurrencia Directa |
| Food | Expansión Semántica de Dishes |
| Restaurant | Matching del verbo "eat" con "serve" |
| Restaurant | Tipo PLACE respuesta a Where |
| DESCRIPTORES NO ENLAZADOS | |
| Chef | Descartado |

Figura VII.10. Descriptores de la Consulta

Con los descriptores de la consulta obtenidos (a través de los sustantivos que posee la consulta y aquellos obtenidos mediante el análisis de los verbos de las relaciones ontológicas y las palabras-pregunta) *se arma una consulta* a la base de datos en la que se persistieron las relaciones entre los conceptos ontológicos y los documentos tratados.

3.3. Interfaz a Otros Dominios

Este componente hace posible el esquema distribuido y es el responsable de propagar la consulta del usuario hacia otros dominios que puedan proveer una respuesta. Para realizar esta tarea se dispone de dos estrategias posibles: propagación completa, la cual implica propagar la consulta a todos los dominios; o propagación selectiva, la cual implica propagar la consulta sólo a dominios con mayor posibilidad de responder la misma. Para esta última alternativa, se propone la estrategia que se presenta en el Algoritmo VII.3. Los detalles técnicos y de implementación se presentan en el Capítulo VIII.

INICIO

El usuario solicita propagación de la consulta hacia otros dominios

Con los descriptores de la consulta /*proceso descrito en el Componente de Recuperación*/ se recuperan de la Base de Casos las consultas con mayor similitud a la actual

Se presentan al usuario aquellos dominios que en ocasiones anteriores han respondido con éxito consultas similares a la actual

El usuario selecciona los dominios deseados y se propaga la consulta

Se arma una consulta a la base de datos en la que se persistieron las relaciones entre los conceptos ontológicos y los objetos de conocimiento de los dominios seleccionados

Se presentan los resultados al usuario solicitándole una valoración para almacenar los casos exitosos en la Base de Casos

FIN

Algoritmo VII.3. Lógica del Componente de Interfaz a Otros Dominios

La estrategia implementada mediante este algoritmo se desencadena ante el requerimiento del usuario de propagación selectiva de la consulta hacia otras Memorias Organizacionales definidas en la organización. Con los *descriptores asociados a la consulta* realizada por el usuario y que se obtuvieron de acuerdo al procedimiento descrito en el Componente de Recuperación, *se consulta la*

Base de Casos para recuperar aquellas consultas que tengan un grado de *similitud* con la actual.

Se presentan al usuario aquellos dominios que, en ocasiones anteriores, han respondido con éxito a consultas similares a la actual. En base a estos resultados, el usuario selecciona aquellos *dominios* a los que quiere propagar la consulta procediéndose a enviar *consultas a las bases de datos* de dichos dominios con los descriptores de la consulta.

Una vez obtenidas las respuestas de los dominios, se presentan los resultados al usuario solicitando su *valoración* para alimentar la Base de Casos con nuevos casos exitosos.

4. Conclusiones

En este capítulo se presentó, a partir de la justificación de la necesidad de un enfoque distribuido para la Gestión del Conocimiento, la arquitectura de un Sistema de Memoria Organizacional Distribuida, y se describieron con detalle las estrategias desarrolladas para la implementación de la funcionalidad de cada uno de los componentes de la arquitectura.

La arquitectura de tres componentes propuesta para una Memoria Organizacional Distribuida aborda dos problemas comunes en las implementaciones de repositorios de estas características: la sobrecarga de trabajo que implica para los trabajadores elicitar su conocimiento para contribuir al repositorio y la falta de contexto asociada al proceso de explicitar conocimiento para formalizarlo.

Para el componente de Representación del Conocimiento se propuso una estrategia de anotado automático basada en una ontología de dominio que realiza expansión semántica basada en un diccionario. Como resultado, obtiene un conjunto de descriptores del documento analizado que constituye la porción del modelo del dominio (ontología) que mejor representa el documento.

Para el componente de Procesamiento y Recuperación de Información se propuso un algoritmo que realiza un tratamiento de la consulta del usuario en lenguaje natural expandiendo la estrategia de anotado anterior con un análisis que aprovecha el poder representativo de las relaciones ontológicas junto con otras técnicas tendientes a comprender la consulta realizada.

Finalmente, en el componente de Interfaz a Otros Dominios, para el caso de requerirse una propagación selectiva, se propone una estrategia de aprendizaje basado en casos que permite propagar de manera selectiva la consulta del usuario a otros dominios.

En el capítulo siguiente se presenta el prototipo desarrollado en función de la arquitectura propuesta: Onto-DOM. Mediante la valoración del desempeño de los diferentes componentes de este sistema se podrá valorar el desempeño de las estrategias propuestas en este capítulo. Dichas valoraciones se presentarán en el Capítulo IX.

Onto – DOM

1. Introducción

El prototipo Onto-DOM (*Ontology-based Distributed Organizational Memory*) implementa la arquitectura propuesta en el Capítulo VII, con el propósito de contribuir al desarrollo de soluciones generales para el tratamiento de consultas en lenguaje natural sobre objetos de conocimiento con descripciones textuales. Acentuando el uso de ontologías, el prototipo provee una funcionalidad de representación, consulta y recuperación basada en contenido yendo más allá del reconocimiento superficial de palabras claves típica en los motores de búsqueda actuales, aunque no se intenta un análisis semántico completo de los textos fuente (Ale y col., 2007a; 2007b).

El objetivo global es desarrollar un Sistema de Información de Memoria Organizacional Distribuida para la representación del conocimiento de dominio, la semántica ontológica contenida en consultas en lenguaje natural, y para la búsqueda basada en textos en una Memoria Organizacional.

Según se describió en el Capítulo VII, el punto focal de cada Memoria Organizacional es una ontología de dominio lo cual trae aparejado la necesidad de gestionar dichas ontologías y la selección de un lenguaje de definición. En el Anexo E se presenta una revisión de los diferentes entornos de desarrollo de

ontologías y los lenguajes más utilizados junto con algunas herramientas de anotado.

En la sección 2 se presenta el prototipo Onto-DOM y en la sección 3 se presentan las conclusiones del presente capítulo.

2. Arquitectura Onto-DOM

El prototipo OntoDOM utiliza un modelo cliente servidor. En cada servidor existe el mismo conjunto de capas y el servicio es prestado utilizando un *Application Server* tal como Sun Application Server o JBoss. Se detalla a continuación la estructura de capas que componen la arquitectura y cuyo esquema puede observarse en la Figura VIII.1. (Ale y col. 2008a; 2008b; 2008c).

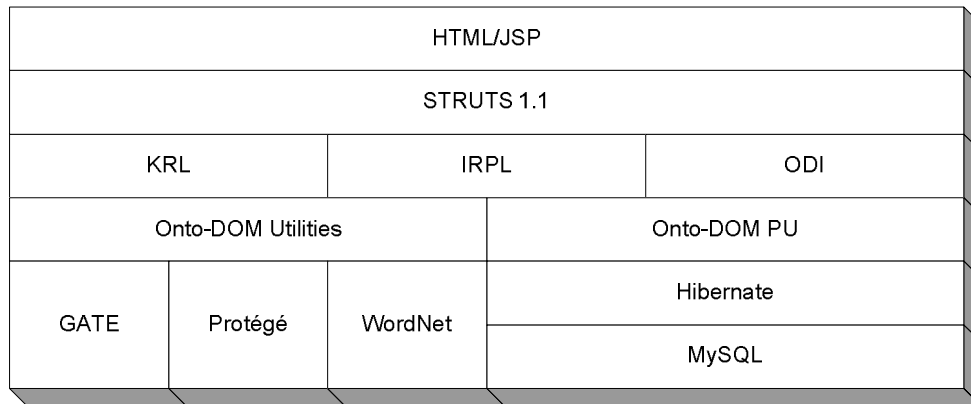


Figura VIII.1. Arquitectura Onto-DOM

La capa más externa, o de presentación de información, está compuesta de simples páginas html o jsp (Java Server Pages). En la siguiente capa se observa el framework *Struts* (Struts, 2000). Todas las peticiones http (o request) son gestionadas utilizando el mapeo de acciones de este framework. Struts es una herramienta de soporte para el desarrollo de aplicaciones Web bajo el patrón Modelo-Vista-Controlador bajo la plataforma J2EE (Java 2, Enterprise Edition).

Struts se desarrollaba como parte del proyecto Jakarta de la Apache Software Foundation, pero actualmente es un proyecto independiente conocido como Apache Struts.

Struts permite reducir el tiempo de desarrollo y su carácter de "software libre" y su compatibilidad con todas las plataformas Java Enterprise lo convierte en una herramienta altamente disponible.

Como se mencionó anteriormente, Struts se basa en el patrón del Modelo-Vista-Controlador (MVC) el cual se utiliza ampliamente y es considerado de gran solidez. De acuerdo con este patrón, el procesamiento se separa en tres secciones diferenciadas, llamadas el modelo, las vistas y el controlador. Logísticamente, separa claramente el desarrollo de interfaz del flujo de trabajo y lógica de negocio permitiendo desarrollar ambas en paralelo o con personal especializado. También es evidente que potencia la reutilización, soporte de múltiples interfaces de usuario (Html, sHtml, Wml, Desktop applications, etc.) y de múltiples idiomas, localismos, etc.

La lógica de la arquitectura se localiza en las capas KRL (Knowledge Representation Layer), IRPL (Information and Retrieval Processing Layer) y ODI (Other Domain Interface) cuya lógica fue descrita en el Capítulo VII y cuya implementación se describe con detalle en las siguientes secciones. La capa *OntoDOM Utilities* es un conjunto de utilidades cuyas funcionalidades son requeridas para completar los procesos de anotado, expansión semántica y contextualización. Esta capa abstrae al resto de las herramientas utilizadas tanto para el tratamiento del lenguaje natural como para la gestión de la ontología de dominio: Gate, Protégé y WordNet. La elección de Protégé para la gestión de la ontología de dominio se justifica en la sección 1 del Anexo E.

Por otro lado, se implementan una serie de capas tendientes al almacenamiento del proceso de anotado de los objetos de conocimiento. *OntoDOM PU (Persistente Unit)* es la capa de persistencia y en ella se encuentran el conjunto de clases que permite almacenar los resultados del anotado y procesamiento de información. Por debajo de ella se encuentra

Hibernate, que es el motor utilizado para asegurar la persistencia de los objetos de la capa de persistencia. Hibernate es una herramienta de mapeo objeto-relacional para la plataforma Java que facilita el mapeo de atributos entre una base de datos relacional tradicional y el modelo de objetos de una aplicación, mediante archivos declarativos (XML) que permiten establecer estas relaciones.

Hibernate permite el desarrollo de clases persistentes siguiendo un lenguaje orientado a objetos incluyendo asociación, herencia, polimorfismo, composición y colecciones. Además permite expresar consultas en una extensión propia y portable de SQL (HQL), como así también en SQL nativo. Hibernate es software libre, distribuido bajo los términos de la licencia GNU LGPL.

Finalmente, MySQL es el motor de base de datos escogido para el almacenamiento de los objetos. Las razones para su elección están relacionadas con sus múltiples ventajas. MySQL es un sistema de administración de base de datos relacional y multihilo con soporte para múltiples usuarios. Usa el lenguaje SQL estandarizado para el almacenamiento, actualización y acceso a la información. MySQL es muy rápido y capaz de almacenar grandes cantidades de datos, dando soporte además, a muchos lenguajes de programación distintos como: C, C++, Eiffel, Java, Perl, PHP, Python y TCL. Es escalable, ofreciendo la posibilidad de manipular bases de datos enormes, del orden de seis mil tablas y alrededor de cincuenta millones de registros, y hasta 32 índices por tabla. Presenta una alta conectividad, es decir, permite conexiones entre diferentes máquinas con distintos sistemas operativos. Es corriente que servidores Linux o Unix, usando MySQL, sirvan datos para ordenadores con Windows, Linux, Solaris, etc. Para ello se usa TCP/IP, tuberías, o sockets Unix. Finalmente, MySQL es un software libre distribuido bajo un esquema de licenciamiento dual GNU GPL.

La arquitectura OntoDOM define un modelo de dominio para la persistencia de la información que es resultado de sus procesos y que se observa en la Figura VIII.2. De los atributos observables, sólo *DocImpl::unlinkedDescriptors* y *ODQueryImpl::unlinkedDescriptors* no son persistidos. El resultado es un conjunto de tablas equivalente al modelo de clases observado ya

que el mapeo de las mismas se hace, como se mencionó anteriormente, con las capacidades básicas de Hibernate (Figura VIII.3).

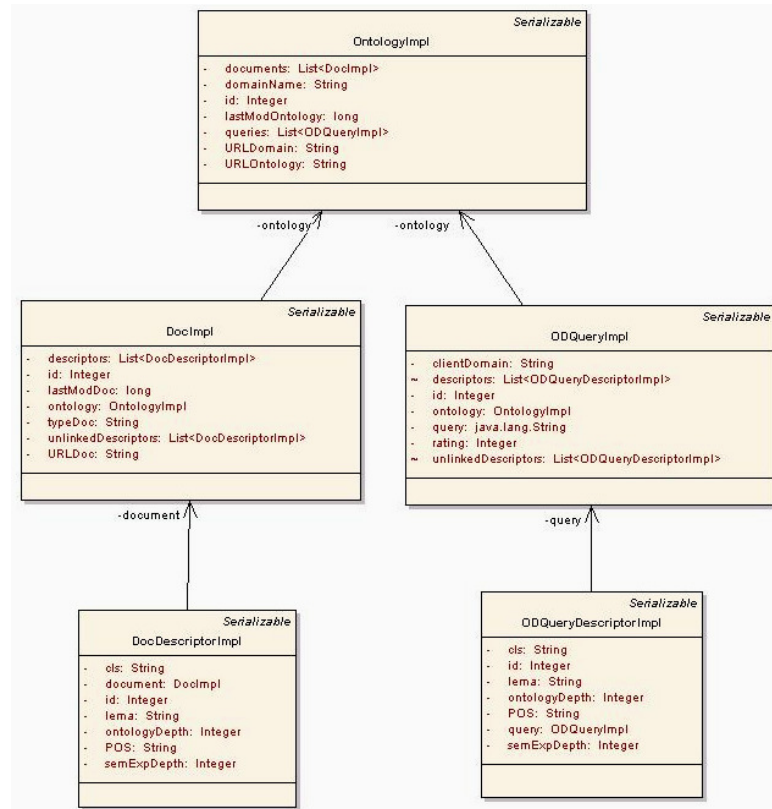


Figura VIII.2. Modelo de Persistencia

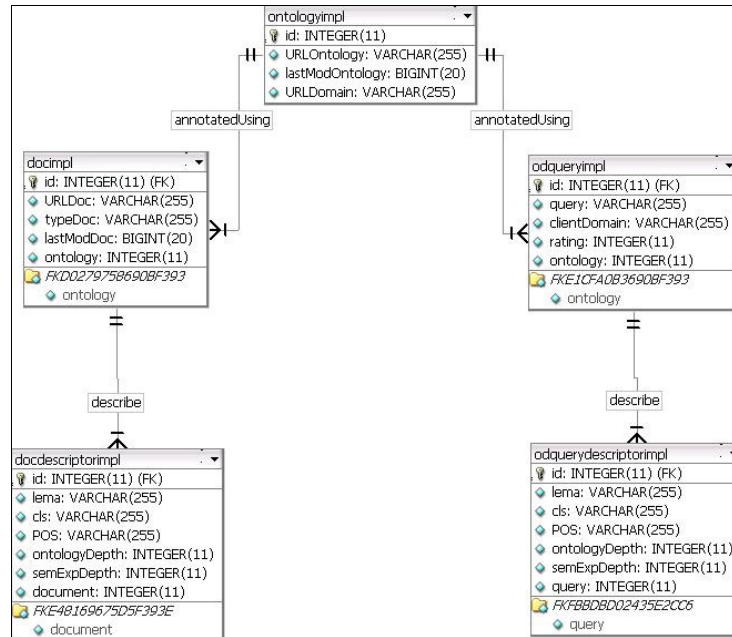


Figura VIII.3. Diseño de Tablas

2.1. Proceso de Inicialización de Onto-DOM

El proceso de inicialización de la herramienta Onto-DOM se describe haciendo referencia a la Figura VIII.4 que presenta el diagrama de colaboración respectivo. Dicho proceso comienza con la creación del dominio (1), un dominio es una instancia de la clase *ontodom.Domain*, cada instancia de esta clase contiene una instancia de las capas KRL, IRPL, y ODI. En otras palabras, se genera una Memoria Organizacional para cada dominio de conocimiento definido. La resolución, tanto del anotado de documentos como del procesamiento y recuperación de los mismos, requiere de la especificación del dominio sobre el que se debe trabajar. Esta instanciación de dominio toma valores definidos en el servidor para inicializar cada herramienta, pueden mencionarse como los más relevantes: la ubicación (URI) de la ontología (archivo OWL), el directorio donde se localiza el ANNIE plug-in, y el directorio de los diccionarios WordNet. La elección de OWL como lenguaje de definición de ontología se justifica en la sección 2 del Anexo E.

En la instanciación de la herramienta de etiquetado, el paso a seguir es la creación de la instancia que resuelve el procesamiento de la capa KRL (1.1). Sus parámetros más relevantes son la dirección (URL) del dominio, por ejemplo: “http://www.travel.com” y la ontología de dominio con la cual se instanciará el *OWLModel*.

Onto-DOM contiene un paquete por cada herramienta a utilizar (Gate, Protégé, WordNet) (1.2). En este paso se instancia el modelo (*ontodom.protege.Model*) y esta clase es la única clase de la herramienta que depende de las librerías de Protégé. En el siguiente paso se parametriza el modelo con la ontología del dominio (1.3) y debido a que la instancia de *Model* inicializada con la dirección (URL) de la ontología requiere del modelo OWL, en este paso se solicita a *ProtegeOWL* que devuelva una instancia de *OWLModel* respectiva a la ontología dada (1.4).

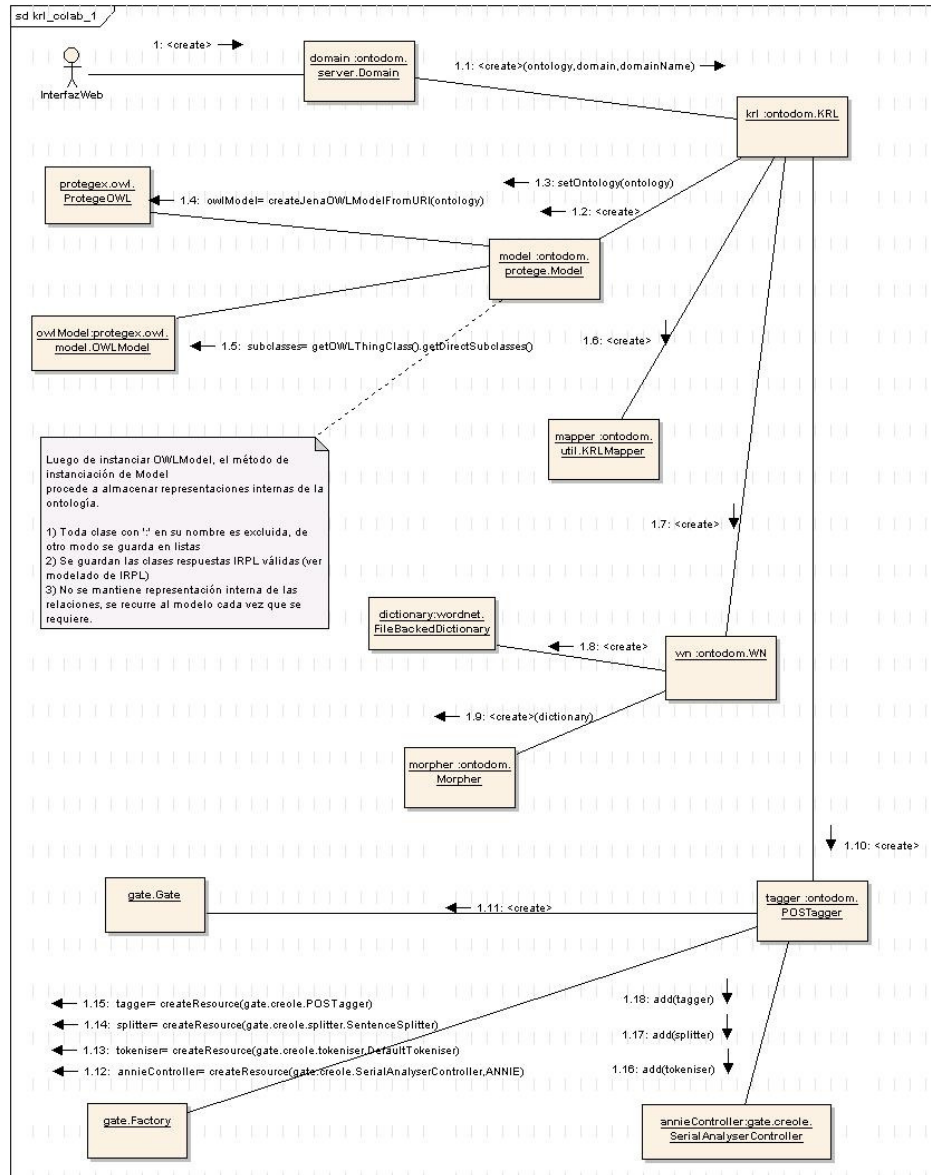


Figura VIII.4. Proceso de Inicialización de la Herramienta Onto-DOM

Para obtener las clases del modelo se pide al *OWLModel* la raíz de la ontología (*Thing*), y a ésta se le solicita el conjunto completo de subclases (1.5). Se procede luego a la creación de la herramienta de mapeo. Dado las diferentes representaciones que ofrecen las tres herramientas (Gate, WordNet y Protégé), es necesaria una clase cuyas instancias puedan realizar comparaciones sin obligar al

resto a conocer las particularidades de las herramientas. Esta función de comparación se vincula directamente con el etiquetado de los documentos (1.6).

Se producen luego una serie de instancias relacionadas con el vínculo de las capas de Onto-DOM con WordNet. En primer lugar se realiza la instanciación del manejador de WordNet, *ontodom.WNImpl*. Esta clase es responsable de la ejecución de las peticiones a los componentes de la librería JWordNet (1.7). Luego se instancia un *wordnet.DictionaryDatabase* para la recuperación de los objetos desde los archivos del diccionario de WordNet (1.8); y finalmente se crea una instancia *ontodom.Morpher*, responsable de los métodos de tratamiento de la morfología de las palabras según lo especifica WordNet (1.9). Esto es necesario ya que WordNet sólo guarda las formas base de las palabras y el procesamiento tiene dos métodos: el manejo de inflexiones y una lista de excepciones.

Comienza luego la inicialización de la relación de Onto-DOM con GATE. En primer lugar se produce la creación de la herramienta de etiquetado, su rol es utilizar las herramientas de GATE (ANNIE) para realizar el etiquetado del documento (1.10). Luego se realiza la inicialización de la librería GATE, paso requerido para la utilización de los recursos de procesamiento de esa librería (1.11). Se crea luego una instancia de ANNIE, un *gate.creole.SerialAnalyserController* responsable de la ejecución del resto de los recursos de procesamiento (1.12).

Se inicializan luego los componentes relacionados a la separación del texto en oraciones y palabras previo a su etiquetación. Se crea el “tokenizer”, *gate.creole.DefaultTokenizer* y, en el caso particular de Onto-DOM el “english tokenizer” responsable de la interpretación de un conjunto de caracteres como un solo “token” a partir de reglas (1.13). Luego se crea un *gate.creole.SentenceSplitter* responsable de la segmentación de cada línea para su posterior etiquetado (1.14); y finalmente un recurso de etiquetado, *gate.creole.POSTagger*, que realiza el ANNIE Part-of-Speech Tagging (1.15).

En la Figura VIII.5 se observa la pantalla de inicio del prototipo Onto-DOM luego de cumplirse todo el proceso de inicialización. A partir de la misma puede accederse al Componente de Representación y al de Recuperación de Información.

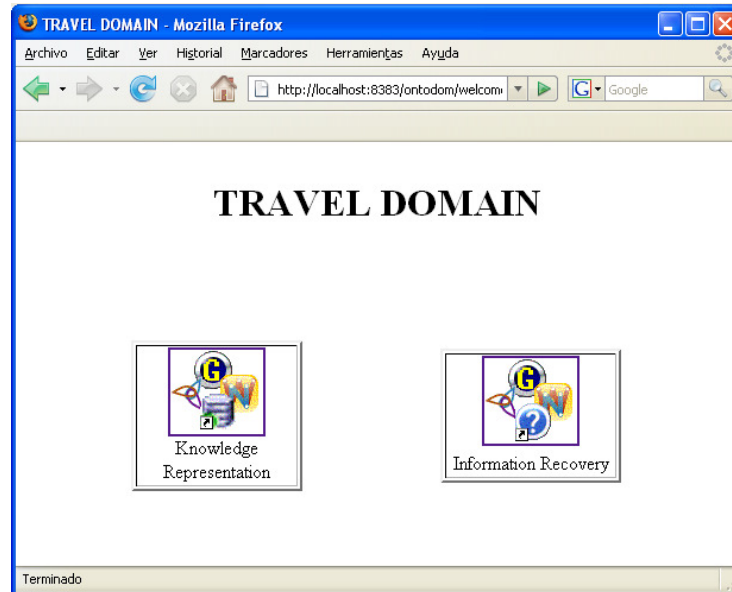


Figura VIII.5. Pantalla de Inicio de Onto-DOM

2.2. Componente de Representación de Conocimiento (KRL)

Cómo se mencionó anteriormente, este módulo tiene el propósito de crear una capa homogénea de conocimiento a partir de las fuentes heterogéneas que posee la organización. Esta adquisición de conocimiento se realiza en forma automática a través de una estrategia de anotado semántico que utiliza la ontología de dominio como estructura principal de clasificación. Los detalles técnicos de implementación se encuentran en el Anexo A.

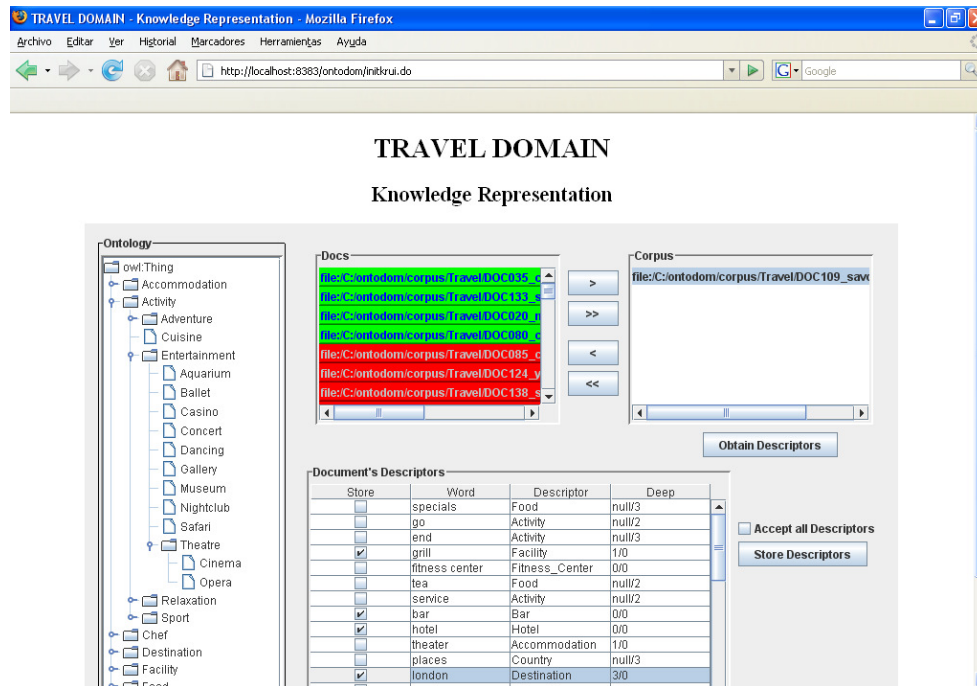


Figura VIII.6. Interfaz KRL

En la Figura VIII.6 puede observarse la interfaz del Componente KRL donde el ingeniero ontológico realiza el anotado de documentos. En la zona izquierda puede verse una representación gráfica de la ontología de dominio que se está utilizando para el proceso de anotado. En la parte superior se muestra el listado de documentos disponibles para anotar entre los cuales el ingeniero ontológico puede optar.

El botón Obtener Descriptores (*Obtain Descriptors*) desencadena el proceso de representación del conocimiento que se describió con detalle en la sección 3.1 del Capítulo VII. El proceso de representación de conocimiento implementado en este módulo puede resumirse en cinco pasos principales: etiquetado, búsqueda de ocurrencias exactas, expansión semántica, inclusión de superclases y persistencia.

2.2.1. Proceso de Etiquetado

El primer paso consiste en la obtención de los sustantivos de los documentos a ser tratados. Para ello se añade el documento al “corpus” que será tratado por el “plugin” Annie de GATE. Dicho “plugin” ejecuta una serie de recursos de procesamiento devolviendo un conjunto de etiquetas sintácticas asociadas a las palabras del documento tipificadas mediante las siglas NN (sustantivos), NNP (sustantivos propios), VB (verbos), etc.

Con las palabras etiquetadas con NN (sustantivos) se crea una lista de *descriptores no enlazados*, esto quiere decir que aún no se ha encontrado ninguna vinculación con los conceptos ontológicos representativos del dominio. En este punto también se tiene en cuenta el tratamiento de sustantivos compuestos (que tienen más de una palabra) al comparar los sustantivos con aquellas clases cuyo nombre tiene más de un término, por ejemplo: *Room_Service*.

2.2.2. Proceso de Búsqueda de Ocurrencias Exactas

Luego del etiquetado sintáctico, el proceso de anotado continua con la búsqueda de las ocurrencias exactas de clases ontológicas dentro de los sustantivos identificados. Por cada sustantivo de la lista se realiza una comparación con el conjunto de clases formateado a tal efecto.

Si se encuentra una coincidencia, el sustantivo es marcado como un posible descriptor del documento asociándolo a la clase ontológica y poniendo en cero todas las profundidades ya que no fue necesario realizar ninguna expansión semántica.

2.2.3. Proceso de Expansión Semántica

Agotada la búsqueda de ocurrencias exactas, el proceso de anotado continua utilizando WordNet como herramienta de expansión semántica. Para cada uno de los descriptores aún no enlazados se realiza una búsqueda en WordNet. Si la búsqueda directa no arroja resultados, la misma se repite usando una forma básica de la palabra en función de las reglas morfológicas y excepciones de WordNet.

Por cada palabra WordNet devuelve un conjunto de sentidos o “synset”. Cada “synset” es un conjunto de sinónimos que define un concepto y que tiene punteros semánticos a otros “synsets” que constituyen diversos tipos de relaciones (hiperonimia, hiponimia, antonimia, etc.).

Para cada sinónimo obtenido se realiza una búsqueda entre los conceptos ontológicos. Si el resultado es positivo, se guarda el sustantivo que dio origen al sinónimo como descriptor, se lo enlaza al concepto ontológico correspondiente y se ponen las profundidades en cero ya que la coincidencia fue encontrada en función de una relación de sinonimia y no se realizaron desplazamientos por la jerarquía ontológica o de WordNet.

Si al procesar un “synset” no se logran resultados positivos, se buscan más términos ontológicos a través de los punteros de *hiperonimia e instancia de hiperonimia* definidos en la instancia del “synset” en proceso subiendo por las relaciones de hiperonimia de la jerarquía dentro del mismo sentido. El tratamiento en cada estadio de la jerarquía es similar, con la única diferencia que se va incrementando el valor de la profundidad en cada lectura de los punteros de hiperonimia. Este proceso de “escalado” continua hasta agotar la información en WordNet para las relaciones halladas a partir del “synset” original y sus padres.

2.2.4. Proceso de Inclusión de Superclases

Debido a la necesidad de construir una base de datos que permita acceder a los documentos a partir de la estructura ontológica se busca maximizar la cantidad de descriptores de cada documento. Para ello se incluyen no sólo las clases halladas, sino también las superclases de las mismas y es en este paso donde se recuperarán las superclases de cada clase hallada. Cada nivel de la relación 'isA' incrementa un contador (*depth*) que se usa para dar información de la profundidad de la relación respecto al término hallado. Esto ofrece un dato importante para el responsable del anotado que le permite decidir si las superclases se añaden a la base de conocimientos como descriptores de un documento en particular en función a cuán lejos están de acuerdo a lo que indica esta profundidad del término original.

Básicamente, en este proceso se realiza la inclusión como descriptores del documento a las superclases de las clases halladas hasta el momento. Se recupera el conjunto de descriptores vinculados al dominio y se comienza a iterar sobre dicho conjunto añadiendo descriptores a medida que se “sube” por la jerarquía ontológica. Finalizado el proceso, los nuevos descriptores son agregados al conjunto de descriptores vinculados al dominio.

2.2.5. Persistencia

En este paso se almacenan en base de datos las referencias a los documentos y los descriptores seleccionados por el usuario. La lista de descriptores a almacenar es asignada a los descriptores del documento y se utiliza el motor de Hibernate para persistir la información.

En la Figura VIII.6 puede observarse el resultado del proceso de anotado sobre un documento del dominio Travel (*DOC109_savoy.txt*). En la sección *Document's Descriptors* pueden verse los sustantivos encontrados en el

documento y los descriptores asociados a los mismos, junto con el dato de la profundidad de búsqueda ontológica y de expansión semántica. En base a esta información el ingeniero ontológica a cargo del proceso de anotado puede decidir qué conjunto de descriptores se asociarán al documento.

2.3. Componente de Recuperación de Información (IRPL)

Este módulo implementa la estrategia de resolución de una consulta en lenguaje natural, que permite la recuperación del objeto de conocimiento buscado, descrita en la sección 3.2 del Capítulo VII. Los detalles técnicos de implementación se encuentran en el Anexo B.

Como puede observarse en la Figura VIII.7 el proceso comienza con una petición del usuario, en lenguaje natural, que es transportada por la interfaz web.



Figura VIII.7. Interfaz de la capa IRPL

A dicha consulta se le realiza un análisis lingüístico en el cual (mediante el procesamiento con el “plugin” Annie de GATE) se le asigna a cada palabra su POS (Part of Speech) correspondiente.

En esta oportunidad se conservan aquellas palabras cuyo POS indique que son sustantivos (NN), verbos (VB) o palabras pregunta del tipo dónde,

cuándo, quién, etc. (Wh) y con ellos se construye una lista de descriptores no enlazados.

La estrategia presentada determina tanto la similitud en función de los conceptos ontológicos como también utilizando las relaciones entre dichos conceptos. El tratamiento de los sustantivos es similar al realizado con los sustantivos del proceso de anotado de documentos, en el cual se tratan de identificar conceptos ontológicos, pero esta vez en el texto de la consulta.

El objetivo en este punto es ir más allá de las relaciones taxonómicas (es-un) haciendo uso de las relaciones semánticas de la ontología para mejorar la comprensión de la consulta. En esencia, se trata de entender la consulta basándose en el conocimiento codificado en la ontología de dominio, los recursos léxicos como WordNet y GATE y las heurísticas asociadas con el tratamiento de las palabras pregunta (Wh, where, when, who, etc.)

Sin embargo, como se mencionó antes, el objetivo principal era ir más allá de una búsqueda por palabras clave o el uso de la ontología de dominio como herramienta de expansión semántica. En este sentido y en primer lugar, se usarán los verbos detectados en la consulta para buscar similitud semántica con las relaciones definidas en la ontología. Por otro lado, se analizarán los conceptos vinculados por esas relaciones para determinar si son del tipo esperado como respuesta de acuerdo a la palabra pregunta (Wh).

Una vez obtenidos los descriptores de la consulta, se arma con ellos una consulta SQL a la base de datos del dominio, donde se persistieron las relaciones entre los documentos y sus respectivos descriptores. El resultado presentado al usuario puede verse en la Figura VIII.8.

TRAVEL DOMAIN

Information Recovery and Processing - Query Answer

| | | |
|--------|---|--|
| Query: | <input type="text" value="Where can I eat vegetarian dishes near an acropolis?"/> | |
| | <input type="button" value="Process"/> | <input type="button" value="Propagate Query"/> |

| Words | URLs |
|--|---|
| [Food, Vegetarian, restaurant, vegetarian] | file:/D:/ontodom/corpus/Travel/Athens-685.html |
| [Food, Vegetarian, vegetarian] | file:/D:/ontodom/corpus/Travel/DOC146_namale_fiji.txt |
| | file:/D:/ontodom/corpus/Travel/Athens_mini_guide-Eating.htm |
| | file:/D:/ontodom/corpus/Travel/DOC149_crystal_serenity_spa.txt |
| [Food, food, restaurant] | file:/D:/ontodom/corpus/Travel/DOC116_athenaeum_hotel_london.txt |
| | file:/D:/ontodom/corpus/Travel/DOC086_cowley_manor.txt |
| | file:/D:/ontodom/corpus/Travel/DOC105_charlotte_street_hotel_london.txt |
| [Food, Vegetarian] | file:/D:/ontodom/corpus/Travel/Restaurants-Athens-Eden_Restaurant-BR-1.html |
| [Food, dishes] | file:/D:/ontodom/corpus/Travel/DOC084_monte_carlo_grand_hotel.txt |
| [Food, restaurant] | file:/D:/ontodom/corpus/Travel/DOC108_the_connaught.txt |
| | file:/D:/ontodom/corpus/Travel/DOC038_half_moon_cay.txt |
| | file:/D:/ontodom/corpus/Travel/DOC114_baglionihotellondon.txt |
| | file:/D:/ontodom/corpus/Travel/DOC115_metropolitan_hotel_london.txt |
| | file:/D:/ontodom/corpus/Travel/DOC074_das_triest_austria.txt |
| | file:/D:/ontodom/corpus/Travel/DOC051_beach_vacation_in_makena.txt |
| | file:/D:/ontodom/corpus/Travel/DOC113_claridges_london.txt |
| | file:/D:/ontodom/corpus/Travel/DOC087_le_manoir_aux_quat_saisons.txt |
| | file:/D:/ontodom/corpus/Travel/DOC101_adare_manor.txt |
| | file:/D:/ontodom/corpus/Travel/DOC092_hotel_joseph.txt |
| | file:/D:/ontodom/corpus/Travel/DOC023_fall_oakland_house.txt |
| | file:/D:/ontodom/corpus/Travel/DOC097_mount_juliet.txt |

Figura VIII.8. Módulo de Procesamiento y Recuperación de la Información

2.4. Componente de Propagación a otros Dominios (ODI)

Una vez realizada la recuperación de información en el dominio local, el sistema permite, mediante el componente de propagación a otros dominios (ODI), realizar la consulta en otros dominios de la organización. El usuario puede propagar la consulta utilizando el botón Propagate Query que se observa en la Figura VIII.8. Dispone de dos estrategias posibles, propagación completa o propagación selectiva, las cuales se describen en las secciones siguientes.

2.4.1 Propagación Completa

La propagación completa implica propagar la consulta a todos los dominios que componen el sistema distribuido. Esta estrategia de propagación no requiere de un algoritmo particular y su resultado depende exclusivamente de la estrategia de procesamiento de la consulta y recuperación de información que implementa el componente de recuperación de cada dominio. En el caso de esta tesis, del Algoritmo VII.2 presentado en el Capítulo VII.

El beneficio de esta estrategia es que se tiene mayor posibilidad de obtener una mejor respuesta a la consulta debido a la expansión a todos los dominios. Una desventaja es que, en caso de haber muchos dominios con muchos documentos anotados, la búsqueda puede requerir más tiempo del deseado por el usuario.

2.4.2 Propagación Selectiva

La propagación selectiva implica propagar la consulta sólo a dominios con mayor posibilidad de responder la misma. Para la propagación selectiva, este componente implementa un esquema de análisis basado en casos (Algoritmo VII.3, Capítulo VII), que le permite sugerir al usuario una lista de dominios con posibilidades de ofrecer la información buscada. A tal fin, cuenta con una base de casos en la cual se han almacenado los datos correspondientes a consultas propagadas con anterioridad. Los detalles técnicos de implementación se encuentran en el Anexo C.

Cada caso almacenado tiene asociado un identificador de la consulta, el dominio remoto que ha dado respuesta, la valoración de la respuesta y los descriptores que fueron asociados a la consulta con su correspondiente profundidad ontológica y semántica en el dominio de local (origen de la consulta).

Las categorías de valoración ofrecidas al usuario, con la cual puede valorar la utilidad de las respuestas ofrecidas por otros dominios, son:

0 – None; 1 – Poor; 2 – Regular; 3 – Good; 4 – Very Good; 5 – Excellent

Ante la solicitud de análisis de propagación de la consulta, se hace un requerimiento a la base de casos recuperando aquellos casos con mayor grado de similitud entre la consulta actual (Qp), que desea ser propagada, y la consulta candidata Qdr , almacenada en la base de casos asociada al dominio remoto dr que dio respuesta. A tal fin se calcula un índice de similitud, $SimQuery(dr, Qp)$, que contempla el número de descriptores coincidentes entre ambas consultas, la representatividad de la consulta Qdr con el dominio remoto dr que le dio respuesta y la valoración que el usuario ha dado al caso.

Siendo:

D : conjunto de dominios de conocimiento o áreas organizacionales gestionados por Onto-DOM.

dl : dominio local desde el que se propaga la consulta Qp , $dl \in D$.

dr : dominio remoto al que se propaga la consulta Qp , $dr \in D$.

Qp : consulta a propagar desde el dominio local dl .

Qdr : consulta previamente ejecutada, almacenada en la base de casos asociada al dominio dr .

$Descriptors(q, d)$: conjunto de descriptores obtenidos mediante análisis semántico de la consulta q en el dominio d .

$OD_i[q, d]$: profundidad ontológica del descriptor i , $i \in Descriptors(q, d)$, en la consulta q mediante el análisis semántico de la consulta en el dominio d .

$SD_i[q, d]$: profundidad semántica del descriptor i , $i \in Descriptors(q, d)$, en la consulta q mediante el análisis semántico de la consulta en el dominio d .

$UR[Qdr]$: valoración del usuario para la consulta almacenada Qdr .

$\#[set]$: tamaño del conjunto set .

$DQdr = \#[Descriptors(Qdr, dl)]$: cantidad de descriptores de la consulta Qdr almacenados en la base de casos.

$MD = \#[Descriptors(Qp, dl) \cap Descriptors(Qdr, dl)]$: cantidad de descriptores que coinciden entre las consultas Qp y Qdr , almacenados en la base de casos.

El índice de similitud entre la consulta Qp a propagar, con la consulta Qdr obtenida de la base de casos que actualmente está implementado en Onto-DOM se calcula según la siguiente fórmula:

$$SimQuery(dr, Qp) = \frac{MD}{DQdr} \left[1 + \sum_{i=1}^{DQdr} (OD_i[Qdr, dl] + SD_i[Qdr, dl]) \right]^{-1} \left(2 - \frac{1}{UR[Qdr]} \right)$$

(a)
 (b)
 (c)

El factor (a) mide el grado de coincidencia de los descriptores de las consultas comparadas. Si los descriptores coinciden completamente el factor vale 1 y en caso contrario tomará valores menores que 1. Si no existen descriptores en común el índice resultará nulo, de modo que sólo serán considerados aquellos casos que contengan al menos un descriptor coincidente.

El factor (b) considera la potencialidad representativa de los descriptores de la consulta almacenada Qdr . Si todos los descriptores coinciden con conceptos ontológicos, la suma de profundidades semántica y ontológica será nula y el factor no afectará al índice de similitud. En caso contrario, a mayor profundidad ontológica y/o semántica este factor provocará una disminución en el valor del índice.

El factor (c) se incluye para considerar la valoración del usuario con respecto a la calidad de la respuesta lograda en el dominio dr .

La efectividad de este índice de similitud para sugerir dominios a propagar se analiza en el Capítulo IX.

En la Figura VIII.9 puede verse la interfaz de propagación donde se le ofrecen al usuario los posibles dominios de propagación para su consulta.

TRAVEL DOMAIN

Others Domains Interface - Propagation Domains Selection

Query: Where can I eat vegetarian dishes near an acropolis?

| Domains without positive rating | | |
|---------------------------------|--|-------------------------------------|
| Domain | URL | Propagate Y/N |
| HISTORY | http://localhost:8484/ontodom/IRPServlet | <input checked="" type="checkbox"/> |
| STRUCTURE | http://localhost:8080/ontodom/IRPServlet | <input type="checkbox"/> |
| GOURMET | http://localhost:8585/ontodom/IRPServlet | <input checked="" type="checkbox"/> |

Figura VIII.9. Módulo de Interfaz a otros Dominios

Los casos a presentarse serían los siguientes:

- i.* que se propague la consulta y la base de casos esté vacía para ese dominio (primera vez que se propaga): no se presentan dominios sugeridos y se propaga a todos los dominios disponibles o bien a los seleccionados por el usuario.
- ii.* la base de casos no está vacía pero no se encuentran similitudes: se le informa al usuario y se le pregunta si desea propagar a todos los dominios.
- iii.* se encuentran similitudes: se presentan las coincidencias encontradas ordenadas de acuerdo al resultado de *SimQuery* y se deja que el usuario decida propagar a todas las coincidencias o a aquellas con valores de *SimQuery* más altos. Además se le presenta una lista de los restantes dominios con el fin de ofrecerle la posibilidad de propagar hacia alguno de ellos.

En la Figura VIII.10 pueden verse resultados del proceso de propagación selectiva junto con la posibilidad de valorar las respuestas recibidas por cada dominio para incrementar la base de casos.

TRAVEL DOMAIN

Others Domains Interface - Propagation Domains Answer

Responses for: Where can I eat vegetarian dishes near an acropolis?

| Rate answer: None <input type="button" value="Submit"/> | <table border="1"> <thead> <tr> <th>Prop. Domain</th> <th>URL</th> </tr> </thead> <tbody> <tr> <td>HISTORY</td> <td>http://localhost:8484</td> </tr> <tr> <th>Words</th> <th>URLs</th> </tr> </tbody> </table> | Prop. Domain | URL | HISTORY | http://localhost:8484 | Words | URLs | | | | | | |
|---|---|--------------|-----|---------|-----------------------|-------|------|---|---|-------------------------------|---|-------------------|---|
| Prop. Domain | URL | | | | | | | | | | | | |
| HISTORY | http://localhost:8484 | | | | | | | | | | | | |
| Words | URLs | | | | | | | | | | | | |
| Rate answer: <input type="button" value="Submit"/> <ul style="list-style-type: none"> None Excellent Very Good Good Regular Poor | <table border="1"> <thead> <tr> <th>Prop. Domain</th> <th>URL</th> </tr> </thead> <tbody> <tr> <td>GOURMET</td> <td>http://localhost:8585</td> </tr> <tr> <th>Words</th> <th>URLs</th> </tr> <tr> <td>[Food, Vegetarian, acropolis, restaurant, vegetarian, view]</td> <td>file/C:/ontodom/corpus/Gourmet/Athens-685.html</td> </tr> <tr> <td>[Food, acropolis, food, view]</td> <td>file/C:/ontodom/corpus/Gourmet/hilton-hotel-athens.html</td> </tr> <tr> <td>[Food, acropolis]</td> <td>file/C:/ontodom/corpus/Gourmet/hilton-athens.en.html</td> </tr> </tbody> </table> | Prop. Domain | URL | GOURMET | http://localhost:8585 | Words | URLs | [Food, Vegetarian, acropolis, restaurant, vegetarian, view] | file/C:/ontodom/corpus/Gourmet/Athens-685.html | [Food, acropolis, food, view] | file/C:/ontodom/corpus/Gourmet/hilton-hotel-athens.html | [Food, acropolis] | file/C:/ontodom/corpus/Gourmet/hilton-athens.en.html |
| Prop. Domain | URL | | | | | | | | | | | | |
| GOURMET | http://localhost:8585 | | | | | | | | | | | | |
| Words | URLs | | | | | | | | | | | | |
| [Food, Vegetarian, acropolis, restaurant, vegetarian, view] | file/C:/ontodom/corpus/Gourmet/Athens-685.html | | | | | | | | | | | | |
| [Food, acropolis, food, view] | file/C:/ontodom/corpus/Gourmet/hilton-hotel-athens.html | | | | | | | | | | | | |
| [Food, acropolis] | file/C:/ontodom/corpus/Gourmet/hilton-athens.en.html | | | | | | | | | | | | |

Figura VIII.10. Resultados del Proceso de Propagación

3. Conclusiones

El prototipo Onto-DOM presentado es un sistema que permite la representación del conocimiento de dominio, la semántica ontológica contenida en consultas en lenguaje natural, y la búsqueda basada en textos en una Memoria Organizacional. Para ello implementa, en sus diferentes capas de la arquitectura que contiene la lógica del sistema, la estrategia de anotado propuesta como aporte de esta tesis (Algoritmo VII.1, Capítulo VII); la estrategia de procesamiento de la consulta y recuperación de información propuesta como aporte de esta tesis (Algoritmo VII.2, Capítulo VII); y las estrategias de propagación completa y de propagación selectiva propuestas como aporte de esta tesis (Algoritmo VII.3, Capítulo VII).

De este modo, el sistema podrá ser utilizado como herramienta para realizar una validación empírica de dichas estrategias, a través de la evaluación del desempeño de los respectivos componentes para un caso de estudio. Este análisis y sus resultados se presentan en el Capítulo IX.

Análisis de Desempeño de Onto – DOM

1. Introducción

El prototipo Onto-DOM presentado en el capítulo anterior, implementa la capa de la arquitectura que contiene la lógica del sistema a través del componente de representación de conocimiento (KRL), el componente de recuperación de información (IRPL) y el componente de propagación a otros dominios (ODI). Según se describió en dicho capítulo, el componente de representación de conocimiento implementa la estrategia de anotado propuesta como aporte de esta tesis (Algoritmo VII.1, Capítulo VII); el componente de recuperación de información implementa la estrategia de procesamiento de la consulta y recuperación de información propuesta como aporte de esta tesis (Algoritmo VII.2, Capítulo VII); y finalmente, el componente de propagación a otros dominios implementa las estrategias de propagación completa y de propagación selectiva propuestas como aporte de esta tesis (Algoritmo VII.3, Capítulo VII). El análisis de desempeño de cada uno de estos componentes,

mediante un caso de estudio, aportará una validación empírica (Recker, 2005) de dichas estrategias.

2. Caso de Estudio

Para el análisis de desempeño de los componentes de Onto-DOM, se tomó como caso de estudio una organización de turismo con tres dominios: un dominio local denominado *Travel*, y dos dominios remotos denominados *África Travel* y *Europa Travel*. Para el dominio local se utilizó una extensión de la ontología *Travel*, la cual se representa en forma de árbol en la Figura IX.1. Si bien esta forma de representación de la ontología permite explicitar solo las relaciones taxonómicas (es-un), no se incluye una representación gráfica de la ontología, en la cual se puedan representar las otras relaciones, debido al gran tamaño de la misma. Para los dominios remotos se utilizaron las ontologías *África Travel* y *Europa Travel* representadas en el Anexo D.

- [owl:Thing](#)
 - [View](#)
 - [Mountain View](#)
 - [Sea View](#)
 - [Ocean View](#)
 - [Panoramic View](#)
 - [Accommodation](#)
 - [Shelter](#)
 - [Tent](#)
 - [Hut](#)
 - [Campground](#)
 - [Hotel](#)
 - [Resort](#)
 - [Motel](#)
 - [Inn](#)
 - [Luxury Hotel](#)
 - [Apartment](#)
 - [Bed and Breadfast](#)
 - [House](#)
 - [Manor](#)
 - [Castle](#)
 - [Condominium](#)
 - [Cabin](#)
 - [Cottage](#)
 - [Bungalow](#)

- [Facility](#)
 - [Villa](#)
 - [Butler Service](#)
 - [Golf Course](#)
 - [Water Garden](#)
 - [Restaurant](#)
 - [Tennis Court](#)
 - [Swimming Pool](#)
 - [Heated Swimming Pool](#)
 - [Internet Access](#)
 - [Bar](#)
 - [Basketball Court](#)
 - [Croquet Court](#)
 - [Spa](#)
 - [Playroom](#)
 - [Mini Bar](#)
 - [Fitness Center](#)
 - [Gym](#)
 - [Room Service](#)
- [Transportation](#)
 - [Ship](#)
 - [Sailboat](#)
 - [Yacht](#)
 - [Ferry](#)
 - [Flight](#)
 - [Train](#)
 - [Bus](#)
 - [Car](#)
 - [Taxi](#)
- [Vacation Period](#)
 - [Month](#)
 - [Season](#)
 - [Summer](#)
 - [Spring](#)
 - [Winter](#)
- [Activity](#)
 - [Sport](#)
 - [Horseback Riding](#)
 - [Shooting](#)
 - [Clay Shooting](#)
 - [Air Rifle Shooting](#)
 - [Tennis](#)
 - [Sailing](#)
 - [Falconry](#)
 - [Polo](#)
 - [Golf](#)
 - [Ice Skating](#)
 - [Basketball](#)
 - [Paddling](#)
 - [Climbing](#)
 - [Rock Climbing](#)
 - [Mountain Climbing](#)
 - [Extreme Sport](#)
 - [Water Skiing](#)
 - [Rock Climbing](#)
 - [Windsurfing](#)

- [Heli Skiing](#)
- [Mountain Climbing](#)
- [Paragliding](#)
- [Parasailing](#)
- [Skiing](#)
 - [Water Skiing](#)
 - [Heli Skiing](#)
- [Biking](#)
 - [Mountain Biking](#)
- [Volleyball](#)
- [Hiking](#)
- [Archery](#)
- [Water Sport](#)
 - [White Water Rafting](#)
 - [Boating](#)
 - [Jet Boating](#)
 - [Canoeing](#)
 - [Swimming](#)
 - [Surfing](#)
 - [Kayaking](#)
 - [Fishing](#)
 - [River Fishing](#)
 - [Fly Fishing](#)
 - [Deep Sea Fishing](#)
 - [Snorkeling](#)
 - [Diving](#)
 - [Scuba Diving](#)
 - [Wreck Diving](#)
 - [Aqua Aerobics](#)
- [Football](#)
- [Relaxation](#)
 - [Yoga](#)
 - [Pilates](#)
 - [Sailing](#)
 - [Fishing](#)
 - [River Fishing](#)
 - [Fly Fishing](#)
 - [Deep Sea Fishing](#)
 - [Spa Treatment](#)
 - [Facial Treatment](#)
 - [Body Treatment](#)
 - [Sauna](#)
 - [Reflexology](#)
 - [Meditation](#)
 - [Massage](#)
 - [Sun Bed](#)
 - [Lymphatic Drainage](#)
 - [Acupressure](#)
 - [Mud Bath](#)
 - [Hydrotherapy](#)
 - [Digitopressure](#)
 - [Steam Bath](#)
 - [Aromatherapy](#)
 - [Jacuzzi](#)
 - [Turkish Bath](#)
 - [Golf](#)
 - [Sunbathing](#)

- [Cuisine](#)
- [Entertainment](#)
 - [Ballet](#)
 - [Dancing](#)
 - [Theatre](#)
 - [Opera](#)
 - [Cinema](#)
 - [Gallery](#)
 - [Aquarium](#)
 - [Safari](#)
 - [Museum](#)
 - [Nightclub](#)
 - [Concert](#)
 - [Casino](#)
- [Adventure](#)
 - [Excursion](#)
 - [Wildlife Watching](#)
 - [Bungee Jumping](#)
 - [Safari](#)
 - [Camping](#)
- [Chef](#)
 - [Travel Class 3](#)
- [Destination](#)
 - [Urban Area](#)
 - [Country](#)
 - [Town](#)
 - [City](#)
 - [Capital](#)
 - [Beach](#)
 - [Rural Area](#)
 - [National Park](#)
 - [Farmland](#)
 - [Budget Hotel Destination](#)
 - [Family Destination](#)
 - [Island](#)
 - [Retiree Destination](#)
 - [Quiet Destination](#)
 - [Backpackers Destination](#)
- [Food](#)
 - [Diet Food](#)
 - [Spicy](#)
 - [Finger Food](#)
 - [Snack](#)
 - [Organic](#)
 - [Seafood](#)
 - [Vegetarian](#)

Figura IX.1. Ontología *Travel* Extendida

Los documentos anotados son de acceso libre en Internet y fueron elegidos mediante un mecanismo de selección aleatoria. Dado que el propósito no es validar patrones de comportamiento ni principios de validez universal, sino

la lógica de los algoritmos que definen la estrategia de anotado, de recuperación y de propagación, la selección aleatoria no tuvo más propósito que el de generar una muestra no homogénea de documentos, lo cual fue corroborado mediante un análisis lingüístico. El *corpus* quedó conformado por 150 documentos (conteniendo 35.091 palabras) para el dominio local y 80 documentos para los dominios remotos.

3. Análisis de Desempeño

3.1 Componente de Representación de Conocimiento

3.1.1 Procedimiento

La valoración del desempeño de la estrategia de anotado, propuesta en esta tesis, se realizó mediante una comparación de los resultados obtenidos del anotado automático, generado utilizando el módulo de representación del conocimiento que implementa dicha estrategia, con los resultados obtenidos de un anotado manual. Se trabajó con el *corpus* de documentos del dominio local *Travel*.

3.1.2 Métricas

Para proceder a la comparación de resultados en términos cuantitativos, se aplicaron las siguientes métricas de acuerdo a las definiciones de Yang (Yang; 1999): *recall*, *precision*, *fallout* y *accuracy*.

El *recall* es una medida de cuán bien la estrategia se desempeña en encontrar conceptos revelantes. El *recall* es 100% cuando cada concepto relevante del documento es anotado. En teoría, es fácil alcanzar un buen *recall* simplemente anotando cada sustantivo en el texto. Por lo tanto, el *recall* en sí

mismo, no es una buena medida de la calidad de la estrategia. La *precision*, por otro lado, es una medida de cuán bien la estrategia se desempeña al no anotar sustantivos que no son relevantes. Finalmente, el *fallout* es la medida de cuán rápidamente cae la *precision* a medida que el *recall* sube, en otras palabras, representa la porción de conceptos no relevantes que fueron anotados. La *accuracy* representa la exactitud del anotado.

3.1.3 Anotado manual

Cada uno de los documentos del *corpus* fue procesado lingüísticamente determinando el conjunto de descriptores correspondiente (sustantivos presentes en el documento y sustantivos generados por expansión semántica). En base a dichos descriptores cada documento fue anotado manualmente a la ontología *Travel*.

3.1.4 Anotado automático

El *corpus* de documentos fue anotado en forma automática a la ontología *Travel* utilizando el módulo de representación del conocimiento que implementa la estrategia propuesta en esta tesis.

3.1.5 Resultados de la comparación

En la Tabla IX.1 se muestran los resultados del anotado automático en comparación al anotado manual, cuantificado a través de las métricas definidas en la sección 3.1.2. Dichas métricas fueron calculadas del siguiente modo:

$$\textit{Recall} \quad R = VP/(VP+FN)$$

$$\textit{Precision} \quad P = VP/(VP+FP)$$

$$\textit{Fallout} \quad F = FP/(FP+VN)$$

$$\text{Accuracy} \quad A = (VP+VN)/N$$

Donde:

VP, Verdadero Positivo: Descriptor semántico anotado que debía ser anotado

FP, Falso Positivo: Descriptor semántico anotado que no debía ser anotado

VN, Verdadero Negativo: Descriptor semántico no anotado que no debía ser anotado

FN, Falso Negativo: Descriptor semántico no anotado que debía ser anotado

N: Número de descriptores semánticos

D: Documento

AC: Cantidad de descriptores asignados a categoría

NoA: Cantidad de descriptores no asignados a categoría

| D | AC | NoA | VP | FP | VN | FN | N | R | P | F | A |
|-----|----|-----|----|----|----|----|-----|--------|--------|--------|--------|
| 001 | 41 | 62 | 33 | 8 | 60 | 2 | 103 | 0,9429 | 0,8049 | 0,1176 | 0,9029 |
| 002 | 17 | 35 | 14 | 3 | 35 | 0 | 52 | 0,9714 | 0,8142 | 0,0983 | 0,9226 |
| 003 | 23 | 39 | 16 | 7 | 37 | 2 | 62 | 0,9439 | 0,7747 | 0,1186 | 0,9000 |
| 004 | 29 | 58 | 21 | 8 | 56 | 2 | 87 | 0,9362 | 0,7620 | 0,1202 | 0,8963 |
| 005 | 17 | 31 | 12 | 5 | 31 | 0 | 48 | 0,9490 | 0,7508 | 0,1239 | 0,8962 |
| 006 | 43 | 59 | 36 | 7 | 56 | 3 | 102 | 0,9446 | 0,7652 | 0,1218 | 0,8972 |
| 007 | 28 | 62 | 21 | 7 | 56 | 6 | 90 | 0,9208 | 0,7630 | 0,1203 | 0,8912 |
| 008 | 15 | 29 | 10 | 5 | 28 | 1 | 44 | 0,9193 | 0,7510 | 0,1242 | 0,8878 |
| 009 | 11 | 42 | 3 | 8 | 41 | 1 | 53 | 0,9005 | 0,6979 | 0,1285 | 0,8814 |
| 010 | 15 | 43 | 11 | 4 | 42 | 1 | 58 | 0,9021 | 0,7014 | 0,1244 | 0,8846 |
| 011 | 17 | 31 | 10 | 7 | 29 | 2 | 48 | 0,8959 | 0,6911 | 0,1307 | 0,8781 |
| 012 | 11 | 32 | 4 | 7 | 29 | 3 | 43 | 0,8688 | 0,6638 | 0,1360 | 0,8688 |
| 013 | 19 | 49 | 16 | 3 | 45 | 4 | 68 | 0,8636 | 0,6775 | 0,1304 | 0,8710 |
| 014 | 16 | 43 | 9 | 7 | 41 | 2 | 59 | 0,8603 | 0,6693 | 0,1315 | 0,8693 |
| 015 | 33 | 67 | 24 | 9 | 63 | 4 | 100 | 0,8601 | 0,6732 | 0,1311 | 0,8694 |
| 016 | 15 | 49 | 8 | 7 | 48 | 1 | 64 | 0,8619 | 0,6644 | 0,1308 | 0,8697 |
| 017 | 12 | 39 | 7 | 5 | 38 | 1 | 51 | 0,8627 | 0,6597 | 0,1300 | 0,8705 |
| 018 | 26 | 66 | 17 | 9 | 63 | 3 | 92 | 0,8620 | 0,6593 | 0,1297 | 0,8704 |
| 019 | 16 | 40 | 12 | 4 | 40 | 0 | 56 | 0,8692 | 0,6641 | 0,1276 | 0,8735 |
| 020 | 21 | 25 | 17 | 4 | 24 | 1 | 46 | 0,8730 | 0,6714 | 0,1284 | 0,8744 |
| 021 | 14 | 39 | 9 | 5 | 38 | 1 | 53 | 0,8743 | 0,6700 | 0,1278 | 0,8750 |
| 022 | 32 | 69 | 24 | 8 | 65 | 4 | 101 | 0,8735 | 0,6737 | 0,1270 | 0,8752 |
| 023 | 23 | 29 | 21 | 2 | 27 | 2 | 52 | 0,8752 | 0,6841 | 0,1245 | 0,8773 |
| 024 | 29 | 47 | 20 | 9 | 44 | 3 | 76 | 0,8750 | 0,6843 | 0,1264 | 0,8759 |
| 025 | 12 | 24 | 9 | 3 | 23 | 1 | 36 | 0,8760 | 0,6869 | 0,1259 | 0,8764 |

| | | | | | | | | | | | |
|--------------------|-----|-----|-----|-----|-----|-----|-----|---------------|---------------|---------------|---------------|
| 026 | 21 | 41 | 15 | 6 | 37 | 4 | 62 | 0,8727 | 0,6880 | 0,1264 | 0,8749 |
| 027 | 21 | 33 | 15 | 6 | 32 | 1 | 54 | 0,8751 | 0,6890 | 0,1276 | 0,8748 |
| 028 | 47 | 76 | 25 | 22 | 70 | 6 | 123 | 0,8726 | 0,6834 | 0,1316 | 0,8711 |
| 029 | 13 | 46 | 8 | 5 | 43 | 3 | 59 | 0,8676 | 0,6810 | 0,1306 | 0,8709 |
| 030 | 19 | 40 | 12 | 7 | 39 | 1 | 59 | 0,8694 | 0,6794 | 0,1314 | 0,8707 |
| 031 | 21 | 31 | 19 | 2 | 28 | 3 | 52 | 0,8693 | 0,6866 | 0,1293 | 0,8717 |
| 032 | 39 | 58 | 26 | 13 | 54 | 4 | 97 | 0,8692 | 0,6860 | 0,1313 | 0,8703 |
| 033 | 55 | 80 | 36 | 19 | 73 | 7 | 135 | 0,8682 | 0,6851 | 0,1336 | 0,8684 |
| 034 | 48 | 77 | 37 | 11 | 73 | 4 | 125 | 0,8692 | 0,6876 | 0,1335 | 0,8687 |
| 035 | 51 | 80 | 38 | 13 | 76 | 4 | 131 | 0,8702 | 0,6892 | 0,1339 | 0,8687 |
| 036 | 17 | 45 | 12 | 5 | 44 | 1 | 62 | 0,8717 | 0,6897 | 0,1330 | 0,8697 |
| 037 | 15 | 24 | 11 | 4 | 22 | 2 | 39 | 0,8710 | 0,6909 | 0,1335 | 0,8691 |
| 038 | 18 | 43 | 12 | 6 | 41 | 2 | 61 | 0,8706 | 0,6902 | 0,1334 | 0,8691 |
| 039 | 25 | 59 | 15 | 10 | 56 | 3 | 84 | 0,8697 | 0,6879 | 0,1339 | 0,8684 |
| 040 | 22 | 32 | 17 | 5 | 32 | 0 | 54 | 0,8729 | 0,6900 | 0,1339 | 0,8694 |
| 041 | 15 | 27 | 10 | 5 | 26 | 1 | 42 | 0,8738 | 0,6895 | 0,1346 | 0,8691 |
| 042 | 22 | 54 | 13 | 9 | 48 | 6 | 76 | 0,8693 | 0,6871 | 0,1351 | 0,8675 |
| 043 | 17 | 37 | 11 | 6 | 32 | 5 | 54 | 0,8651 | 0,6862 | 0,1356 | 0,8659 |
| 044 | 20 | 29 | 14 | 6 | 28 | 1 | 49 | 0,8666 | 0,6865 | 0,1366 | 0,8657 |
| 045 | 10 | 64 | 3 | 7 | 50 | 14 | 74 | 0,8513 | 0,6779 | 0,1363 | 0,8624 |
| 046 | 19 | 60 | 9 | 10 | 57 | 3 | 79 | 0,8491 | 0,6735 | 0,1365 | 0,8618 |
| 047 | 18 | 27 | 9 | 9 | 26 | 1 | 45 | 0,8502 | 0,6698 | 0,1391 | 0,8600 |
| 048 | 50 | 77 | 33 | 17 | 75 | 2 | 127 | 0,8521 | 0,6696 | 0,1401 | 0,8598 |
| 049 | 55 | 112 | 38 | 17 | 107 | 5 | 167 | 0,8528 | 0,6700 | 0,1400 | 0,8600 |
| 050 | 12 | 40 | 8 | 4 | 40 | 0 | 52 | 0,8557 | 0,6699 | 0,1390 | 0,8612 |
| 051 | 15 | 36 | 9 | 6 | 36 | 0 | 51 | 0,8585 | 0,6686 | 0,1391 | 0,8616 |
| 052 | 29 | 58 | 17 | 12 | 57 | 1 | 87 | 0,8602 | 0,6670 | 0,1398 | 0,8614 |
| 053 | 21 | 40 | 13 | 8 | 39 | 1 | 61 | 0,8615 | 0,6661 | 0,1403 | 0,8613 |
| 054 | 20 | 47 | 11 | 9 | 46 | 1 | 67 | 0,8625 | 0,6639 | 0,1408 | 0,8611 |
| 055 | 21 | 35 | 9 | 12 | 34 | 1 | 56 | 0,8632 | 0,6597 | 0,1430 | 0,8594 |
| 056 | 14 | 36 | 8 | 6 | 34 | 2 | 50 | 0,8620 | 0,6581 | 0,1431 | 0,8590 |
| 057 | 19 | 30 | 13 | 6 | 30 | 0 | 49 | 0,8645 | 0,6585 | 0,1435 | 0,8593 |
| 058 | 42 | 56 | 33 | 9 | 51 | 5 | 98 | 0,8645 | 0,6607 | 0,1436 | 0,8593 |
| 059 | 39 | 48 | 31 | 8 | 46 | 2 | 87 | 0,8658 | 0,6630 | 0,1437 | 0,8597 |
| 060 | 31 | 47 | 19 | 12 | 43 | 4 | 78 | 0,8651 | 0,6622 | 0,1449 | 0,8587 |
| 061 | 26 | 51 | 18 | 8 | 46 | 5 | 77 | 0,8638 | 0,6627 | 0,1450 | 0,8582 |
| 062 | 27 | 58 | 21 | 6 | 56 | 2 | 85 | 0,8646 | 0,6645 | 0,1442 | 0,8590 |
| 063 | 27 | 46 | 21 | 6 | 44 | 2 | 73 | 0,8654 | 0,6663 | 0,1438 | 0,8595 |
| 064 | 20 | 51 | 11 | 9 | 48 | 3 | 71 | 0,8641 | 0,6645 | 0,1440 | 0,8590 |
| 065 | 54 | 97 | 44 | 10 | 95 | 2 | 151 | 0,8655 | 0,6668 | 0,1433 | 0,8600 |
| 066 | 27 | 36 | 21 | 6 | 36 | 0 | 63 | 0,8676 | 0,6685 | 0,1433 | 0,8607 |
| 067 | 12 | 82 | 9 | 3 | 82 | 0 | 94 | 0,8695 | 0,6697 | 0,1417 | 0,8623 |
| 068 | 34 | 75 | 21 | 13 | 71 | 4 | 109 | 0,8691 | 0,6689 | 0,1419 | 0,8620 |
| 069 | 14 | 31 | 11 | 3 | 28 | 3 | 45 | 0,8679 | 0,6706 | 0,1412 | 0,8621 |
| Valor medio | | | | | | | | 0,8679 | 0,6706 | 0,1412 | 0,8621 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |

Tabla IX.1. Resultados del Anotado Automático en Comparación al Anotado Manual.

En las Figuras IX.2, IX.3, IX.4 y IX.5 se muestra la evolución del valor medio de las diferentes métricas en función de la cantidad de documentos anotados. Como puede observarse, los valores muestran una convergencia que se va consolidando en la medida que aumenta el número de documentos procesados. Dichos valores de convergencia se muestran estables a partir de los 60 documentos procesados, y se mantuvieron invariantes durante los próximos 90 documentos, completando el total de 150 documentos procesados. Esto indica que la cantidad de documentos procesados ha sido suficiente para determinar los valores medios de las diferentes métricas.

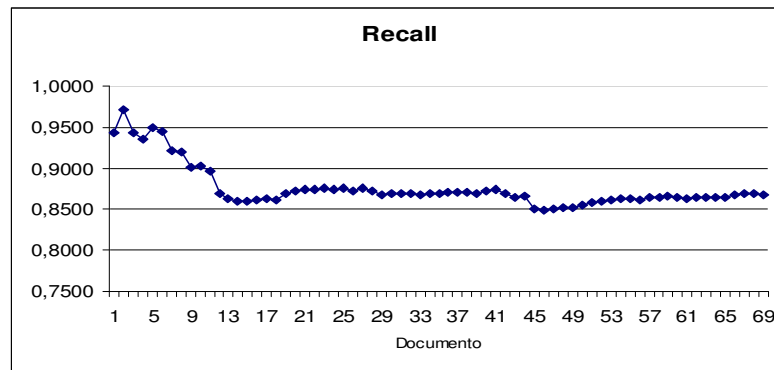


Figura IX.2. Valor Promedio del *Recall* de la Estrategia de Anotado.

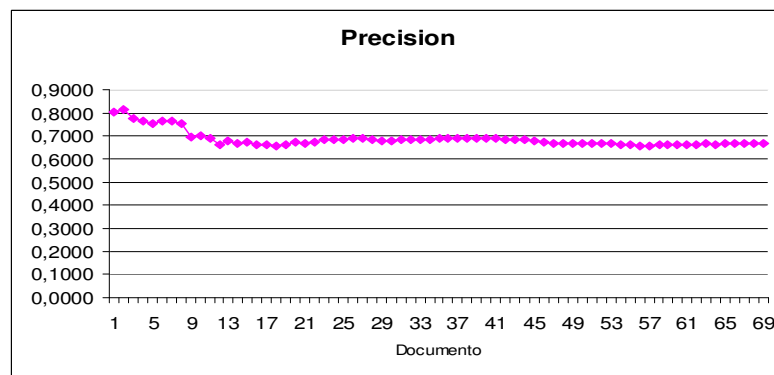


Figura IX.3. Valor Promedio del *Precision* de la Estrategia de Anotado.

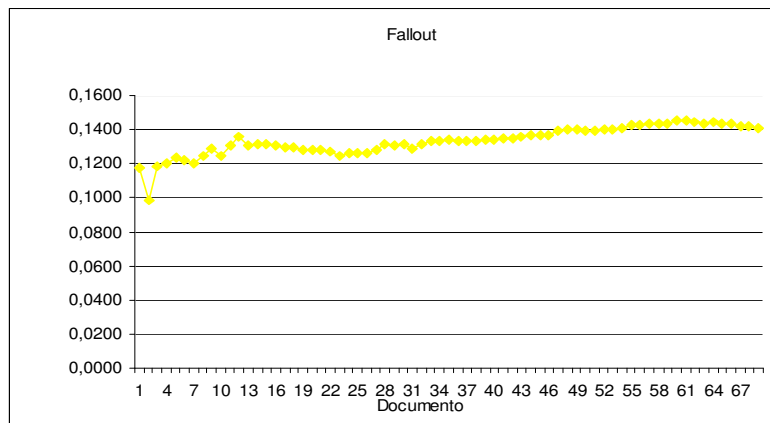


Figura IX.4. Valor Promedio del *Fallout* de la Estrategia de Anotado.

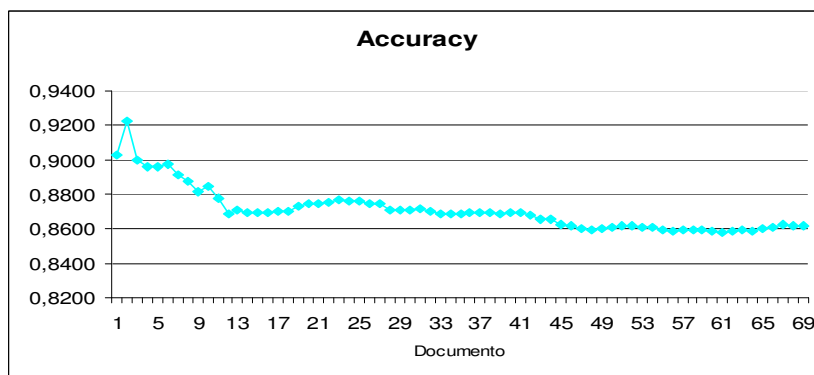


Figura IX.5. Valor Promedio del *Accuracy* de la Estrategia de Anotado.

En la Tabla IX.2 puede verse el valor de convergencia de las diferentes métricas, resultante de la evaluación realizada sobre los 150 documentos del *corpus*.

| Recall | Precision | Fallout | Accuracy |
|--------|-----------|---------|----------|
| 87% | 70% | 14% | 86% |

Tabla IX.2. Resultados del Análisis de Desempeño de la Estrategia de Anotado

Un análisis más exhaustivo del valor obtenido como *recall* indicó que el 82% de los conceptos relevantes no anotados corresponden a nombres de destinos turísticos que, o bien eran lugares que WordNet no reconocía (por ejemplo: Caicos) o tipos de destino que no se tomaron en cuenta en la ontología de dominio (por ejemplo: archipiélago). De acuerdo a lo anterior, el *recall* puede

ser mejorado mediante la utilización de listas de vocabulario asociado al dominio y extendiendo la ontología de dominio.

3.2 Componente de Recuperación de Información

3.2.1 Procedimiento

El desempeño de la estrategia, propuesta en esta tesis, para procesamiento de una consulta formulada en lenguaje natural y de recuperación del objeto de conocimiento buscado, fue valorado comparando los resultados obtenidos del procesamiento automático, realizado utilizando el módulo de recuperación de información que implementa dicha estrategia, con los resultados obtenidos de una búsqueda manual.

Los documentos fueron clasificados en relación a la consulta en dos grupos: *relacionados* y *no relacionados*. Los documentos relacionados son los objetos de conocimiento que responden la consulta y los documentos no relacionados son los que no contienen información relevante para la consulta. Para cada consulta se consideran los m_j primeros documentos recuperados dentro de los cuales se encuentran todos los documentos *relacionados* con la consulta, que fueron determinados por la búsqueda manual.

3.2.2 Métricas

Para proceder a la comparación de resultados en términos cuantitativos se utilizó la métrica más estándar en la comunidad TREC (Text REtrieval Conference) la cual es el *promedio de la precisión media*, MAP (Mean Average Precision) (Manning, 2008). Esta métrica permite determinar la relación entre la *precision* y el *recall* del algoritmo de búsqueda, definiendo un valor promedio de la *precision* para valores del *recall* que varían en el rango de 0 a 100 %.

3.2.3 Búsqueda manual

En la Tabla IX.3 se muestran a modo de ejemplo 25 de las consultas analizadas. Por cada consulta se realizó un análisis semántico de los documentos del *corpus* a efectos de determinar el conjunto de documentos que respondían correctamente dicha pregunta.

| ID_consulta | Consulta |
|-------------|--|
| q01 | Where can I find massage therapy? |
| q02 | Where can I do a hut reservation? |
| q03 | What Island can I play volleyball? |
| q04 | Where can I do deep sea fishing? |
| q05 | Where can I eat seafood in Turkey? |
| q06 | Where can I visit a gallery in summer? |
| q07 | What beach has bungalow with wireless Internet access and telephone? |
| q08 | What hotel has ice-skating? |
| q09 | Where can I do rock climbing? |
| q10 | Where can I do extreme sport? |
| q11 | Where can I walk through the woods and wildlife watching from the beach? |
| q12 | What beach can I play golf? |
| q13 | What beach can I take a golf course? |
| q14 | What hotel has panoramic view in its rooms? |
| q15 | What city has a nightclub? |
| q16 | What hotel has jacuzzi? |
| q17 | What spa has sauna? |
| q18 | Where can I play football? |
| q19 | What island offers water sport activities? |
| q20 | Where can I play tennis? |
| q21 | Where can I do Yoga? |
| q22 | What restaurant can I eat? |
| q23 | Where can I watch a theatre? |
| q24 | Where can I do safari? |
| q25 | Where can I play basketball? |
| | |

Tabla IX.3. Consultas Analizadas para Evaluar la Estrategia de Recuperación

3.2.4 Búsqueda automática

Por cada consulta presentada en la Tabla IX.3 se realizó una búsqueda automática utilizando el módulo de recuperación de información que implementa la estrategia propuesta en esta tesis.

3.2.5 Resultados de la comparación

En la Tabla IX.4 se muestran los resultados de la búsqueda. Para cada consulta q_j se representa con 1 una respuesta *relacionada* y con 0 una *no relacionada*. También muestra la posición que ocupó en la respuesta cada documento d_k recuperado, y el total de documentos recuperados m_j para la consulta q_j dentro de los cuales se encuentran todos los documentos *relacionados* a dicha consulta.

| q_j | d_k | | | | | | | | | | | | | | | m_j |
|-------|-------|---|---|---|---|---|---|---|---|----|----|----|----|----|----|-------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | |
| q01 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | | | | | | | | 8 |
| q02 | 1 | 0 | 0 | 0 | 1 | | | | | | | | | | | 5 |
| q03 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | | | | | | | | 8 |
| q04 | 1 | 1 | | | | | | | | | | | | | | 2 |
| q05 | 1 | | | | | | | | | | | | | | | 1 |
| q06 | 1 | | | | | | | | | | | | | | | 1 |
| q07 | 1 | | | | | | | | | | | | | | | 1 |
| q08 | 0 | 0 | 1 | | | | | | | | | | | | | 3 |
| q09 | 1 | 1 | | | | | | | | | | | | | | 2 |
| q10 | 1 | 1 | | | | | | | | | | | | | | 2 |
| q11 | 1 | | | | | | | | | | | | | | | 1 |
| q12 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | | | | | 7 |
| q13 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | | | | | | 10 |
| q14 | 1 | | | | | | | | | | | | | | | 1 |
| q15 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | | | | | 11 |
| q16 | 1 | 0 | 1 | 1 | | | | | | | | | | | | 4 |
| q17 | 1 | 1 | 1 | 1 | 1 | | | | | | | | | | | 5 |
| q18 | 1 | | | | | | | | | | | | | | | 1 |
| q19 | 1 | | | | | | | | | | | | | | | 1 |
| q20 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | | | | | | 6 |
| q21 | 1 | | | | | | | | | | | | | | | 1 |
| q22 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 15 |
| q23 | 0 | 0 | 1 | | | | | | | | | | | | | 3 |
| q24 | 1 | 1 | | | | | | | | | | | | | | 2 |
| q25 | 1 | 1 | | | | | | | | | | | | | | 2 |

Tabla IX.4. Resultados Cuantitativos de la Búsqueda

Los datos de la Tabla IX.4 permiten calcular la métrica MAP, definida en la sección 3.2.2. Dicha métrica fue calculada del siguiente modo:

$$MAP(Q) = (1/|Q|) * \sum (1/m_j * \sum (Precision(R_{jk}); k=1, \dots, m_j); j=1, \dots, |Q|)$$

Donde:

q_j = consulta j .

Q = Conjunto de consultas = $\{q_1, q_2, \dots, q_j, \dots, q_n\}$.

n = número de consultas.

$|Q|$ = Cantidad de elementos del conjunto Q .

m_j = Cantidad de documentos recuperados para la consulta q_j dentro de los cuales se encuentran todos los documentos *relacionados* a dicha consulta.

d_k = Documento recuperado en la posición k , para $k = 1, \dots, m_j$.

R_{jk} : Cantidad de documentos *relacionados* recuperados para la consulta q_j desde la primer posición hasta la posición de d_k . Esto es, $R_{jk} = \{d_1, \dots, d_k\}$.

En la Tabla IX.5 se muestran los valores de R_{jk} para el conjunto de consultas Q .

| q_j | d_k | | | | | | | | | | | | | | |
|-------|-------|---|---|---|---|---|---|---|---|----|----|----|----|----|----|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| q01 | 1 | 2 | 3 | 4 | 4 | 5 | 6 | 7 | | | | | | | |
| q02 | 1 | 1 | 1 | 1 | 2 | | | | | | | | | | |
| q03 | 1 | 1 | 1 | 1 | 2 | 3 | 4 | 5 | | | | | | | |
| q04 | 1 | 2 | | | | | | | | | | | | | |
| q05 | 1 | | | | | | | | | | | | | | |
| q06 | 1 | | | | | | | | | | | | | | |
| q07 | 1 | | | | | | | | | | | | | | |
| q08 | 0 | 0 | 1 | | | | | | | | | | | | |
| q09 | 1 | 2 | | | | | | | | | | | | | |
| q10 | 1 | 2 | | | | | | | | | | | | | |
| q11 | 1 | | | | | | | | | | | | | | |
| q12 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | | | | | | | | |
| q13 | 1 | 2 | 3 | 4 | 4 | 5 | 6 | 6 | 7 | 8 | | | | | |
| q14 | 1 | | | | | | | | | | | | | | |
| q15 | 1 | 2 | 2 | 2 | 2 | 2 | 3 | 4 | 5 | 6 | 7 | | | | |
| q16 | 1 | 1 | 2 | 3 | | | | | | | | | | | |

| | | | | | | | | | | | | | | | |
|-----|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|
| q17 | 1 | 2 | 3 | 4 | 5 | | | | | | | | | | |
| q18 | 1 | | | | | | | | | | | | | | |
| q19 | 1 | | | | | | | | | | | | | | |
| q20 | 1 | 2 | 3 | 4 | 5 | 6 | | | | | | | | | |
| q21 | 1 | | | | | | | | | | | | | | |
| q22 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| q23 | 0 | 0 | 1 | | | | | | | | | | | | |
| q24 | 1 | 2 | | | | | | | | | | | | | |
| q25 | 1 | 2 | | | | | | | | | | | | | |

Tabla IX.5. Valores de R_{jk}

La Tabla IX.6 muestra la evolución del valor de la métrica *Precision*(R_{jk}) en función de la cantidad de documentos *relacionados* recuperados d_k para la consulta q_j .

| q_j | d_k | | | | | | | | | | | | | | |
|-------|-------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|----|----|----|----|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| q01 | 1 | 1 | 1 | 1 | 0,8 | 0,8 | 0,9 | 0,9 | | | | | | | |
| q02 | 1 | 0,5 | 0,3 | 0,3 | 0,4 | | | | | | | | | | |
| q03 | 1 | 0,5 | 0,3 | 0,3 | 0,4 | 0,5 | 0,6 | 0,6 | | | | | | | |
| q04 | 1 | 1 | | | | | | | | | | | | | |
| q05 | 1 | | | | | | | | | | | | | | |
| q06 | 1 | | | | | | | | | | | | | | |
| q07 | 1 | | | | | | | | | | | | | | |
| q08 | 0 | 0 | 0,3 | | | | | | | | | | | | |
| q09 | 1 | 1 | | | | | | | | | | | | | |
| q10 | 1 | 1 | | | | | | | | | | | | | |
| q11 | 1 | | | | | | | | | | | | | | |
| q12 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | | | | |
| q13 | 1 | 1 | 1 | 1 | 0,8 | 0,8 | 0,9 | 0,8 | 0,8 | 0,8 | | | | | |
| q14 | 1 | | | | | | | | | | | | | | |
| q15 | 1 | 1 | 0,7 | 0,5 | 0,4 | 0,3 | 0,4 | 0,5 | 0,6 | 0,6 | 0,6 | | | | |
| q16 | 1 | 0,5 | 0,7 | 0,8 | | | | | | | | | | | |
| q17 | 1 | 1 | 1 | 1 | 1 | | | | | | | | | | |
| q18 | 1 | | | | | | | | | | | | | | |
| q19 | 1 | | | | | | | | | | | | | | |
| q20 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | | | | | |
| q21 | 1 | | | | | | | | | | | | | | |
| q22 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| q23 | 0 | 0 | 0,3 | | | | | | | | | | | | |
| q24 | 1 | 1 | | | | | | | | | | | | | |
| q25 | 1 | 1 | | | | | | | | | | | | | |

Tabla IX.6. Valores de *Precision*(R_{jk})

La Tabla IX.7 muestra la evolución del valor de la métrica *Recall*(R_{jk}) en función del número de documentos recuperados d_k para cada consulta q_j .

| q_j | d_k | | | | | | | | | | | | | | |
|-------|-------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|----|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| q01 | 0,1 | 0,3 | 0,4 | 0,6 | 0,6 | 0,7 | 0,9 | 1 | | | | | | | |
| q02 | 0,5 | 0,5 | 0,5 | 0,5 | 1 | | | | | | | | | | |
| q03 | 0,2 | 0,2 | 0,2 | 0,2 | 0,4 | 0,6 | 0,8 | 1 | | | | | | | |
| q04 | 0,5 | 1 | | | | | | | | | | | | | |
| q05 | 1 | | | | | | | | | | | | | | |
| q06 | 1 | | | | | | | | | | | | | | |
| q07 | 1 | | | | | | | | | | | | | | |
| q08 | 0 | 0 | 1 | | | | | | | | | | | | |
| q09 | 0,5 | 1 | | | | | | | | | | | | | |
| q10 | 0,5 | 1 | | | | | | | | | | | | | |
| q11 | 1 | | | | | | | | | | | | | | |
| q12 | 0,1 | 0,3 | 0,4 | 0,6 | 0,7 | 0,9 | 1 | | | | | | | | |
| q13 | 0,1 | 0,3 | 0,4 | 0,5 | 0,5 | 0,6 | 0,8 | 0,8 | 0,9 | 1 | | | | | |
| q14 | 1 | | | | | | | | | | | | | | |
| q15 | 0,1 | 0,3 | 0,3 | 0,3 | 0,3 | 0,3 | 0,4 | 0,6 | 0,7 | 0,9 | 1 | | | | |
| q16 | 0,3 | 0,3 | 0,7 | 1 | | | | | | | | | | | |
| q17 | 0,2 | 0,4 | 0,6 | 0,8 | 1 | | | | | | | | | | |
| q18 | 1 | | | | | | | | | | | | | | |
| q19 | 1 | | | | | | | | | | | | | | |
| q20 | 0,2 | 0,3 | 0,5 | 0,7 | 0,8 | 1 | | | | | | | | | |
| q21 | 1 | | | | | | | | | | | | | | |
| q22 | 0,1 | 0,1 | 0,2 | 0,3 | 0,3 | 0,4 | 0,5 | 0,5 | 0,6 | 0,7 | 0,7 | 0,8 | 0,9 | 0,9 | 1 |
| q23 | 0 | 0 | 1 | | | | | | | | | | | | |
| q24 | 0,5 | 1 | | | | | | | | | | | | | |
| q25 | 0,5 | 1 | | | | | | | | | | | | | |

Tabla IX.7. Valores de *Recall*(R_{jk})

La Figura IX.6 muestra la evolución del valor de la métrica *Precision*(R_{jk}) vs. *Recall*(R_{jk}) para las consulta q12 y q1. Si bien ambas consultas tienen 7 documentos relacionados (q01 y q12 en Tabla IX.4), el proceso de recuperación por error recuperó para la consulta q1 un documento no relacionado, por lo que el valor de su métrica *Precision* se ve afectada tomando valores menores a 1. Es también necesario mostrar que el documento no relacionado fue ubicado en la posición $k=5$ (q01 en Tabla IX.4). En la Figura IX.6 se ha graficado como q1' el supuesto caso en que el documento no relacionado recuperado para la consulta q1 hubiese sido ubicado en la posición

$k=2$ a efectos de mostrar un resultado que sería menos favorable que el que realmente sucedió. La métrica MAP utilizada, captura esta diferencia contabilizando el área debajo de la curva $Precision(R_{jk})$ vs. $Recall(R_{jk})$, la cual, como puede verse de los diferentes gráficos de la Figura IX.6, se reduce en la medida que el resultado es más desfavorable. Esto muestra la bondad de esta métrica para evaluar los resultados de la recuperación.

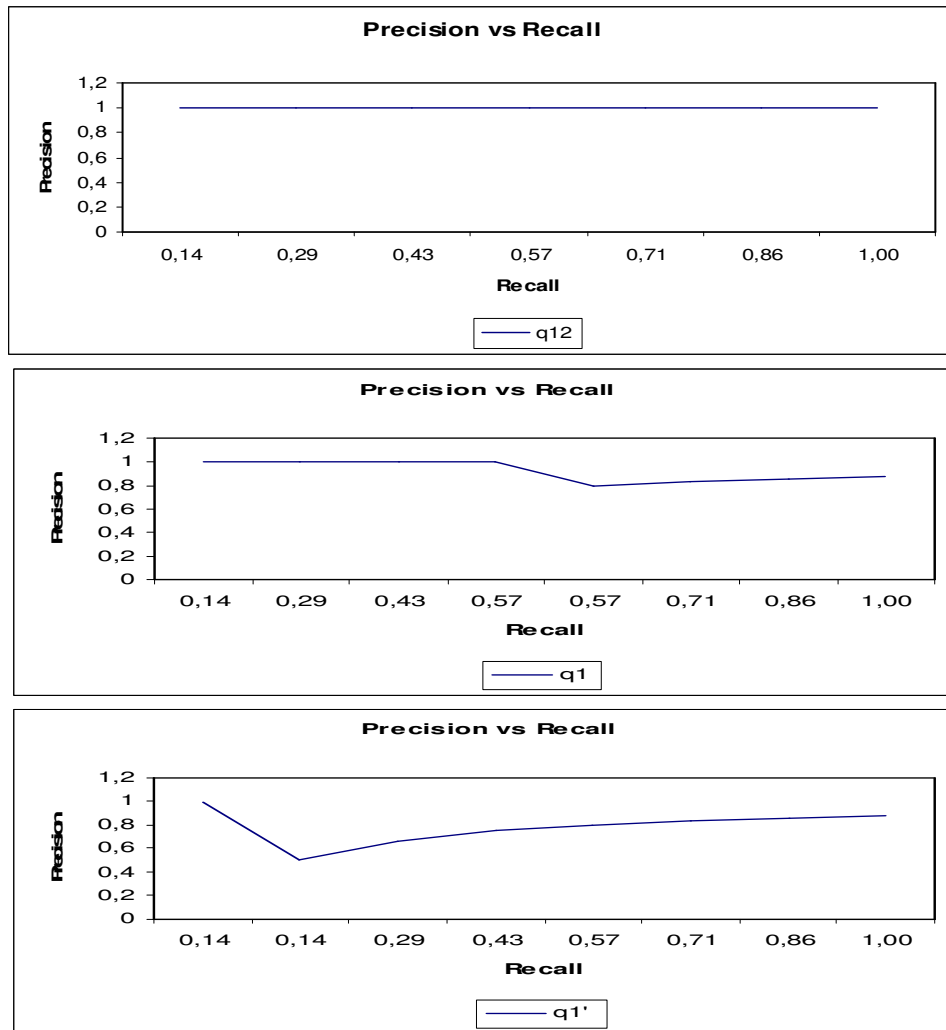


Figura IX.6. Gráfica de la Métrica $Precision(R_{jk})$ vs. $Recall(R_{jk})$

La Figura IX.7 muestra la evolución del MAP en función de la cantidad de consultas procesadas. Como puede observarse, los valores muestran una

convergencia que se va estabilizando en torno a 0,86 a medida que aumenta el número de consultas procesadas.

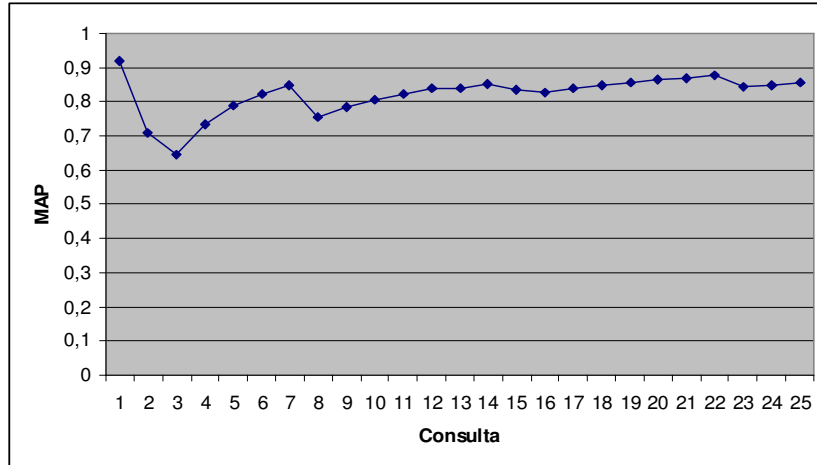


Figura IX.7. Valores del MAP en Función del Número de Consultas

A efectos de una mejor evaluación del desempeño de la estrategia para procesamiento de una consulta formulada en lenguaje natural y de recuperación del objeto de conocimiento buscado, se realizó un estudio similar al descrito en esta sección, utilizando el módulo de recuperación de información en los dominios remotos *África Travel* y *Europa Travel*. Las consultas formuladas y los resultados de la recuperación para los dominios remotos se muestran en el Anexo D.

En las Figuras IX.8 y IX.9 se muestra la evolución del MAP en función de la cantidad de consultas procesadas para los dominios remotos. Como puede observarse, los valores muestran una convergencia que se va estabilizando en torno a 0,84 para el dominio remoto *África Travel* y 0,88 para *Europa Travel*, a medida que aumenta el número de consultas procesadas.

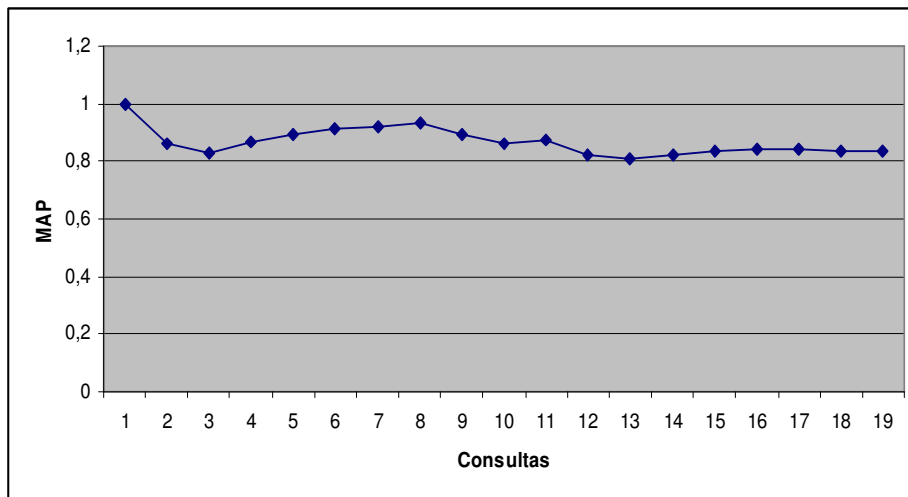


Figura IX.8. Valores del MAP en Función del Número de Consultas para *África Travel*

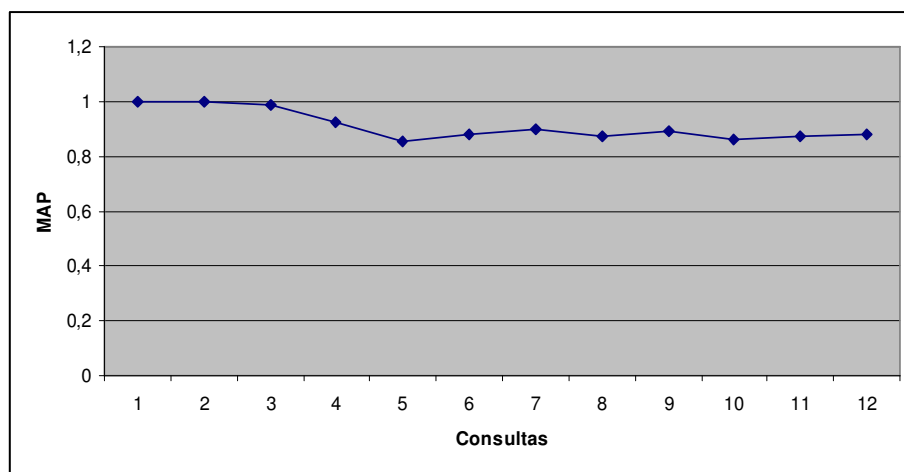


Figura IX.9. Valores del MAP en Función del Número de Consultas para *Europa Travel*

3.3 Componente de Propagación a otros Dominios

El componente de propagación a otros dominios implementa las estrategias de propagación completa y de propagación selectiva. De éstas, la estrategia de propagación selectiva utiliza un esquema de análisis basado en

casos (Sección 2.4.2, Capítulo VIII) del cual depende su desempeño. En esta sección se describe el proceso de valoración de desempeño de dicha estrategia.

3.3.1 Procedimiento

El desempeño de la estrategia de propagación selectiva, propuesta en esta tesis, fue valorado comparando los resultados obtenidos de propagar las consultas a los dominios sugeridos por el componente que implementa dicha estrategia versus los resultados obtenidos de propagar las consultas a todos los dominios. Las consultas fueron formuladas desde el dominio local *Travel* y propagadas a los dominios remotos *África Travel* y *Europa Travel*. El proceso seguido en la experimentación fue el siguiente:

INICIO

Para cada consulta formulada desde el dominio local *Travel*.

El sistema calculó el índice de similitud de la consulta para cada dominio remoto

Se propagó la consulta a todos los dominios remotos y se recuperó información en cada dominio utilizando el módulo de recuperación de información de Onto-DOM

Se valoraron los resultados de las búsquedas remotas en función de los documentos recuperados, según las categorías definidas en Onto-DOM: *excelente (E)*, *Very Good (VG)*, *Good (G)*, *Regular (R)*, *Poor (P)* y *None (N)* cuando no se recuperaron documentos.

Se almacenó en la base de casos de Onto-DOM la consulta valorada como *excelente (E)* o *Very Good (VG)*. Esto permitió incrementar progresivamente la base de casos

Se compararon las valoraciones otorgadas a la información recuperada para la consulta en cada dominio remoto, con los respectivos valores de los índices de similitud.

FIN

3.3.2 Métricas

Para proceder a la comparación de resultados en términos cuantitativos se utilizó la métrica *accuracy*, utilizada previamente en la sección 3.1.2, la cual representa la exactitud de las sugerencias realizadas por el componente de propagación.

Para cada consulta q_j el sistema calcula el índice de similitud $SimQuery(dr, q_j)$ asociado a cada dominio d_r , el cual sirve de sugerencia al usuario para seleccionar el dominio al cual propagar la consulta. Se considera que dicha sugerencia es correcta si el dominio con mayor índice de similitud coincide con el dominio que el usuario ha valorado mejor luego de analizar los documentos recuperados.

3.3.3 Resultados de la comparación

En la Tabla IX.8 se muestran 26 de las consultas propagadas.

| ID_consulta | Consulta |
|-------------|--|
| q01 | Where can I find a caravan site? |
| q02 | Where can I visit medieval buildings? |
| q03 | Where can I do a mount resort reservation? |
| q04 | What Island can I do angling? |
| q05 | Where can I drink wine? |
| q06 | Where can I eat seafood in Africa? |
| q07 | Where can I visit a museum? |
| q08 | Where can I watch antelopes? |
| q09 | What city has famous gardens? |
| q10 | Where can I watch cheetahs? |
| q11 | What city can I visit famous castles? |
| q12 | Where can I do bird watching? |
| q13 | Where can I do a guided tour? |
| q14 | What river can I take rafting classes? |
| q15 | What city can I visit famous castles? |
| q16 | What city offers Unesco world heritage excursions? |
| q17 | Where can I do gorge swing? |

| | |
|------|---|
| q18 | What city has a cabaret? |
| q19 | What cities offer national park excursions? |
| q20 | Where can I visit a famous cathedral? |
| q21 | Where can I do a ballooning safari? |
| q22 | Where can I see famous paintings? |
| q23 | What city offers shipwreck excursions? |
| q24 | Where can I do a train excursion? |
| q25 | Where can I do tiger fishing? |
| q26 | Where can I do a day trip? |
| | |

Tabla IX.8. Consultas Propagadas

El índice de similitud calculado por el sistema, para cada consulta y dominio, se muestra en la Tabla IX.9. En dicha tabla se indican también los dominios mejor posicionados por el sistema, identificados como AT (*África Travel*) y ET (*Europa Travel*).

| q_j | $SimQuery(AT, q_j)$ | $SimQuery(ET, q_j)$ | Dominio mejor posicionado |
|-------|---------------------|---------------------|---------------------------|
| q01 | 0.000 | 0.000 | AT ET |
| q02 | 0.000 | 0.000 | AT ET |
| q03 | 0,163 | 0,450 | ET |
| q04 | 0.000 | 0.000 | AT ET |
| q05 | 0,050 | 0.000 | AT |
| q06 | 0,054 | 0.000 | AT |
| q07 | 0,109 | 0,450 | ET |
| q08 | 0,050 | 0,072 | ET |
| q09 | 0.000 | 0.000 | AT ET |
| q10 | 0,050 | 0,072 | ET |
| q11 | 0.000 | 1,800 | ET |
| q12 | 0,163 | 0,144 | AT |
| q13 | 0,100 | 0,072 | AT |
| q14 | 0,050 | 0,072 | ET |
| q15 | 0.000 | 1,800 | ET |
| q16 | 0,050 | 0,072 | ET |
| q17 | 0,05 | 0,16363 | ET |

| | | | |
|------|-------|----------|------|
| q18 | 0,05 | 0,2 | ET |
| q19 | 0,05 | 0,072 | ET |
| q20 | 0,05 | 0,072 | ET |
| q21 | 0,1 | 0,085714 | AT |
| q22 | 0,3 | 0,45 | ET |
| q23 | 0,3 | 0,654545 | ET |
| q24 | 0,072 | 0,189474 | ET |
| q25 | 0 | 0,41667 | ET |
| q26 | 1,8 | 0,072 | AT |
| | | | |

Tabla IX.9. Índices de Similitud

En la Tabla IX.10 se muestran las valoraciones asignadas a los resultados obtenidos del proceso de recuperación de información en cada dominio y su comparación con las sugerencias propuestas por el sistema.

| q _i | Valoración información recuperada de AT | Valoración información recuperada de ET | Dominio mejor valorado | Dominio mejor posicionado por sistema | Resultado de la sugerencia |
|----------------|---|---|------------------------|---------------------------------------|----------------------------|
| q01 | E | P | AT | AT ET | correcta |
| q02 | N | E | ET | AT ET | correcta |
| q03 | E | P | AT | ET | incorrecta |
| q04 | E | P | AT | AT ET | correcta |
| q05 | E | N | AT | AT | correcta |
| q06 | E | P | AT | AT | correcta |
| q07 | G | E | ET | ET | correcta |
| q08 | E | N | AT | ET | incorrecta |
| q09 | E | E | AT ET | AT ET | correcta |
| q10 | E | N | AT | ET | incorrecta |
| q11 | P | E | ET | ET | correcta |
| q12 | E | P | AT | AT | correcta |
| q13 | E | P | AT | AT | correcta |
| q14 | E | P | AT | ET | incorrecta |
| q15 | G | E | ET | ET | correcta |
| q16 | P | E | ET | ET | correcta |
| q17 | E | P | AT | ET | incorrecta |
| q18 | P | E | ET | ET | correcta |
| q19 | E | P | AT | ET | incorrecta |
| q20 | E | E | AT ET | ET | correcta |
| q21 | E | P | AT | AT | correcta |
| q22 | N | E | ET | ET | correcta |

| | | | | | |
|------|------|------|------|------|------------|
| q23 | E | P | AT | ET | incorrecta |
| q24 | P | E | ET | ET | correcta |
| q25 | VG | N | AT | ET | incorrecta |
| q26 | E | P | AT | AT | correcta |
| | | | | | |

Tabla IX.10. Resultados de la Propagación Selectiva

Los valores de la *accuracy* se muestran en la Tabla IX.11 y en la Figura IX.10. Puede observarse que el sistema tiende a realizar sugerencias correctas en aproximadamente el 70% de los casos.

| q _j | Cantidad de sugerencias correctas acumuladas | <i>Accuracy</i> |
|----------------|--|-----------------|
| q01 | 1 | 1.000 |
| q02 | 2 | 1.000 |
| q03 | 2 | 0,666 |
| q04 | 3 | 0,750 |
| q05 | 4 | 0,800 |
| q06 | 5 | 0,833 |
| q07 | 6 | 0,857 |
| q08 | 6 | 0,750 |
| q09 | 7 | 0,777 |
| q10 | 7 | 0,700 |
| q11 | 8 | 0,727 |
| q12 | 9 | 0,750 |
| q13 | 10 | 0,769 |
| q14 | 10 | 0,714 |
| q15 | 11 | 0,733 |
| q16 | 12 | 0,750 |
| q17 | 12 | 0,705 |
| q18 | 13 | 0,722 |
| q19 | 13 | 0,684 |
| q20 | 14 | 0,700 |
| q21 | 15 | 0,714 |
| q22 | 16 | 0,727 |
| q23 | 16 | 0,695 |
| q24 | 17 | 0,708 |
| q25 | 18 | 0,720 |
| q26 | 18 | 0,692 |
| | | |

Tabla IX.11. Valores Promedios de la *Accuracy* para la Propagación Selectiva

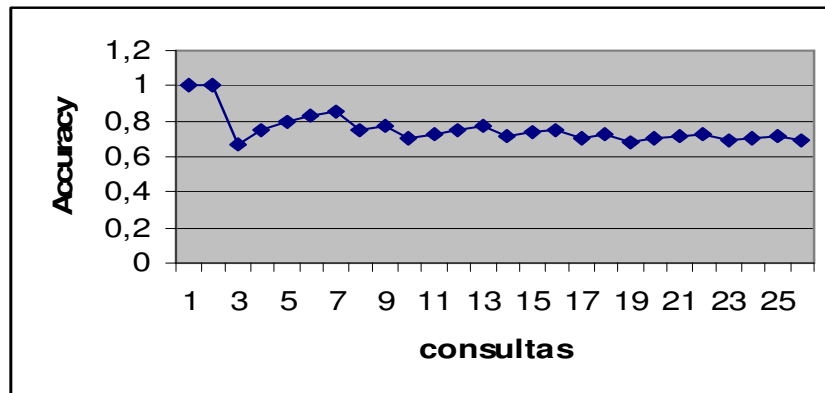


Figura IX.10. Evolución de los Valores de la *accuracy* para la Propagación Selectiva

4. Conclusiones

El análisis de desempeño de los componentes de Onto-DOM mediante un caso de estudio, permitió validar de manera empírica la estrategia de anotado (Algoritmo VII.1, Capítulo VII); la estrategia de procesamiento de la consulta y recuperación de información (Algoritmo VII.2, Capítulo VII); y finalmente, la estrategia de propagación selectiva (Algoritmo VII.3, Capítulo VII); propuestas como aporte de esta tesis.

Para cada una de las estrategias a evaluar se diseñó un procedimiento que permitiese comparar los resultados de dichas estrategias con resultados confiables. Para generar estos últimos, en los dos primeros procedimientos se recurrió a procesamiento manual, y en el tercero a una propagación exhaustiva (a todos los dominios).

Abdou y col. (2008) presentan un modelo probabilístico de recuperación de información que denominan $I(n)B2$, y comparan el desempeño de dicho modelo con otros diez modelos diferentes de recuperación de información. Presentan un valor del $MAP = 0.397$ obtenido para dicho modelo de una valoración empírica, indicando una mejora del 170% en relación al valor del MAP obtenido para el clásico modelo de recuperación de información basado en la frecuencia de términos en los documentos. En la sección 3.2, el desempeño del

módulo de recuperación de información, que implementa la estrategia basada en ontologías propuesta en esta tesis, obtuvo un MAP = 0.860, lo cual indica una mejora sustancial del proceso de recuperación de información. Esta mejora obtenida es consecuencia del efecto combinado de la estrategia de anotado y de la estrategia de procesamiento de la consulta y recuperación de información propuestas en esta tesis.

En relación a la propagación de la consulta a otros dominios, desde el punto de vista de la recuperación de información, la propagación completa es la mejor alternativa cuando el número de dominios es bajo. Esta estrategia puede no ser muy conveniente para el caso de una organización con gran cantidad de dominios, porque podría producir una excesiva carga de trabajo, por lo cual, una estrategia de propagación selectiva podría ser más apropiada para este caso. En esta tesis se ha propuesto una estrategia de propagación selectiva, con la cual se han obtenido buenos resultados (sección 3.3) de recuperación de información. No obstante, es posible seguir mejorando dicha estrategia, quedando esto como línea de trabajo futuro.

Conclusiones y Trabajos Futuros

1. Introducción

El objetivo de este capítulo es resumir las principales contribuciones del trabajo realizado en el marco de la presente tesis. Dicho resumen se presenta en la sección 2. En la sección 3, se detallan algunos aspectos relevantes de la tesis que constituyen puntos de partida para el desarrollo de trabajos futuros.

2. Principales Contribuciones

La necesidad de gestionar el conocimiento organizacional es innegable, sin embargo, en los últimos años la Gestión del Conocimiento ha atravesado por una serie de cuestionamientos debido al excesivo énfasis en soluciones basadas exclusivamente en tecnologías de información o la proliferación de modelos que no ofrecen el marco teórico apropiado para implementar iniciativas de Gestión de Conocimiento o el desarrollo de aplicaciones prácticas. El desafío actual es trascender estos problemas o visiones limitantes para integrar lo social con lo tecnológico. Es por ello que, en esta tesis se han hecho contribuciones no sólo desde el punto de vista de los sistemas de información sino también desde el aspecto social de la Gestión del Conocimiento, definiendo un modelo conceptual

que sirve de guía para implementar una iniciativa de Gestión de Conocimiento en una organización; para evaluar la situación actual, con relación al tema, en un momento dado de la vida organizacional; o como marco de referencia para el desarrollo de tecnologías de información.

2.1. Modelo Conceptual de Gestión del Conocimiento Distribuido

Uno de los aportes de esta tesis, desde el punto de vista de la Gestión del Conocimiento como disciplina organizacional, es el resultado del análisis de la situación actual de la Gestión del Conocimiento, realizado a lo largo de los tres primeros capítulos, que ha identificado un conjunto de requisitos que un modelo conceptual de gestión del conocimiento organizacional debería satisfacer para constituirse en marco de referencia para una implementación de Gestión de Conocimiento y para el desarrollo de tecnologías de información. Dichos requisitos fueron presentados como aporte de esta tesis en el Capítulo V.

Estos requisitos fueron utilizados, para analizar el alcance de los modelos conceptuales presentes en la literatura, presentados en el Capítulo III, definidos para dar soporte a la Gestión del Conocimiento organizacional. Del análisis se desprendieron las dimensiones o aspectos que cada modelo enfatiza o contempla, observándose que cada uno de ellos se enfoca en ciertos elementos de la Gestión del Conocimiento pero ninguno los resume a todos, limitando de esta forma el alcance de los mismos.

Se identificó entonces como problemática concreta, en el Capítulo V, la falta de un modelo conceptual unificado y más abarcativo que satisfaga todos los requisitos identificados que sirva como marco de referencia para las iniciativas de Gestión de Conocimiento y de desarrollo de tecnologías de información para su implementación concreta.

Con el propósito de contribuir a la solución de esta problemática, en el Capítulo VI se ha propuesto un Modelo Conceptual de Gestión del Conocimiento Distribuido organizacional que verifica todos los requisitos identificados. Este

modelo fue generado tomando en consideración las limitaciones de otros modelos y las críticas más comunes asociadas al fracaso de las iniciativas de Gestión del Conocimiento sintetizadas en el Capítulo V. Dicho modelo propone una integración entre los aspectos tecnológicos y sociales de la Gestión del Conocimiento vinculando las actividades, comúnmente asociadas con dicha gestión, con otros procesos organizacionales de naturaleza altamente social, como son la enseñanza, el aprendizaje, la tutoría y el entrenamiento. De este modo, el modelo propone mover el foco desde un excesivo énfasis en las tecnologías de información hacia la estructura social organizativa compuesta por comunidades de práctica y conocimiento que hace posible la creación, uso y distribución del conocimiento. En este sentido, la aplicación del modelo propuesto conlleva la identificación de dominios de conocimiento dentro de la organización (junto con las comunidades de práctica y conocimiento que los componen) y su conexión en red potenciando de esta forma los procesos organizacionales relacionados con la Gestión del Conocimiento. Para ello, el modelo se sustenta en un *Sistema de Gestión del Conocimiento*, que da soporte al proceso de creación de conocimiento, a través de cuatro procesos de conversión entre comunidades configuradas en una red de conocimiento; y un *Sistema de Memoria Organizacional Distribuida* que da soporte a la red de conocimiento que ayuda a canalizar los esfuerzos de los trabajadores.

El modelo conceptual propuesto sirve entonces como marco de referencia para las iniciativas de Gestión de Conocimiento y de desarrollo de tecnologías de información. Para este último propósito, el *Sistema de Memoria Organizacional Distribuida* constituye la parte implementable del modelo conceptual.

Definido el modelo conceptual, la tesis se centró en el desarrollo de una propuesta de una arquitectura para el *Sistema de Memoria Organizacional Distribuida*, su implementación prototipo y su validación empírica. Quedó fuera del alcance de esta tesis la implementación del *Sistema de Gestión del Conocimiento*.

2.2. Sistema de Memoria Organizacional Distribuida

Utilizando como marco de referencia el modelo conceptual propuesto en el Capítulo VI, en el Capítulo VII se presentó, como otro aporte de esta tesis, una arquitectura para el *Sistema de Memoria Organizacional Distribuida* y se describieron con detalle las estrategias desarrolladas para la implementación de la funcionalidad de cada uno de los componentes de la arquitectura.

La arquitectura de tres componentes propuesta para una Memoria Organizacional Distribuida aborda dos problemas comunes en las implementaciones de repositorios de estas características: la sobrecarga de trabajo que implica para los trabajadores elicitar su conocimiento para contribuir al repositorio y la falta de contexto asociada al proceso de explicitar conocimiento para formalizarlo.

Para el componente de Representación del Conocimiento se propuso una estrategia de anotado automático basada en una ontología de dominio que realiza expansión semántica basada en un diccionario (Algoritmo VII.1). Como resultado, obtiene un conjunto de descriptores del documento analizado que constituye la porción del modelo del dominio (ontología) que mejor representa el documento.

Para el componente de Procesamiento y Recuperación de Información se propuso un algoritmo que realiza un tratamiento de la consulta del usuario en lenguaje natural expandiendo la estrategia de anotado anterior con un análisis que aprovecha el poder representativo de las relaciones ontológicas junto con otras técnicas tendientes a comprender la consulta realizada (Algoritmo VII.2).

Finalmente, en el componente de Interfaz a Otros Dominios, para el caso de requerirse una propagación selectiva, se propone una estrategia de aprendizaje basado en casos que permite propagar de manera selectiva la consulta del usuario a otros dominios (Algoritmo VII.3).

Las principales características del modelo de repositorio generado por esta arquitectura, y que lo diferencian de otras propuestas, son las siguientes:

- ✓ Se presenta un esquema distribuido que posibilita el soporte de una red de conocimiento conformada por los dominios de conocimiento dentro de la organización (que aglutinan las comunidades de práctica y de conocimiento) evitando los costos de mantenimiento de un repositorio centralizado y respetando la natural distribución del conocimiento.
- ✓ Se provee contexto para los objetos de conocimiento incorporados en la Memoria Organizacional a través de una representación de los mismos basada en un modelado ontológico.
- ✓ A través del desarrollo de una estrategia de anotado propia se posibilita la incorporación automática de los objetos de conocimiento a la Memoria Organizacional, liberando de esta forma a los trabajadores de la sobrecarga de trabajo que implica elicitarse su conocimiento para contribuir a la misma.
- ✓ Se proporciona al usuario una interacción amigable y transparente con la Memoria Organizacional aceptando requerimientos en lenguaje natural. Asimismo, se oculta la complejidad de la estructura distribuida automatizando la propagación de consultas a través de un mecanismo de propagación total, o bien mediante uno de propagación selectiva que sugiere al usuario los posibles destinos de propagación de su consulta en base a un mecanismo de aprendizaje basado en casos.

2.3. Onto-DOM: Prototipo de Memoria Organizacional

Otra contribución importante, desde el punto de vista de los sistemas de información, es el desarrollo del prototipo Onto-DOM, presentado en el Capítulo VIII, que implementa la arquitectura de Memoria Organizacional Distribuida propuesta en el Capítulo VII. Dicho sistema, implementado utilizando software libre, presenta las características principales siguientes:

- ✓ Permite el anotado automático de fuentes de conocimiento heterogéneas o no estructuradas en función de un modelo de dominio

- ✓ Permite el uso de lenguaje natural para las consultas haciendo innecesario el aprendizaje de cualquier formalismo
- ✓ Permite el uso de fuentes de conocimiento de otros dominios mediante la propagación de la consulta a otras memorias organizacionales

Onto-DOM implementa la capa de la arquitectura que contiene la lógica del sistema a través del componente de representación de conocimiento (KRL), el componente de recuperación de información (IRPL) y el componente de propagación a otros dominios (ODI). El componente de representación de conocimiento implementa la estrategia de anotado propuesta como aporte de esta tesis (Algoritmo VII.1, Capítulo VII); el componente de recuperación de información implementa la estrategia de procesamiento de la consulta y recuperación de información propuesta como aporte de esta tesis (Algoritmo VII.2, Capítulo VII); y finalmente, el componente de propagación a otros dominios implementa las estrategias de propagación completa y de propagación selectiva propuestas como aporte de esta tesis (Algoritmo VII.3, Capítulo VII).

En el Capítulo IX se presentó un análisis de desempeño de cada uno de los componentes del sistema OntoDOM, mediante un caso de estudio, como proceso de validación empírica (Recker, 2005) de las estrategias (Algoritmos) propuestos como aporte de esta tesis en el Capítulo VII.

3. Trabajos Futuros

Los trabajos futuros de investigación a seguir, a partir de la presente tesis, se enmarcan en los puntos que se detallan a continuación:

Mejorar el prototipo en los siguientes aspectos: utilización de los resultados del proceso de anotado para la sugerencia de cambios (adición/remoción de conceptos) en la ontología de dominio; desarrollo de heurísticas que mejoren el proceso de análisis de la consulta del usuario; y mejoramiento de la estrategia de propagación de la consulta.

Implementar el sistema OntoDOM mejorado a nuevos casos de estudio a efectos de realizar nuevas validaciones empíricas del mismo.

Avanzar en la implementación del *Sistema de Gestión del Conocimiento*, que quedó fuera del alcance de esta tesis, en un caso de estudio real a efectos de poder realizar una validación empírica del mismo.

Habiendo reconocido la importancia del aspecto social de la Gestión del Conocimiento incorporando los procesos de enseñanza, aprendizaje, entrenamiento y tutoría, queda por analizar la inclusión de otros procesos sociales que se dan dentro de la organización. Los mismos surgen de procesos de conversión entre formas intermedias de conocimiento y por el natural desarrollo de las actividades diarias en la organización.

Componente de Representación de Conocimiento

1. Proceso de Etiquetado

Como se observa en el Figura A.1, el método central del proceso es *CRIdentification(urlDoc)* el cual implementa la estrategia de anotado descrita anteriormente, por cada documento devuelve una lista de palabras, las clases obtenidas y una profundidad asociada al proceso de búsqueda de las mismas. La profundidad devuelta por la búsqueda está relacionada a una herramienta, de esta manera hay una profundidad asociada la ontología (isA) y una asociada a la búsqueda realizada en WordNet (sinonimia, hiperonimia).

Con respecto específicamente al etiquetado sintáctico, el primer paso es la creación de una instancia de *ontodom.DocImpl* para contener la información resultado del procesamiento. Esta clase, junto con otras que se describirán más adelante, constituyen la unidad de persistencia de la herramienta. Cada instancia es identificable por medio de la instancia de *ontodom.OntologyImpl* y la URL del documento en procesamiento (1.1).

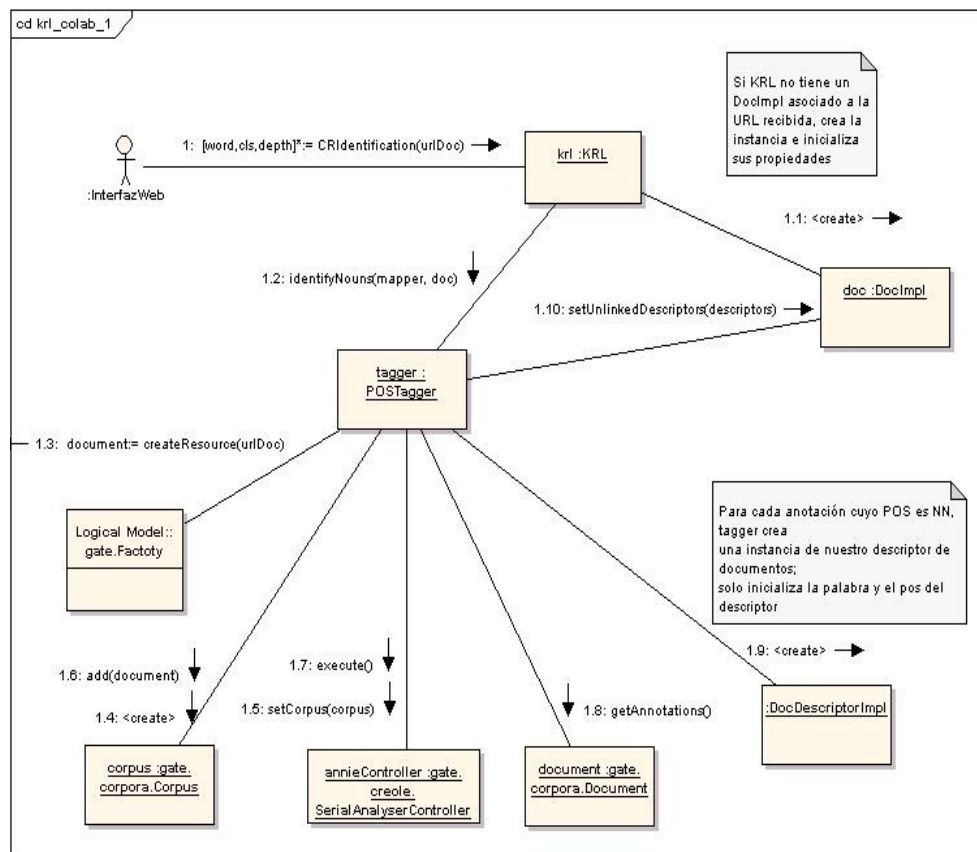


Figura A.1. Proceso de Etiquetado

El siguiente paso es la realización del etiquetado para obtener los sustantivos del documento, responsabilidad de *ontodom.POSTagger* que utiliza el “plugin” ANNIE y la instancia de mapeo (1.2). Se utiliza luego la URL del documento para crear una instancia de la clase *Document* (*gate.corpora.Document*) (1.3) y se crea el *Corpus* (*gate.corpora.Corpora*) a procesar (1.4).

La instancia *annieController* (*gate.creole.SerialAnalyserController*) es inicializada con dicho corpus (1.5) añadiéndole a continuación el *Document* al *Corpus* (1.6). Se solicita luego a *annieController* la ejecución del procesamiento de etiquetado, lo que implica la ejecución de los recursos de procesamiento (RP) sobre el documento (1.7). Dicha ejecución almacena su resultado en la instancia de *Document*, y en este punto se solicita a la misma el conjunto de etiquetas

obtenidas por ANNIE. El conjunto de etiquetas está tipificado mediante siglas NN, NNP, VB, etc (1.8).

Para cada etiqueta cuyo nombre tipo comienza con NN se crea un descriptor (*ontodom.DocDescriptorImpl*), tanto su lema (sustantivo), como su POS son inicializados en este punto. El lema o la palabra del documento que da razón al descriptor es formateada para poder comparar libremente entre las distintas herramientas, este paso se realiza utilizando la instancia de mapeo y ha sido omitido en la figura por simplicidad (1.9). Estos descriptores aún no han sido vinculados al dominio ya que todavía no se ha encontrado relación con la ontología de dominio. El conjunto de estos descriptores es incorporado a la lista de descriptores no enlazados del documento (1.10).

La instancia *tagger* utiliza, para el etiquetado de términos con nombres que contengan más de una palabra (sustantivos compuestos), el método *identifyPOS*. El conjunto de etiquetas es un grafo, *identifyPOS* recorre dicho grafo, en primer lugar, observa si el “token” leído concuerda con el prefijo de alguna clase cuyo nombre contenga más de una palabra, (esto se hace con el fin de recuperar del texto aquellos sustantivos del dominio que estén formados por más de una palabra), esta función se cumple en colaboración con la instancia de mapeo *mapper*. Si el primer paso falla, se guarda la palabra observada sólo si el POS de la misma comienza con NN, es decir, si se trata de un sustantivo.

2. Proceso de Búsqueda de Ocurrencias Exactas

Como se mencionó con anterioridad, el proceso de anotado comienza con la búsqueda de ocurrencias de clases ontológicas en el texto. En la Figura A.2 se muestra la colaboración que resuelve dicha operación.

La instancia de KRL solicita a la instancia de mapeo la búsqueda de ocurrencias exactas en el documento en proceso (1). En primer lugar se necesitan los descriptores hallados en el paso anterior, estos se solicitan a la instancia de *ontodom.DocImpl*, y como ya se mencionó, estos descriptores aún no están

enlazados al dominio. En este punto *mapper* comienza a iterar sobre la colección de descriptores (1.1).

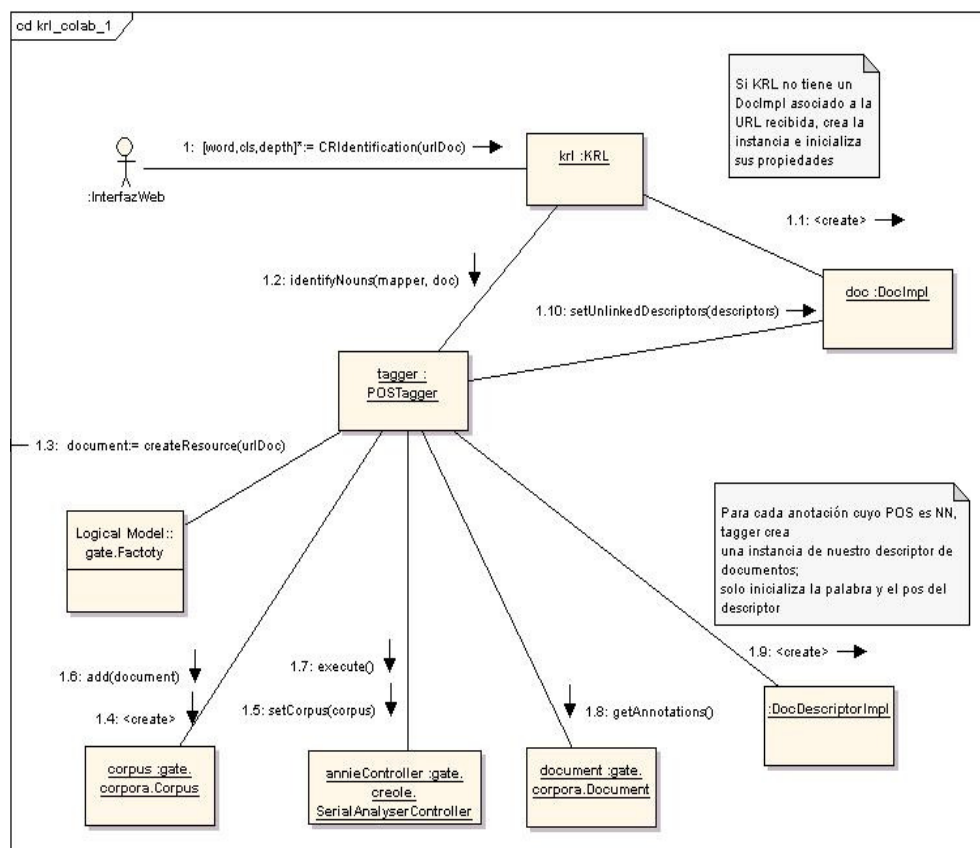


Figura A.2. Proceso de Búsqueda de Ocurrencias Exactas

Al descriptor observado en la iteración, se le solicita su lema (sustantivo) (1.2) y se compara dicho lema con el conjunto de nombres de clases formateados para comparación (1.3). Si concuerda, se recupera el nombre real de la clase ontológica (1.4) y como se ha encontrado una ocurrencia exacta, se da valor a la propiedad *cls* del descriptor (1.5).

A continuación se dan valores a las profundidades de búsqueda. En cuanto a la profundidad asociada a la búsqueda en la ontología, como no se ha subido ningún nivel taxonómico, dicho valor es cero (1.6) al igual que el valor de la profundidad de búsqueda en WordNet que tampoco se ha utilizado (1.7). Terminada la iteración, el conjunto de descriptores vinculados al dominio es

restado del conjunto de descriptores no vinculados (1.8) y se usa para iniciar el conjunto de descriptores vinculados al dominio que es propiedad de la instancia *doc* (1.9).

3. Proceso de Expansión Semántica

Como se observa en la Figura A.3, agotada la búsqueda de ocurrencias exactas, el proceso de anotado continua utilizando WordNet como herramienta de expansión semántica. La instancia de KRL solicita la búsqueda de expansiones semánticas sobre el documento que está siendo procesado (1).

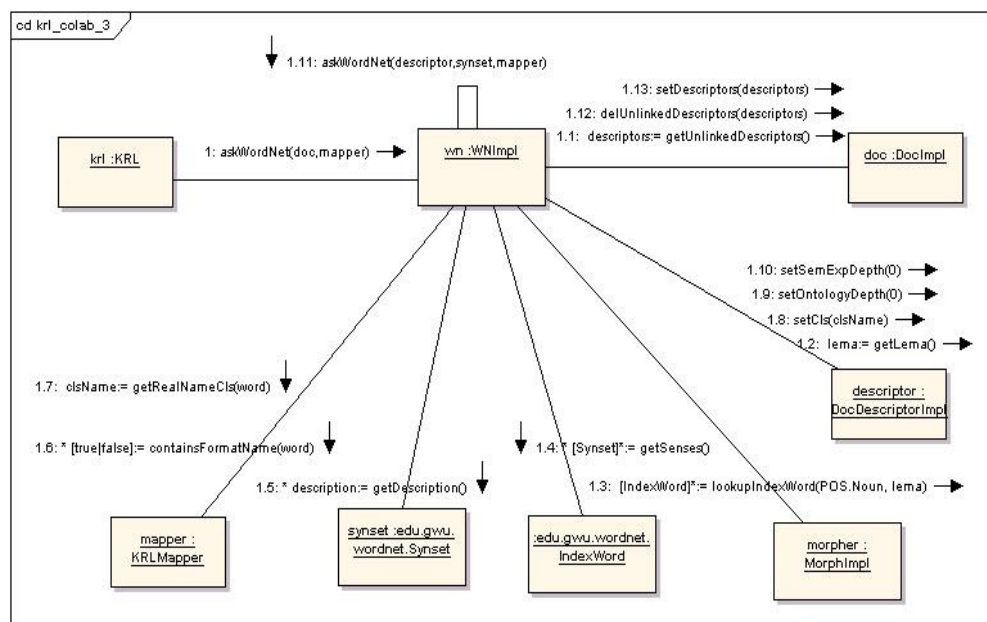


Figura A.3. Proceso de Expansión Semántica

La instancia *wn* (*ontodom.WNImpl*) pide al documento el conjunto de descriptores que aún no han sido vinculados al dominio y el proceso de expansión comienza a iterar sobre estos descriptores (1.1). Para el descriptor en observación, se solicita el lema o sustantivo (1.2). El lema es utilizado para obtener los *IndexWords* asociados a cada sustantivo, y se informa a la librería (*JWordNet*) que se está buscando un sustantivo (*POS.NOUN*). Se utiliza una

instancia propia denominada *morpher* responsable de la recuperación de la información de WordNet y del tratamiento de la morfología de las palabras (1.3).

Morpher solicita los *IndexWords* asociados a la palabra y el POS solicitados utilizando una instancia de *DictionaryDatabase* de la librería *JWordNet*. Si esta búsqueda fracasa, se repite buscando excepciones, esto es, formas básicas de la palabra en cuestión para el POS dado. Esto lo hace solicitando a la misma instancia de *DictionaryDatabase* que busque una forma básica de la palabra.

Si se encontró una forma base de la palabra, se repite la búsqueda usando esa forma básica y si la búsqueda continua siendo infructuosa se hace un último intento en base a la morfología de la palabra. En la Tabla A.1 se listan las reglas definidas por WordNet (WordNet, 2006):

| POS | Suffix | Ending | POS | Suffix | Ending | POS | Suffix | Ending |
|------|--------|--------|------|--------|--------|-----|--------|--------|
| NOUN | "s" | "" | VERB | "s" | "" | ADJ | "er" | "" |
| NOUN | "ses" | "s" | VERB | "ies" | "y" | ADJ | "est" | "" |
| NOUN | "xes" | "x" | VERB | "es" | "e" | ADJ | "er" | "e" |
| NOUN | "zes" | "z" | VERB | "es" | "" | ADJ | "est" | "e" |
| NOUN | "ches" | "ch" | VERB | "ed" | "e" | | | |
| NOUN | "shes" | "sh" | VERB | "ed" | "" | | | |
| NOUN | "men" | "man" | VERB | "ing" | "e" | | | |
| NOUN | "ies" | "y" | VERB | "ing" | "" | | | |

Tabla A.1. Reglas Morfológicas de WordNet

La palabra resultante de la aplicación de estas reglas es utilizada para buscar los *IndexWords* en el diccionario.

A partir de este punto, se hablará de palabra y no de sustantivo, el hecho de que se trata de un sustantivo es determinado por el parámetro POS, de este modo los métodos mencionados en adelante pueden ser utilizados por IRPL en el tratamiento tanto de verbos como de sustantivos.

A cada instancia de *edu.gwu.wordnet.IndexWord* hallada se le solicita el conjunto de sentidos que define la palabra. Este conjunto de sentidos son instancias de *edu.gwu.wordnet.Synset* (*JWordNet*). En WordNet cada *Synset* es

un conjunto de sinónimos que define un concepto y que tiene punteros semánticos a otros *Synsets* (hiperonimia, hiponimia, antonimia) (1.4). A cada *synset* se le solicita su descripción (1.5). Dicha descripción se trata del conjunto de sinónimos que definen el sentido y no del texto que define ese sentido. Dicho texto es también utilizado pero por IRPL y se tratará más adelante.

Para cada sinónimo, se pregunta a *mapper* si existe una clase de nombre equivalente a ese sinónimo (1.6) y si se encuentra el sinónimo dentro de los nombres de clases formateados, se pide a *mapper* el nombre real de la clase ontológica hallada por WordNet. En este caso se ha encontrado una instancia ontológica por relación semántica de sinonimia utilizando WordNet (1.7) y se guarda en el descriptor el nombre de la clase hallada (1.8).

Como aún no se ha expandido la búsqueda mediante la taxonomía ontológica, permanece el valor cero en la profundidad de búsqueda asociada (1.9) y como se trata de sinonimia, también se da el valor cero a la profundidad de búsqueda en WordNet (1.10). En este punto, si para un *synset* dado no se ha tenido éxito, la búsqueda procede por el puntero de hiperonimia del *synset* (1.11).

Si al procesar un *synset* no se logran resultados, se buscan más términos ontológicos a través de los punteros de *hiperonimia* e *instancia de hiperonimia* definidos en la instancia de *Synset* en proceso. El proceso sube por las relaciones de hiperonimia siempre dentro del mismo sentido, no se vuelve a entrar a los índices (*IndexWords*). En cada lectura del puntero se recupera un conjunto de *targets*, estos son otro conjunto de *Synsets* y su tratamiento es el mismo por lo que se repiten los pasos desde 1.5 a 1.11 para cada *synset* o *hyper synset*. La única diferencia es que el anotado va incrementando el valor de la profundidad en cada lectura de los punteros de hiperonimia. Este proceso de escalado continua hasta agotar la información en WordNet para las relaciones halladas a partir del *Synset* original y sus padres.

Finalizada esta serie de iteraciones (descriptores, *index words*, *synsets*, palabras, *synsets* hallados por hiperonimia), el conjunto de descriptores

exitosamente enlazados al dominio es eliminado del conjunto de descriptores que no lo han sido (1.12). Estos descriptores obtenidos se incorporan a la colección de descriptores vinculados al dominio que es propiedad del emulador de documento (1.13).

4. Proceso de Inclusión de Superclases

Debido a la necesidad de construir una base de datos que permita acceder a los documentos a partir de la estructura ontológica, son necesarias no sólo las clases halladas, sino también las superclases de las mismas y es en este paso donde comienza la colaboración para recuperar las superclases de cada clase hallada. Cada nivel de la relación *isA* incrementa un contador, *depth*, que se usa para dar información de la profundidad de la relación respecto al término hallado. Esto ofrece un dato importante para el responsable del anotado que le permite decidir si las superclases se añaden a la base de conocimientos como descriptores de un documento en particular.

Básicamente, como se muestra en la Figura A.4, en este proceso se solicita a la instancia de modelo (*ontodom.Model*) la inclusión como descriptores del documento a las superclases de las clases halladas hasta el momento (1). El primer paso es recuperar el conjunto de descriptores vinculados al dominio. A continuación *model* comienza a iterar sobre dicho conjunto añadiendo descriptores a medida que sube por la jerarquía ontológica (1.1). A cada descriptor se solicita el lema (1.2) y el nombre de la clase ontológica (1.3).

Utilizando el modelo (*edu.stanford.smi.protege.model.OWLModel*) y el nombre de clase, se recupera una instancia de *edu.stanford.smi.protege.model.Cls* que representa una *OWLNamedClass* (1.4). A esta clase se le solicitan sus superclases (1.5) y se le pide su nombre (1.6). A cada superclase hallada corresponde un nuevo descriptor (1.7). Dicho descriptor tendrá por lema el mismo que el del descriptor original y el nombre de su clase será el nombre de la superclase (1.8).

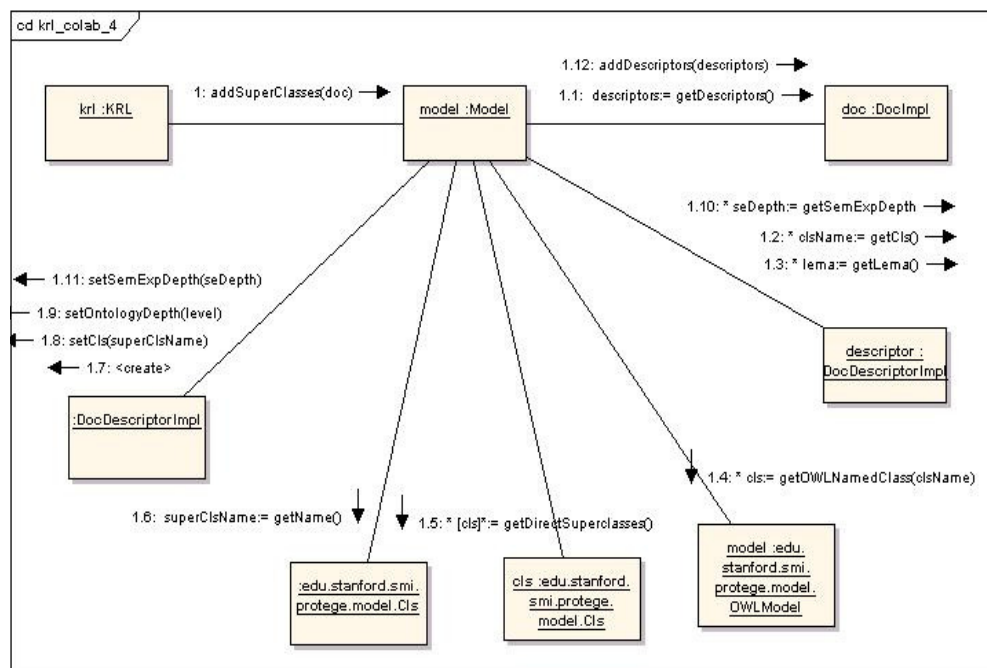


Figura A.4. Proceso de Inclusión de Superclases

El proceso sube por la jerarquía taxonómica, y en este punto el nivel de pasos ascendidos es guardado dentro del descriptor (1.9). Se recupera la profundidad de búsqueda por expansión semántica del descriptor original (1.10) y el valor de dicha profundidad para el nuevo descriptor es el mismo (1.11). Finalizado el proceso, los nuevos descriptores son agregados al conjunto de descriptores vinculados al dominio (1.12).

5. Persistencia

En esta colaboración se almacenan en base de datos los documentos y los descriptores seleccionados por el usuario. Como se muestra en la Figura A.5, la instancia de *ontodom.KRL* recibe desde la interfaz la URL del documento que el usuario desea almacenar junto con una lista de palabras. Esta lista tendrá dos palabras por cada descriptor que se desea persistir. Estas palabras son el lema (sustantivo) y el nombre de la clase (1).

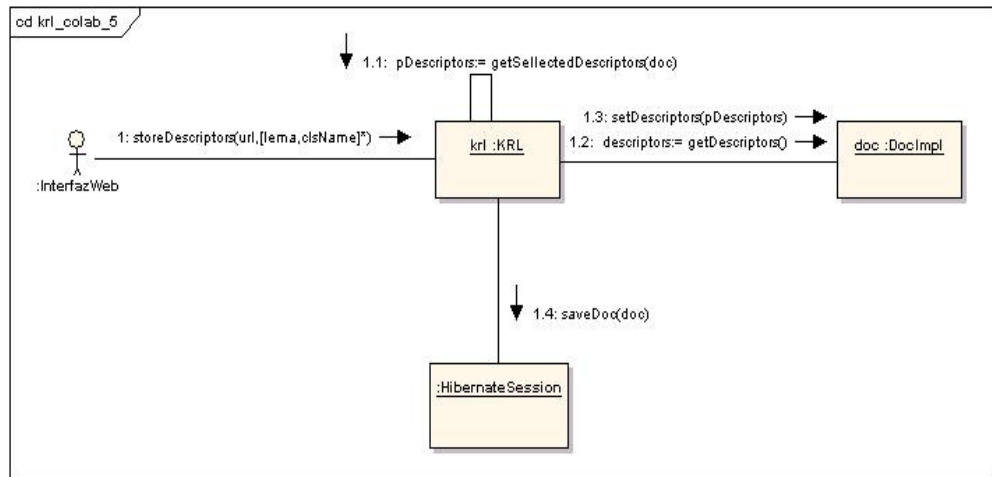


Figura A.5. Persistencia

A continuación se crea una lista con los descriptores a almacenar. Para esto se itera sobre las palabras recibidas desde la interfaz. Cada par de palabras han de pertenecer a uno de los descriptores incluidos en el conjunto de descriptores vinculados al dominio pertenecientes al documento cuya URL coincida con la URL recibida desde la interfaz. Cada descriptor cuyo lema y nombre de clase sean hallados entre las palabras recibidas será incluido en la lista de descriptores a almacenar (1.1) (1.2). La lista de descriptores a almacenar es asignada a los descriptores del documento (1.3) y se utiliza el motor de Hibernate para persistir el documento (1.4).

Componente de Procesamiento y Recuperación de Información

1. Proceso General

Como se muestra en la Figura B.1, el proceso comienza con la petición del usuario por medio de una consulta en lenguaje natural, la cual es transportada por la interfaz web. Esta requiere una instancia de *ontodom.Query*, construida por la instancia *irpl* (*ontodom.IRPL*) del dominio consultado (1).

Inmediatamente, se crean instancias de *ontodom.ODQueryImpl*, la cual es persistente y contendrá información relevante a la descripción de la consulta a partir de la ontología de dominio (1.1) y de *ontodom.Query*. Este objeto es responsable de la recuperación de información a partir de la inicialización que *ontodom.IRPL* hace con sus propiedades (1.2).

La instancia *irpl* realiza el correspondiente análisis lingüístico de la consulta. Guarda información del POS de las palabras de la consulta en *query* (información de procesamiento) y en *odquery* (información de representación), y, además, guarda información sobre el significado de los verbos que contiene la consulta (1.3 – 1.4) (Ver sección 2 - Proceso de Análisis Lingüístico)

En este paso, *irpl* identifica las relaciones de las clases ontológicas vinculadas a la consulta guardando información del procesamiento en *query* e información de representación en *odquery* (1.5).

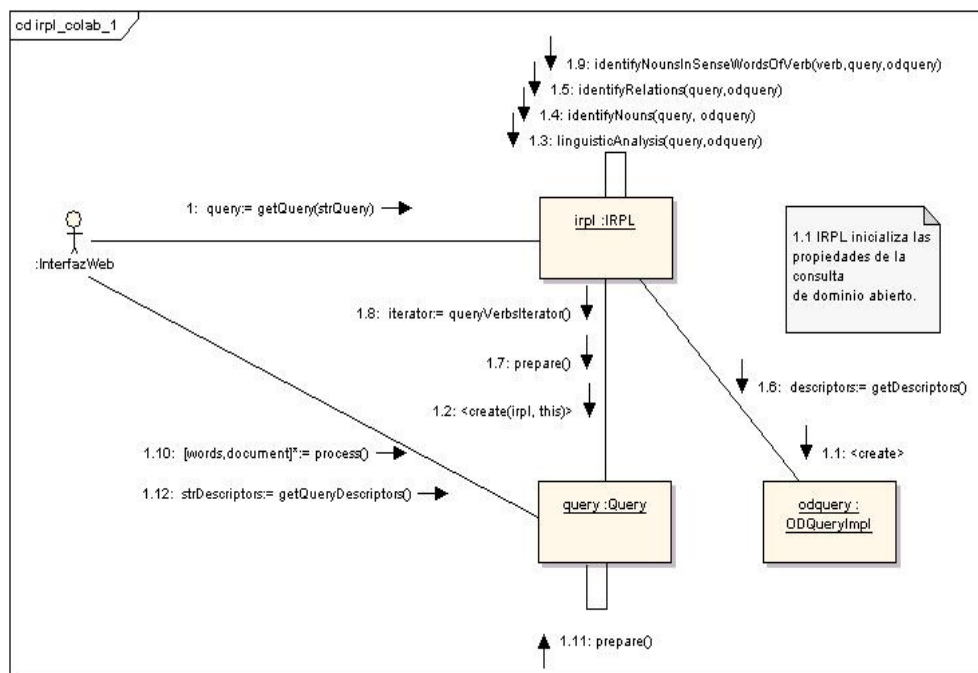


Figura B.1. Proceso General de Recuperación de Información

Una vez realizado esto, se verifica si se tuvo éxito y existen descriptores vinculados al dominio guardados dentro de *odquery* (1.6), si no es así, se realiza la preparación de la consulta (1.7) (ver detalle en sección 6 – Proceso de Preparación de la Consulta). Este paso es normalmente realizado antes del procesamiento de la misma y se ejecuta a esta altura del proceso debido a la ausencia de descriptores.

A continuación del paso 1.7 se comienza a iterar sobre los verbos de la consulta aplicando la estrategia sobre la información que WordNet da para cada verbo (1.8) (Ver detalle en sección 5 – Identificación de Sustantivos de Definición de un Verbo)).

En cada paso de la iteración, se usa el verbo en cuestión para obtener las palabras de su definición en WordNet. Se obtienen los sustantivos contenidos en dicha definición y se realiza la búsqueda de ocurrencias exactas para terminar con búsqueda de sinónimos usando WordNet (siguiendo un proceso similar al usado en el anotado de documentos). La información de procesamiento obtenida

es puesta en *query* mientras que la información de representación es puesta en *odquery*. El proceso vuelve al paso 1.4 (*identifyNouns*), esta vez con más información obtenida por medio del significado que WordNet da al verbo de la consulta (1.9).

Finalmente, la interfaz web solicita al objeto *query* el conjunto de documentos que han sido etiquetados con los descriptores hallados (1.10). Como paso inicial del procesamiento de información, se realiza la preparación de la consulta, este paso inicializa variables para la ejecución del proceso de recuperación (1.11) (ver detalle en sección 6 – Proceso de Preparación de la Consulta) y adicionalmente, los descriptores pueden ser solicitados por la interfaz web para ser mostrados al usuario (1.12).

2. Proceso de Análisis Lingüístico

Como se observa en la figura B.2, el análisis lingüístico es solicitado al objeto *tagger* (*ontodom.POSTagger*), este análisis (descrito con más detalle en la sección 2.1. – Análisis Lingüístico) - asigna un POS a los elementos relevantes de la consulta (sustantivos, verbos y palabras pregunta ‘wh-words’) utilizando el “plugin” ANNIE. Por cada una de estas palabras se agrega un descriptor a *odquery* (1.1).

Si se considera como ejemplo la consulta “*Where can I eat vegetarian dishes?*”, este paso da como resultado:

```
1 [lema: where, POS: WRB, cls: null, ontologyDepth: null, semExpDepth: null]
2 [lema: can, POS: MD, cls: null, ontologyDepth: null, semExpDepth: null]
3 [lema: i, POS: PRP, cls: null, ontologyDepth: null, semExpDepth: null]
4 [lema: eat, POS: VB, cls: null, ontologyDepth: null, semExpDepth: null]
5 [lema: vegetarian, POS: NN, cls: null, ontologyDepth: null, semExpDepth: null]
6 [lema: dishes, POS: NNS, cls: null, ontologyDepth: null, semExpDepth: null]
7 [lema: ?, POS: ., cls: null, ontologyDepth: null, semExpDepth: null]
```

El proceso continua con una iteración sobre los descriptores producidos en el paso anterior, estos descriptores aún no han sido vinculados a la ontología

por lo que pertenecen a la lista de descriptores no vinculados del objeto *odquery* (1.2).

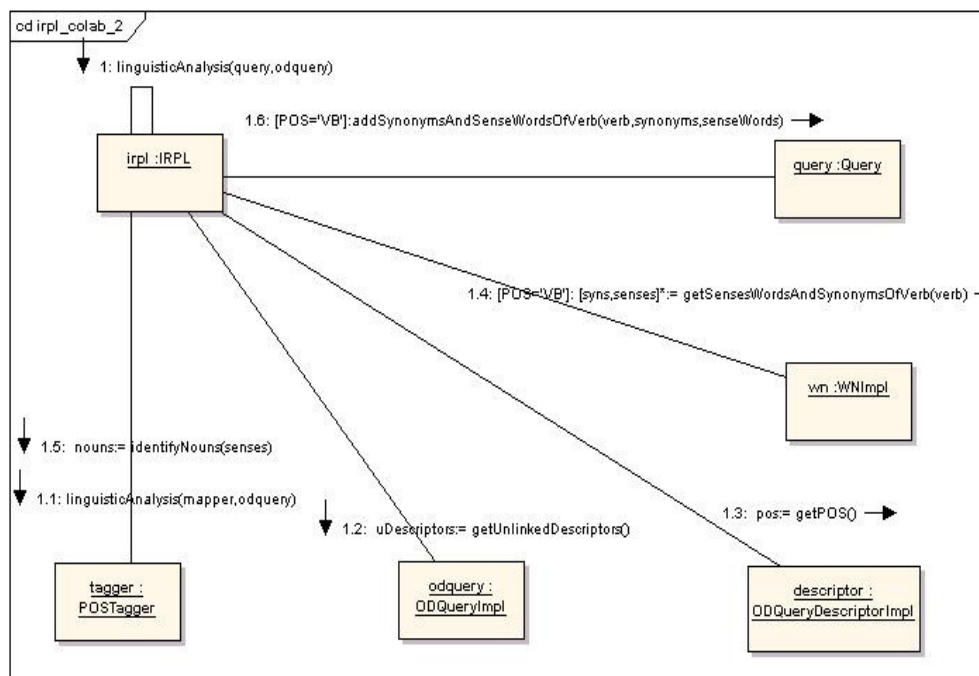


Figura B.2. Proceso General de Análisis Lingüístico

Al descriptor observado en la iteración, se le solicita su POS (1.3) y si el POS comienza con VB, se solicita a WordNet el conjunto de palabras que describen a este verbo (ver descripción detallada en la sección 2.2. – Obtención de Palabras de un Verbo). El objeto responde con una lista de sinónimos y una lista de oraciones descriptivas. WordNet también cuenta con ejemplos pero estos fueron excluidos debido a su ambigüedad (1.4).

De la información recuperada por WordNet sólo interesan los sustantivos, se solicita a *tagger* la identificación de los mismos, y éste a su vez utilizará el “plugin” ANNIE para devolver una lista de sustantivos (1.5) (ver descripción detallada en la sección 2.3. – Obtención de Sustantivos de Definición de un verbo). Esta información de descripción del verbo es almacenada en *query* para el paso de procesamiento (1.6).

2.1. Análisis Lingüístico

Como se observa en la Figura B.3, el primer paso del análisis lingüístico mediante ANNIE es la obtención de un documento de la librería GATE (*gate.corpora.Document*) a partir del texto de la consulta (1.1). A continuación se necesita una instancia de *Corpus* (*gate.corpora.Corpora*) para procesarla posteriormente usando ANNIE (1.2). Se agrega el documento al *Corpus* (1.3) y se inicializa el controlador de ANNIE con el *Corpus* (1.4). Se ejecuta luego el conjunto de recursos de procesamiento sobre el *Corpus* y el etiquetado resultante formará parte del documento existente en el *Corpus* (1.5).

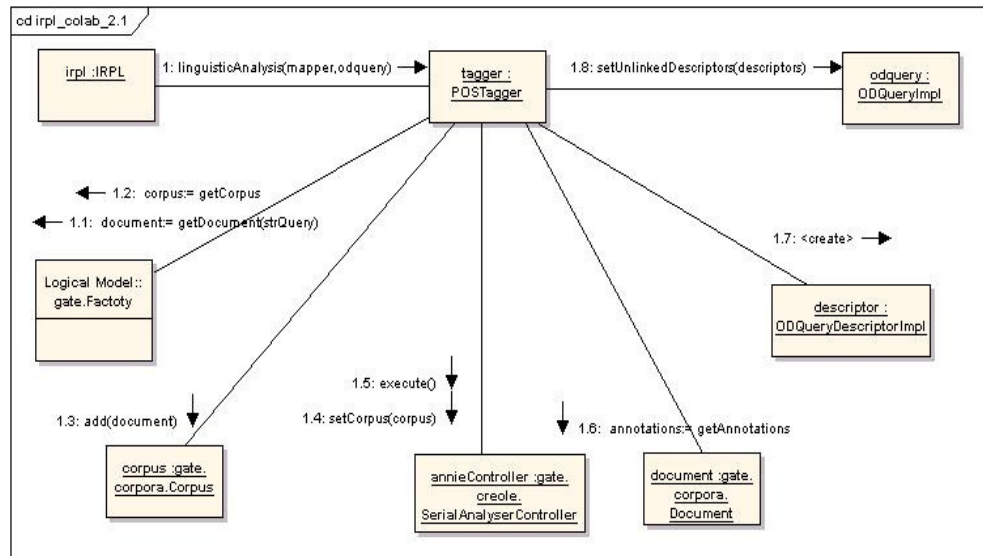


Figura B.3. Análisis Lingüístico

A continuación se pide al documento la lista de anotaciones obtenidas (1.6) y por cada etiqueta cuyo POS comience con NN o VB o WH, se crea un descriptor (*ontodom.ODQueryImpl*), este descriptor es inicializado con la palabra y el POS (1.7). Finalmente, el conjunto de descriptores hallados es colocado en la lista de descriptores no vinculados al dominio perteneciente a la instancia *odquery* (1.8).

2.2. Obtención de Palabras de Definición de un Verbo

Como se observa en la Figura B.4, el objetivo es utilizar el verbo de la consulta como elemento de vinculación con el dominio respecto de la acción o actividad que implique, y por lo tanto, son necesarios los sustantivos que definen el significado de dicho verbo.

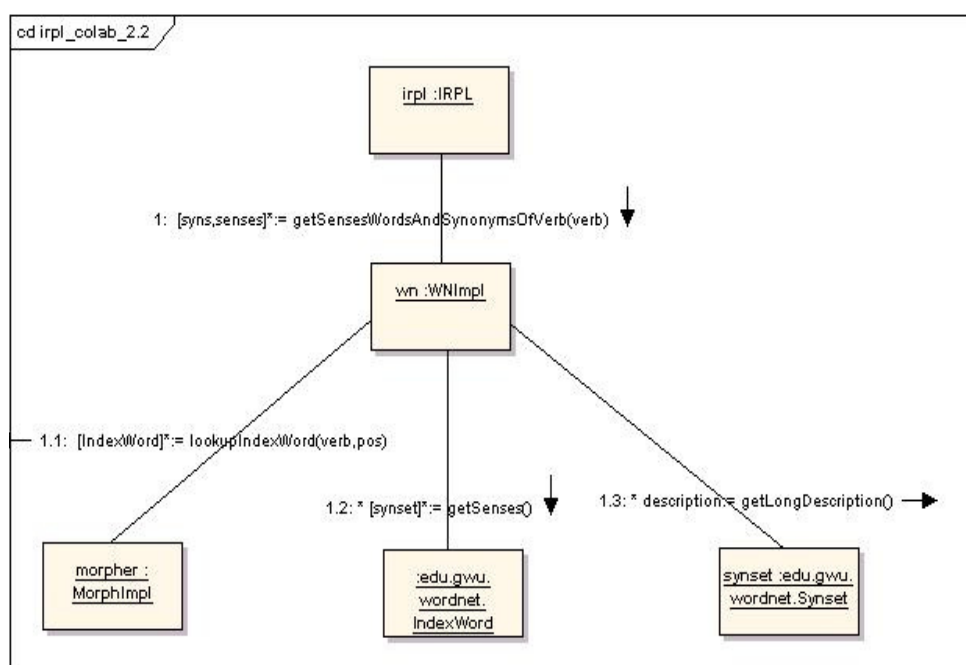


Figura B.4. Obtención de Palabras de Definición de un Verbo

El primer paso es solicitar a *morpher* la recuperación del conjunto de palabras del archivo *índice* (*index.verb*) que apuntan al archivo de datos de verbos (*data.verb*). Estos índices mencionan a los términos base del verbo encontrado en la consulta. El trabajo de *morpher* en este punto es obtener la forma base o raíz del verbo y usarla para obtener los índices de WordNet sobre su archivo de datos de verbos (1.1). A cada *indexWord* (*edu.gwu.WordNet.IndexWord*) se le solicita la lista de *synsets* que definen los

sentidos apuntados (*data.verb*) por el mismo (1.2) y a cada *synset* (*edu.gwu.WordNet.Synset*) se le solicita su descripción. Esta incluye el conjunto de sinónimos y el conjunto de oraciones que definen el sentido. La descripción también incluye ejemplos pero los mismos fueron dejados de lado debido a su ambigüedad (1.3).

2.3. Obtención de los Sustantivos de Definición de un Verbo

Como se observa en la Figura B.5, del conjunto de palabras obtenidas en el paso anterior sólo se consideran útiles los sustantivos. En este punto se utiliza una vez más a ANNIE y su proceso de etiquetado.

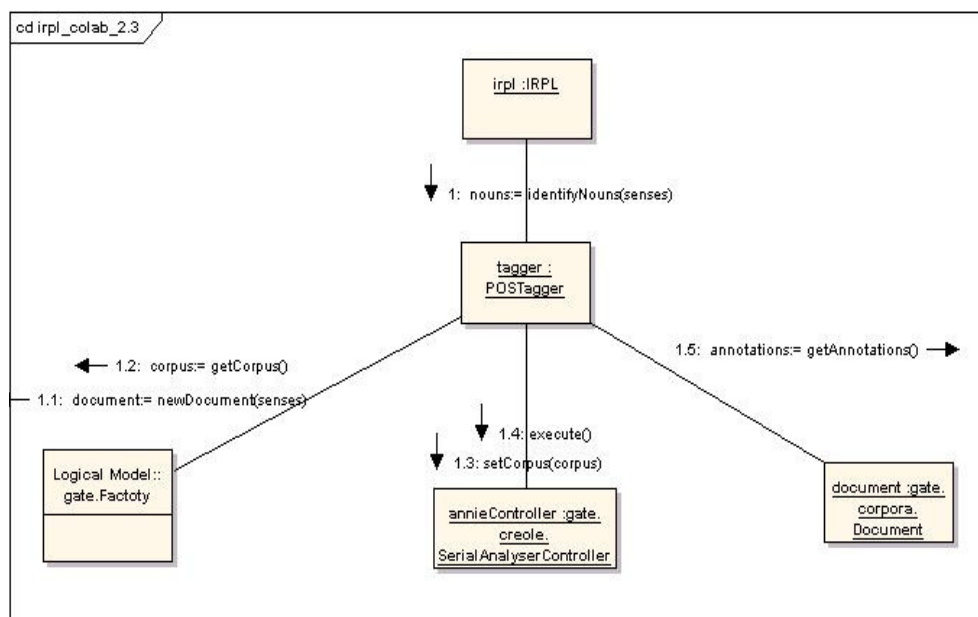


Figura B.5. Obtención de Sustantivos de Definición de un Verbo

El proceso comienza con la creación del objeto *document* a partir de las palabras halladas (1.1), luego se crea el *corpus* y se adiciona al mismo el objeto *document* (1.2). Se inicializa el objeto *annieController* utilizando el *corpus* (1.3) y se ejecutan los recursos de procesamiento de ANNIE sobre el *document* para el

etiquetado correspondiente (1.4). Finalmente, una vez obtenidas las etiquetas, el conjunto será filtrado teniendo en cuenta la condición de que el POS comience con NN (1.5).

3. Proceso de Identificación de Sustantivos

Como se observa en la Figura B.6, este proceso es equiparable al segundo proceso de etiquetado (Proceso de Búsqueda de Ocurrencias Exactas), la diferencia central es el tipo de objetos que realizan la colaboración.

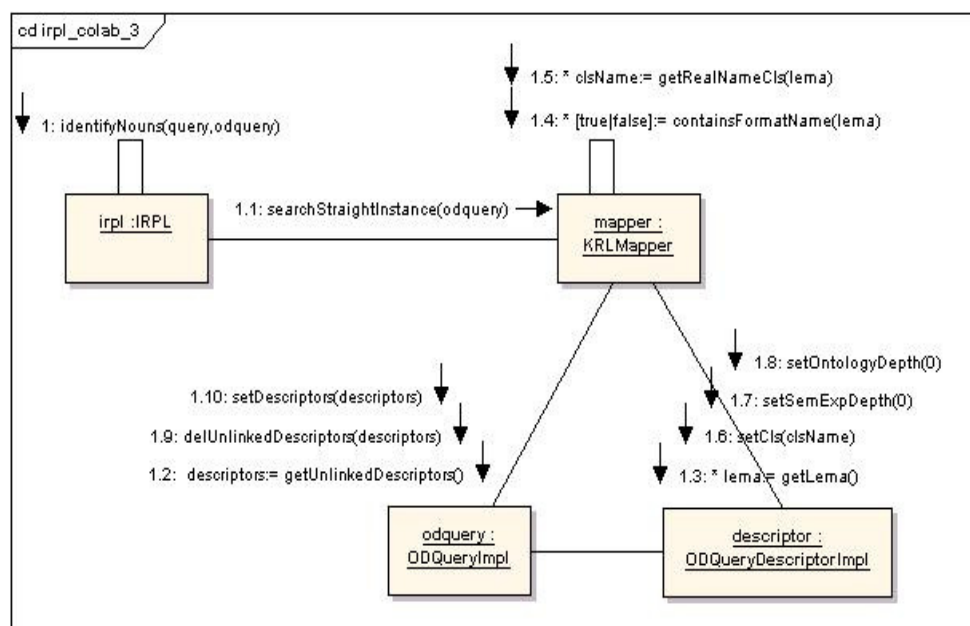


Figura B.6. Proceso de Identificación de Sustantivos

La instancia de IRPL solicita a la instancia de mapeo la búsqueda de ocurrencias exactas en la consulta en proceso (1.1). En primer lugar se necesitan los descriptores hallados en la sección 1 – Proceso General, los que se solicitan a la instancia *odquery* (*ontodom.ODQueryImpl*). Como ya se mencionó, estos descriptores aún no están enlazados al dominio. En este punto *mapper* comienza a iterar sobre la colección de descriptores (1.2).

Al descriptor observado en la iteración, se le solicita su *lema* (1.3) y se compara dicho *lema* con el conjunto de nombres de clases formateados para su comparación (1.4). Si concuerda, se recupera el nombre real de la clase ontológica (1.5) y como se ha encontrado una ocurrencia exacta, se da valor a la propiedad *cls* del descriptor (1.6).

A continuación se da valores a los campos relacionados con la profundidad de búsqueda. A la profundidad de búsqueda en la ontología, como no se ha subido ningún nivel taxonómico, se le otorga el valor cero (1.7) y como no se realizó expansión semántica la profundidad de búsqueda en WordNet también toma el valor cero (1.8).

Terminada la iteración, el conjunto de descriptores vinculados al dominio es restado del conjunto de descriptores no vinculados (1.9) y este conjunto se usa para iniciar el conjunto de descriptores vinculados al dominio que es propiedad de la instancia *odquery* (1.10).

4. Proceso de Identificación de Relaciones

Como se observa en la Figura B.7, aquí el análisis trata de identificar las relaciones ontológicas vinculables a la consulta y a la respuesta de la misma.

Este proceso tiene sentido sólo sobre el conjunto de descriptores asociados al dominio, por lo tanto, en este paso se recuperan esos descriptores (1.1). A cada descriptor se solicita su lema (1.2) y su nombre de clase asociada (1.3).

Al objeto *model* (*ontodom.Model*) se le solicita las relaciones de la clase y por cada relación se devuelve el nombre de la misma y la lista de clases que permite acceder (1.4) (ver descripción en la sección 4.1. – Obtención de Relaciones Ontológicas).

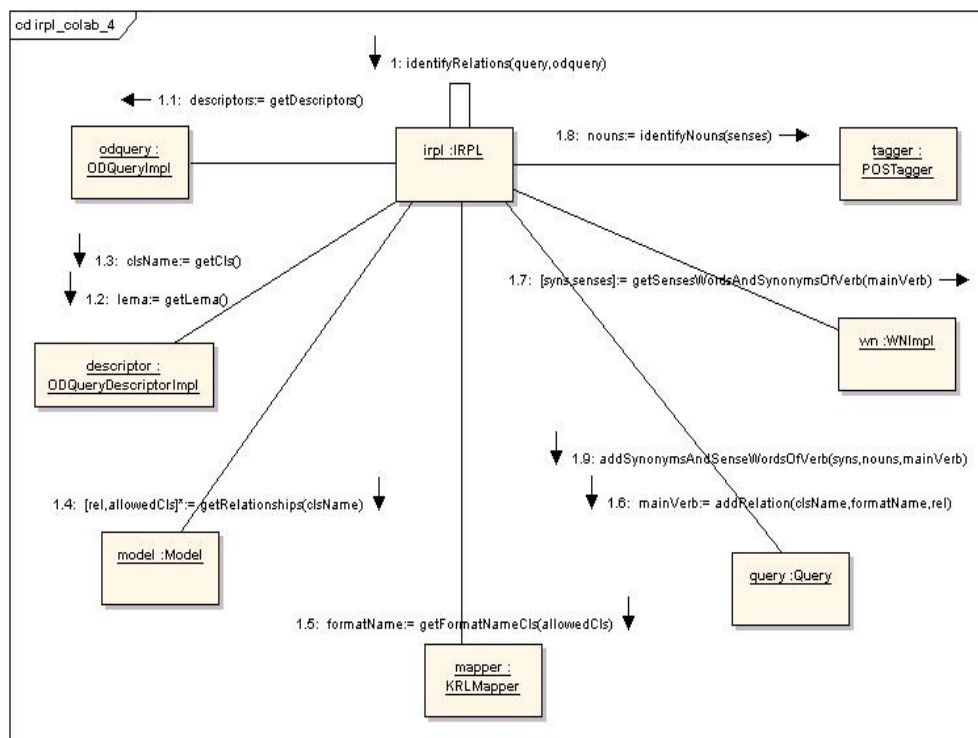


Figura B.7. Proceso de Identificación de Relaciones

A continuación se construyen las relaciones halladas: se formatea el nombre de cada clase accesible mediante la relación hallada (1.5) y se agrega la relación a la información de procesamiento de *query*. Cada relación añade la terna (clase, clase accesible, relación). Este paso devuelve el verbo principal de la relación acorde con la siguiente convención (1.6):

- i.* Relación *es un* (is a): relaciones omitidas del análisis de verbo, se dan por válidas (más adelante se explicará este criterio), el verbo principal es nulo;
- ii.* Relaciones que comienzan con verbos *ser o estar* (is, are, be): se toma la siguiente palabra;
- iii.* Relaciones que comienzan con verbos *tener* (have, has): son omitidas por ofrecer pocas posibilidades de análisis, el verbo principal es nulo;

- iv. De otro modo se toma por verbo principal a la primer palabra del nombre de la relación;

Para el ejemplo “*Where can I eat vegetarian dishes?*” se citan los siguientes resultados a partir de la ontología de ejemplo Travel⁶:

```
(Chef , Food, specialized(specializedIn))
(Chef , Vegetarian, specialized(specializedIn))
(Cuisine, Food, cook(cook))
(Cuisine, Vegetarian, cook(cook))
(Destination, Food, null(hasTypeOfFood))
(Destination, Vegetarian, null(hasTypeOfFood))
(Restaurant , Food, serves(serves))
(Restaurant , Vegetarian, serves(serves))
(Vegetarian, Food, null(isA))
```

Tomando el verbo principal de la relación, se utiliza WordNet para obtener información que describa al verbo a fin de tratar de vincular éste a la consulta en proceso (1.7). De la descripción entregada por WordNet, esto es, sinónimos más oraciones de explicación del sentido, sólo se toman los sustantivos. En este punto se vuelve a usar a ANNIE para obtener los sustantivos de este conjunto de palabras (1.8).

4.1. Obtención de Relaciones Ontológicas

Como se observa en la Figura B.8, a continuación se describe el proceso de recuperación de relaciones ontológicas a fin de mostrar el uso de las librerías de Protégé.

Dado un nombre de clase ontológica, el primer paso es solicitar al modelo (*edu.stanford.smi.protege.model.OWLModel*) una instancia de *OWLNamedClass* a fin de poder recuperar las relaciones definidas para la clase

⁶ Disponible, en su versión original, en <http://protege.stanford.edu/plugins/owl/owl-library/index.html>

(1.1). A dicha clase se le solicita el conjunto de superclases (1.2) y el mismo es usado para incorporar las relaciones tipo *isA* a la respuesta a IRPL (1.3).

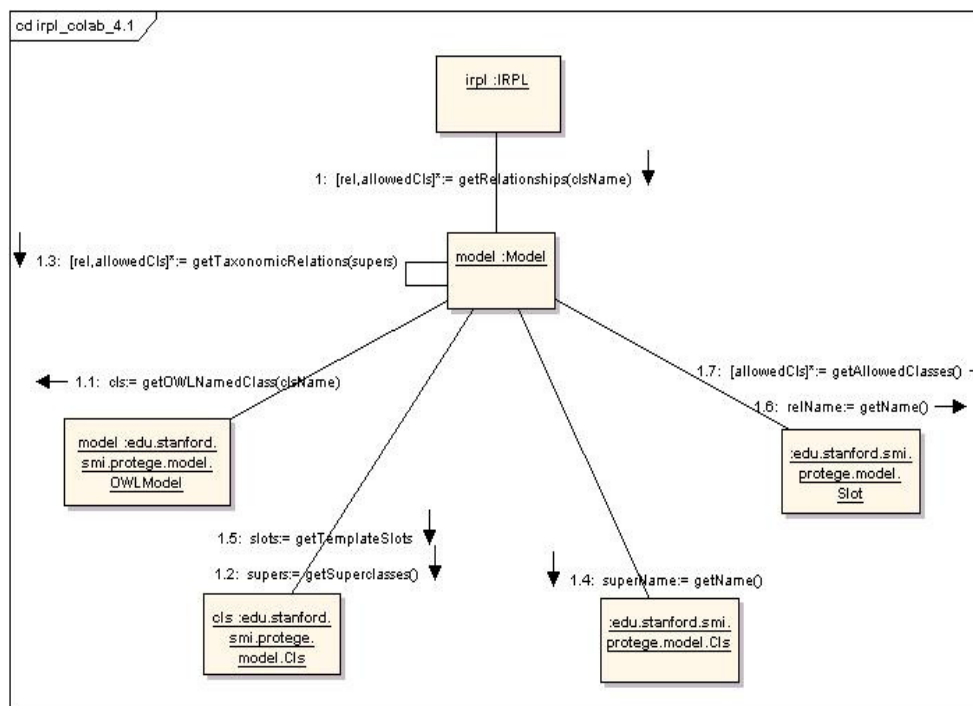


Figura B.8. Obtención de Relaciones Ontológicas

A cada superclase se le solicita su nombre, ese es el parámetro requerido por la librería, no se precisa el nombre de la relación que es asignado directamente con el valor *isA* (1.4). A continuación se vuelve a trabajar con la *OWLNamedClass* original. En este punto solo interesan las relaciones no taxonómicas, con este fin se solicitan los *slots* definidos para la clase (1.5).

Cada *slot* (*edu.stanford.smi.protege.model.Slot*) define una relación entre clases, por lo que una clase es accesible mediante esa relación y en este punto se solicita el nombre de la relación (1.6). La relación vincula a la clase en estudio con un conjunto de clases, el mismo es solicitado por *model*. Este objeto añade, por cada clase relacionada, el par (relación, nombre de clase). Este conjunto es unido al conjunto de relaciones taxonómicas y finalmente devuelto a IRPL (1.7).

5. Identificación de Sustantivos de Definición de un Verbo

A continuación se describirá el detalle del paso 1.8 del procesamiento de la consulta.

Como se observa en la Figura B.9, dado el verbo, el primer paso es solicitar al objeto de interfaz con WordNet la búsqueda de las palabras de definición del verbo. Este proceso es el mismo desarrollado en la sección 2.2. – Obtención de Palabras de Definición de un Verbo (1.1). Del conjunto e información que WordNet ofrece, sólo interesan los sustantivos, este proceso es el mismo explicado en la sección 3 – Proceso de Identificación de Sustantivos, la diferencia es que la instancia de *gate.corpora.Document* utilizada es construida en base a la información de definición del verbo (1.2).

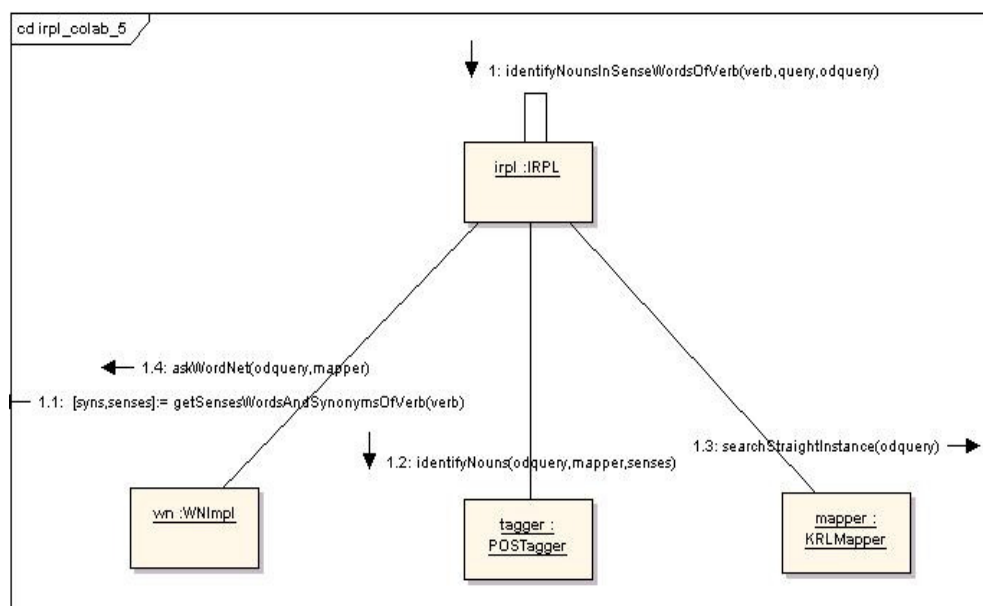


Figura B.9. Identificación de Sustantivos de Definición de un Verbo

A continuación se realiza la búsqueda de ocurrencias exactas de términos ontológicos utilizando los descriptores hasta ahora hallados (ver el proceso detallado en sección 3 – Proceso de Obtención de Sustantivos) (1.3). Sobre los descriptores de la consulta no vinculados a la ontología, se procede a realizar el proceso de expansión semántica utilizando WordNet. Este proceso es el mismo

descripto para KRL en la sección 3 - Proceso de Obtención de Sustantivos, la diferencia sólo se centra en las instancias con las que se trabaja, esto es *ODQueryImpl* en lugar de *DocImpl* y *ODQueryDescriptorImpl* en lugar de *DocDescriptorImpl* (1.4).

Hasta ahora se ha descripto la ejecución del procesamiento de la consulta, se comenzará ahora a describir el proceso de recuperación de los documentos.

6. Proceso de Preparación de Consulta

Esta colaboración corresponde a los pasos 1.7 y 1.11 del proceso general (sección 1). Como se observa en la Figura B.10, este proceso se limita a inicializar las variables necesarias para el proceso de recuperación.

En primer lugar, este proceso trabaja con el conjunto completo de descriptores, y en todos los pasos que siguen a continuación se itera sobre esta lista (1.1).

Se solicita el POS del descriptor (1.2) y la profundidad del descriptor respecto de la taxonomía ontológica. Si bien esta operación se realiza y en el paso 1.7 es un criterio, no se han incluido los padres de los términos hallados en la ontología, debe observarse que no se ha agregado un subproceso tal como el Proceso de Inclusión de Superclases de KRL (sección 4), la consecuencia de esto es que siempre será cero. (1.3).

Se solicita la profundidad del descriptor respecto de la profundidad de búsqueda en WordNet, esto es la cantidad de veces que se subió en la taxonomía de WordNet mediante la relación de hiperonimia entre *synsets* (1.4).

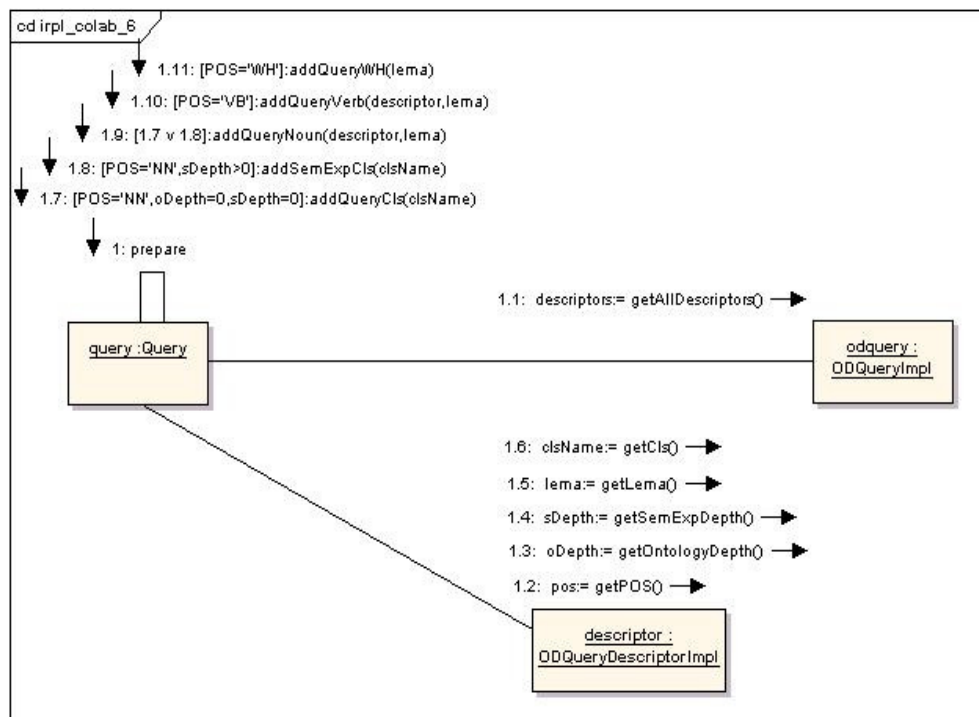


Figura B.10. Proceso de Preparación de Consulta

Se solicita el lema del descriptor, esto es, una palabra perteneciente a la consulta (1.5) y el nombre del término ontológico asociado al descriptor, válido sólo si el lema es un sustantivo (1.6).

Si el POS etiqueta al lema como un sustantivo y el nivel de profundidad de búsqueda de expansión semántica es cero al igual que el nivel de profundidad taxonómica, entonces se trata de un sustantivo de la consulta que es a la vez un término ontológico, esto se cumple para los términos hallados por aparición exacta y para los términos encontrados por sinonimia mediante WordNet (1.7).

Si el POS etiqueta al lema como un sustantivo y el nivel de profundidad de búsqueda de expansión semántica es mayor que cero, entonces se trata de un sustantivo de la consulta que es hipónimo de un término ontológico hallado mediante la relación de hiperonimia de WordNet (1.8).

Finalmente, si se cumplen 1.7 o 1.8 el término de la consulta es un sustantivo útil de la misma y es tomado en cuenta para la recuperación de

documentos junto con las clases halladas (1.9). Por otro lado, si el POS etiqueta al lema como un verbo, el lema se lista para procesamiento posterior (1.10), o si el POS etiqueta al lema como una palabra pregunta (wh-word), el lema se lista para un proceso de validación que se describe a continuación (1.11).

Nota: Como los verbos y las palabras preguntas no son ligados a la ontología no son persistidos.

6.1. Proceso de Validación Inicial

Este proceso selecciona las relaciones halladas en el Proceso de Identificación de Relaciones (sección 4), el análisis separa relaciones válidas de inválidas. Las relaciones inválidas podrán usarse si se logra vincular el verbo de la relación con el significado de la consulta.

Son relaciones válidas aquellas que cumplan uno de los siguientes criterios:

- i.* El origen y el destino de la relación están incluidos entre las clases de la consulta o entre las clases halladas por expansión semántica;
- ii.* Toda relación *isA*.

Para la consulta de ejemplo, los resultados son los siguientes:

- ✓ La relación (Food, Vegetarian, null(isA)) es válida;
- ✓ La relación (Cuisine, Vegetarian, cook(cook)) necesita validarse;
- ✓ La relación (Chef, Vegetarian, specialized(specializedIn)) necesita validarse;
- ✓ La relación (Destination, Vegetarian, null(hasTypeOfFood)) necesita validarse;
- ✓ La relación (Restaurant, Vegetarian, serves(serves)) necesita validarse;
- ✓ La relación (Cuisine, Food, cook(cook)) necesita validarse;
- ✓ La relación (Chef, Food, specialized(specializedIn)) necesita validarse;
- ✓ La relación (Destination, Food, null(hasTypeOfFood)) necesita validarse;
- ✓ La relación (Restaurant, Food, serves(serves)) necesita validarse;

7. Proceso de Recuperación

Como se observa en la Figura B.11, este proceso es en su mayoría algorítmico y comienza cuando la consulta *query* ya ha sido inicializada. El proceso no valida las relaciones *isA* encontradas y todo término ontológico hallado es incluido junto con sus superclases. Ahora, todas las relaciones no taxonómicas cuyos verbos principales no hayan sido incluidos por ser muy generales deben ser validadas. Con esto quiere decirse que la inclusión o no de las clases accesibles mediante estas relaciones deben ser evaluadas respecto de la ontología de dominio.

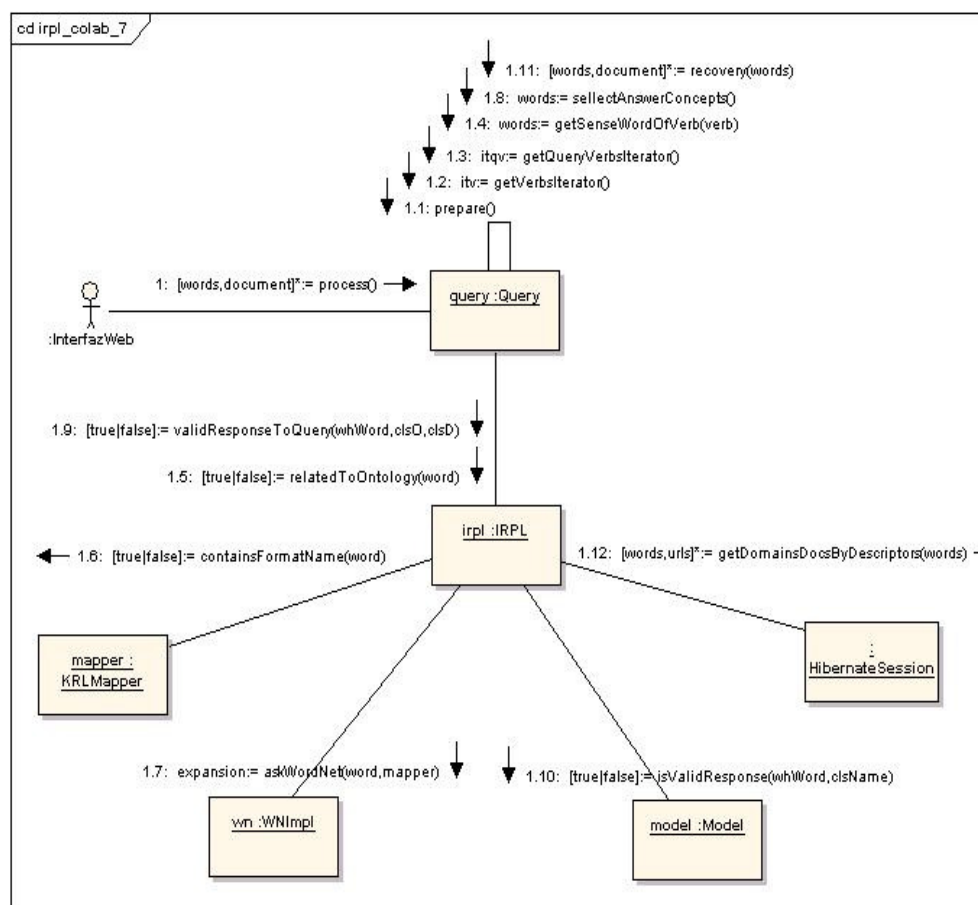


Figura B.11. Proceso de Recuperación

En este punto es cuando finalmente se utiliza la información dada por WordNet para los verbos hallados en la consulta y en las relaciones ontológicas.

En primer lugar se prepara la consulta, si bien se mencionó al describir el proceso general el caso en que no se logra obtener clases ontológicas a partir de los sustantivos de la consulta, no se espera que este sea el caso general. Aquí se procede a estudiar la información recogida como se describió antes al explicar el proceso de Preparación de Consulta (sección 6) (1.1)

El proceso comienza iterando sobre los verbos principales de las relaciones (1.2) y el siguiente nivel de iteración se hace sobre los verbos hallados en la consulta (1.3). Luego se itera sobre los sustantivos que se han obtenido para cada verbo. Se comienza pidiendo la lista de sustantivos de cada verbo, se itera por las mismas (dos niveles más de iteración) tomando una palabra de cada una de las listas a la vez. Si las palabras son iguales, (esto es, un sustantivo descriptivo del verbo de la consulta es igual al sustantivo descriptivo del verbo de una relación) entonces sólo resta ver si este sustantivo está vinculado a la ontología (1.4).

El paso siguiente es pedir a IRPL que informe si la palabra es vinculable o no a la ontología (1.5). El ejemplo más claro de correspondencia es que la palabra pueda ser un término ontológico, para ello se observa si *mapper* contiene un nombre de clase formateado igual a la palabra estudiada, en caso verdadero no se continúa buscando. Se incrementa un contador en cada relación exitosa en este proceso de contextualización que será usado en 1.8 (1.6).

Si el proceso anterior fracasa se expande la búsqueda usando WordNet, este proceso es igual que otros en los que se utiliza a WordNet. La diferencia es que aquí no interesan ni descriptores ni sustantivos, sólo importa si entre el conjunto de sinónimos e hiperónimos de esa palabra existe al menos un término ontológico. Se incrementa un contador en cada relación exitosa en este proceso de contextualización que será usado en 1.8 (1.7).

Para la consulta de ejemplo se citan los siguientes éxitos:

Relación: (Cuisine, Vegetarian, cook(cook))

Verbo principal: cook
 Verbo de la consulta: eat
 Sustantivo en común: meal
 Relación: (Chef, Vegetarian, specialized(specializedIn))
 Verbo principal: specialized
 Verbo de la consulta: eat
 Sustantivo en común: way
 Relación: (Restaurant, Vegetarian, serves(serves))
 Verbo principal: serves
 Verbo de la consulta: eat
 Sustantivo en común: food

En el siguiente paso se seleccionan los descriptores, los mismos están expresados mediante relaciones, esto es ternas (*clsO*, *clsD*, verbo principal). Se itera sobre estas relaciones: si la relación es válida (sección 6 – Proceso de Preparación de Consulta) se procede a incorporar su clase origen y destino como parte de los descriptores seleccionados para la recuperación de documentos. Las relaciones inválidas son incluidas de la misma forma si, primero fueron exitosas en el proceso de contextualización descripto arriba, y segundo, al menos una de sus clases es respuesta válida para el tipo de pregunta (1.8).

El segundo requisito de aceptación de las clases de una relación inválida se realiza en este paso donde se solicita a IRPL que busque dentro del modelo *owl*, si al menos una de las clases de la relación es respuesta válida para el tipo de pregunta tipificado por la palabra pregunta (*wh-word*) (1.9). IRPL traslada la petición anterior al modelo, éste ya tiene almacenado los nombres de las clases que son respuesta válida para cada tipo de pregunta. Esto se hizo en el proceso de inicialización de la herramienta.

La razón para esto es que se quiere dejar de lado aquellas relaciones que se han observado como exitosas en los pasos 1.6 y 1.7 pero que no son respuestas aceptables para el tipo de pregunta.

Para la consulta de ejemplo, la relación: (*Cuisine*, *Vegetarian*, *cook(cook)*) no pasa el criterio porque tanto *Cuisine* como *Vegetarian* no son respuestas válidas para una pregunta tipo *Where*. Ahora, la relación (*Restaurant*,

Vegetarian, serves(serves)), sí pasa el criterio gracias a que *Restaurant* es una respuesta válida a una pregunta tipo *Where*.

Después de una fase de prueba se encontró que apoyarse en WordNet para esta función producía resultados inestables. Se decidió por lo tanto, que es una de las responsabilidades del ingeniero ontológico, al construir la ontología, etiquetar las clases con los tipos de respuestas que ésta puede responder. En la ontología de ejemplo, *Travel*, se ha realizado la clasificación mostrada en la Tabla B.1.

| DataTypeProperty | OWLNamedClass |
|-------------------------|--|
| IRPL_whatValidResponse | Todas las clases de primer nivel |
| IRPL_whenValidResponse | Vacation_Period |
| IRPL_whereValidResponse | Acommodation; Destination; Facility. |
| IRPL_whoValidResponse | Chef |
| IRPL_wichValidResponse | Todas las clases de primer nivel |

Tabla B.1. Clasificación Tipo de Respuesta

Si bien las propiedades *DataTypeProperty* no son heredables, todas las subclases de las clases etiquetadas con uno de los tipos son consideradas como respuesta válida para el tipo de pregunta etiquetado (1.10).

El conjunto de nombres de clases obtenido hasta el momento más el conjunto de sustantivos de la consulta que fueron vinculados a la ontología mediante expansión semántica es utilizado para recuperar los descriptores. En este paso se construye esta lista de palabras y se solicita a una sesión de Hibernate la recuperación de las URLs de los documentos que mencionan en su anotado a estos descriptores (1.11).

Este es el paso de recuperación en sí mismo y se trata de una consulta en el lenguaje HQL (Hibernate Query Language) (1.12).

```
Ej:  
SELECT D  
FROM DocImpl AS D, OntologyImpl AS O, DocDescriptorImpl AS K  
WHERE (D.ontology.id = O.id) AND (O.URLDomain = :domain) AND  
(K.document.id = D.id) AND (  
  (K.lemma LIKE 'dishes') OR (K.cls LIKE 'dishes')OR  
  (K.lemma LIKE 'Food') OR (K.cls LIKE 'Food')OR  
  (K.lemma LIKE 'Restaurant') OR (K.cls LIKE 'Restaurant ')OR  
  (K.lemma LIKE 'Vegetarian') OR (K.cls LIKE 'Vegetarian')  
)
```


Componente de Interfaz a Otros Dominios Propagación Selectiva

1. Valoración de los Dominios

Como se observa en la Figura C.1, la propagación comienza cuando el usuario la solicita. La interfaz web usa el servicio *propagate()* del objeto *odi* (*ontodom.ODI*) del dominio consultado. El servicio recibe como parámetro la instancia *odquery* creada en IRPL (componente de recuperación de información) recuperando la misma desde la sesión del usuario.

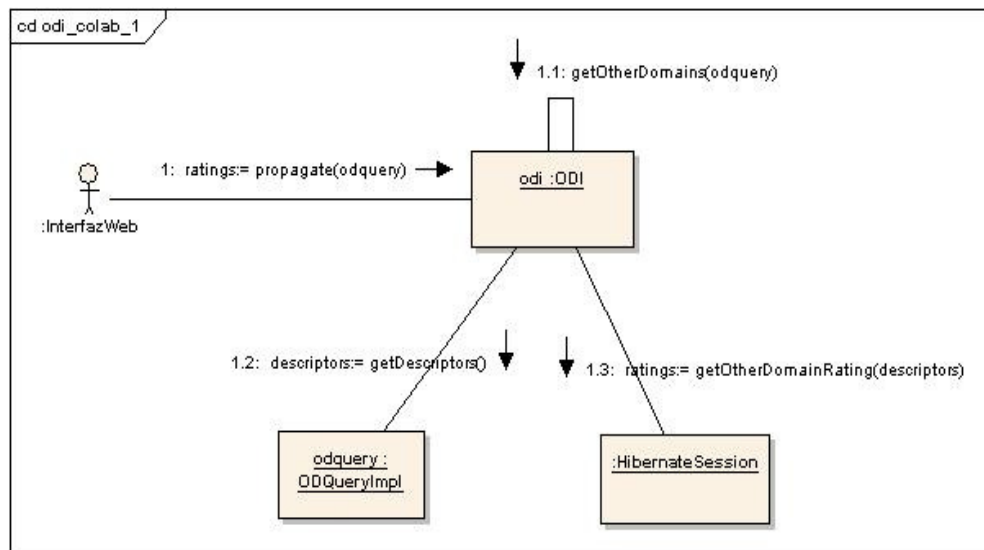


Figura C.1. Valoración de los Dominios

Este paso devuelve, para cada dominio, la valoración del usuario y el índice de similitud de la mejor consulta almacenada en el dominio, calculado en base a la fórmula presentada en la sección 5.4.2. del Capítulo VIII:

$$SimQuery(dr, Qp) = \frac{MD}{DQdr} \left[1 + \sum_{i=1}^{DQdr} (OD_i[Qdr, dl] + SD_i[Qdr, dl]) \right]^{-1} \left(2 - \frac{1}{UR[Qdr]} \right)$$

El proceso *getOtherDomains(odquery)* prepara la propagación instanciando las variables necesarias (1.1). El proceso continúa con la evaluación de similitud donde sólo se toman en cuenta los descriptores de la consulta que está siendo propagada (Qp). En este paso, los mismos son recuperados desde la instancia *odquery* (1.2). Finalmente, el criterio de elección de los dominios candidatos se traduce en una consulta HQL (Hibernate Query Language) (1.3).

2. Propagación

Una vez visualizados los dominios de propagación, el usuario escoge aquellos a los que desea propagar, la interfaz web ejecuta una petición de procesamiento de la consulta a cada dominio seleccionado, así la colaboración esquematizada en la Figura C.2 ocurre tantas veces como dominios haya escogido el usuario.

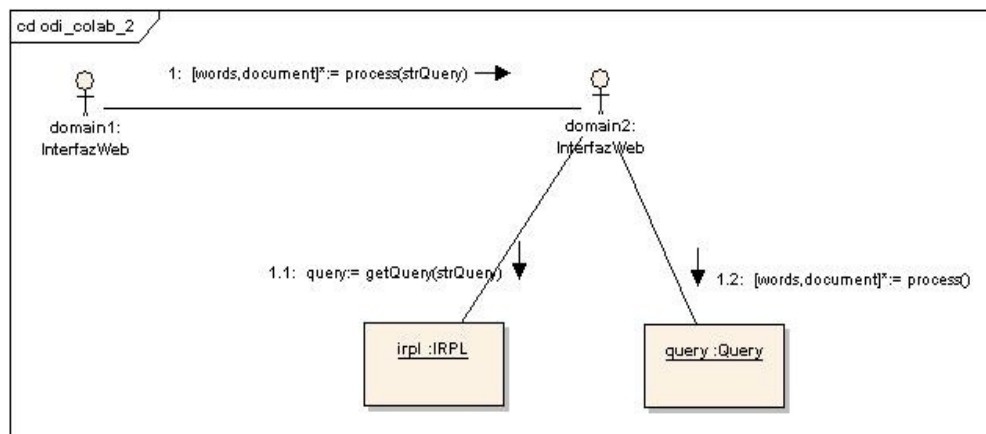


Figura C.2. Propagación de la Consulta

En la Figura C.2 puede verse cómo las interfaces web colaboran entre sí para resolver la petición del usuario. Aquí podemos reconocer a OntoDOM como a un sistema distribuido.

Cualquier servidor HTTP que responda a la interfaz esperada (consulta *strQuery*; respuesta [*descriptores(palabras),documento(URL)*]⁷) es un dominio de propagación potencial. En el sistema de ejemplo, todos los servidores son réplicas exactas diferenciadas solo en la ontología y el *corpus*, pero no hay ninguna limitación al respecto. Esto no quiere decir que puede ser cualquier servidor, todo dominio al que pretenda propagarse debe ser configurado previamente.

3. Valoración de la consulta

La capa ODI incluye entre sus capacidades el aprendizaje en casos, y para ello se apoya en el usuario para obtener una valoración de las respuestas dadas por un dominio determinado. Esta información es almacenada para ser recuperada más tarde como ya se describió en la sección 1 de este Apéndice.

⁷ El lector que requiera más información puede encontrarla en la sección OntoDOM – Proceso de Procesamiento y Recuperación de Información (Anexo B).

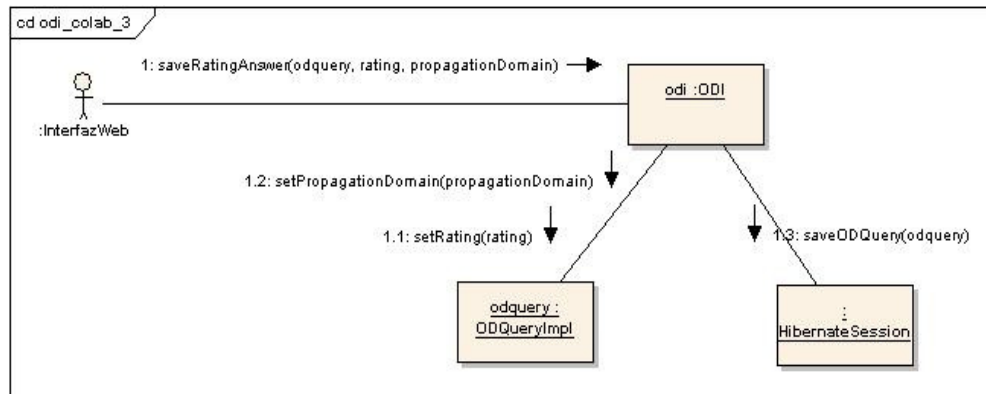


Figura C.3. Valoración de la Consulta

Como se observa en la Figura C.3, la colaboración comienza cuando el usuario valora el conjunto de respuestas (documentos) de un determinado dominio. Las valoraciones (rating) posibles son: "Excellent", "Very Good", "Good", "Regular", "Poor", "None" con valuaciones de 5 a 0 respectivamente. El proceso recupera la instancia de representación de la consulta (*ontodom.ODQuery*) desde la sesión de usuario a fin de almacenarla. El último parámetro requerido es el dominio que se está valorando a fin de asociarlo a la valoración del usuario y a los descriptores a persistir. Todos estos objetos son dados a la instancia de ODI del servidor por la interfaz web (1).

Para poder almacenar la valoración del usuario, esta se guarda en el objeto *odquery* (1.1). De la misma forma, se almacena el objeto *odquery* la URL del dominio valorado (1.2). El objeto *odquery* es persistido en base de datos mediante Hibernate (1.3.).

Anexo D

Datos Complementarios Caso Estudio

1. Ontología Europa-Travel

owl:Thing

- View
 - Mountain_View
 - Sea_View
 - Ocean_View
 - Panoramic_View
 - Scenic_View
- Accommodation
 - Shelter
 - Tent
 - Hut
 - Campground
 - Hotel
 - Resort
 - Mountain_Resort
 - Motel
 - Inn
 - Luxury_Hotel
 - Apartment
 - Bed_and_Breakfast
 - House
 - Manor
 - Castle
 - Condominium
 - Cabin
 - Cottage

-
- Bungalow
 - Villa
 - Facility
 - Butler_Service
 - Golf_Course
 - Restaurant
 - Tennis_Court
 - Swimming_Pool
 - Heated_Swimming_Pool
 - Internet_Access
 - Basketball_Court
 - Bar
 - Mini_Bar
 - Snack_Bar
 - Croquet_Court
 - Spa
 - Garden
 - Water_Garden
 - Playroom
 - Fitness_Center
 - Gym
 - Room_Service
 - Transportation
 - Ship
 - Yacht
 - Launch
 - Luxury_Glass_Domed_Motor_Launch
 - Boat
 - Jet_Boat
 - Sailboat
 - Motor_Boat
 - Public_Boat
 - Ferry
 - High_Speed_Ferry
 - Train
 - Cogwheel_Train
 - Electric_Railway
 - Rack_Railway
 - Aircraft
 - Bicycle
 - Bus
 - Double_Decker_Bus
 - Open_Top_Double_Decker_Bus
 - Motor_Coach
 - Car
 - Cable_Car
 - Minivan

-
- Air_Conditional_Minivan
 - Taxi
 - Vacation_Period
 - Month
 - Season
 - Summer
 - Spring
 - Winter
 - Activity
 - Sport
 - Horseback_Riding
 - Shooting
 - Clay_Shooting
 - Air_Rifle_Shooting
 - Tennis
 - Sailing
 - Falconry
 - Polo
 - Golf
 - Ice_Skating
 - Basketball
 - Paddling
 - Extreme_Sport
 - Water_Skiing
 - Rock_Climbing
 - Windsurfing
 - Heli_Skiing
 - Mountain_Climbing
 - Paragliding
 - Parasailing
 - Climbing
 - Rock_Climbing
 - Mountain_Climbing
 - Skiing
 - Water_Skiing
 - Heli_Skiing
 - Biking
 - Mountain_Biking
 - Volleyball
 - Hiking
 - Archery
 - Water_Sport
 - Boating
 - Jet_Boating
 - Rafting
 - White_Water_Rafting
 - Canoeing

- Swimming
- Surfing
- Kayaking
- Fishing
 - River_Fishing
 - Fly_Fishing
 - Deep_Sea_Fishing
- Diving
 - Scuba_Diving
 - Wreck_Diving
- Snorkeling
- Aqua_Aerobics
- Football
- Relaxation
 - Yoga
 - Strolling
 - Pilates
 - Sailing
 - Fishing
 - River_Fishing
 - Fly_Fishing
 - Deep_Sea_Fishing
 - Spa_Treatment
 - Facial_Treatment
 - Body_Treatment
 - Sauna
 - Reflexology
 - Meditation
 - Massage
 - Sun_Bed
 - Acupressure
 - Lymphatic_Drainage
 - Mud_Bath
 - Hydrotherapy
 - Digitopressure
 - Steam_Bath
 - Aromatherapy
 - Jacuzzi
 - Turkish_Bath
- Golf
- Sunbathing
- Picnic
- Entertainment
 - Theatre
 - Opera
 - Cinema
 - Aquarium

-
- Cabaret
 - Disco
 - Concert
 - Casino
 - Cafe
 - Dancing
 - Ballet
 - Gallery
 - Safari
 - Museum
 - Modern_Art
 - Ancient_Art
 - Nightclub
 - Sightseeing
 - Adventure
 - Excursion
 - Boat_Ride
 - Archaeological_Journey
 - River_Cruise
 - Visit
 - Landmark_Visit
 - Cemetery
 - Ruin
 - Volcano
 - Aqueduct
 - Historic_Center
 - Monument
 - Statue
 - Carving
 - Bridge
 - UNESCO_World_Heritage
 - Park
 - UNESCO_Cultural_Heritage
 - Winery
 - Market
 - Antique_Market
 - Flea_Market
 - Horse_Market
 - Flower_Market
 - Village_Market
 - Bird_Market
 - Cavern
 - Fountain
 - Geyser
 - Square
 - Gallery
 - Painting

- Fresco
- Museum
 - Modern_Art
 - Ancient_Art
- Concentration_Camp
- Bath
- Building_Visit
 - Town_Hall
 - Chapel
 - Geothermal_Plant
 - Church
 - Palace
 - Royal_Palace
 - UNESCO_World_Heritage
 - Tower
 - Arch
 - UNESCO_Cultural_Heritage
 - Forum
 - Convent
 - Monastery
 - Cathedral
 - Basilica
 - Castle
 - Temple
 - Opera_House
 - House_of_Parliament
 - Amphitheater
 - Congress_Hall
 - Train_Station
 - Railway_Station
- Day_Trip
- Prehistoric_Mound
- Tour
 - Day_Tour
 - Full_Day_Tour
 - Half_Day_Tour
 - Bike_Day_Tour
 - Gallery_Tour
 - Guided_Tour
 - Guided_Coach_Tour
 - Walking_Tour
 - Guided_Walking_Tour
 - Double_Decker_Bus_Tour
 - City_Tour
 - Hop_On_Hop_Off_Tour
 - Boat_Tour
 - Panoramic_Tour

-
- Wildlife_Watching
 - Bungee_Jumping
 - Safari
 - Hunting
 - Camping
 - Chef
 - Travel_Class_3
 - Destination
 - Canal
 - Lake
 - Valley
 - Urban_Area
 - Country
 - Town
 - City
 - Capital
 - Mount
 - Loch
 - Beach
 - Village
 - Rural_Area
 - National_Park
 - Farmland
 - Forest
 - Family_Destination
 - Island
 - Coast
 - Retiree_Destination
 - Backpackers_Destination
 - Waterfall
 - Gulf
 - Mountain
 - Peak
 - Peninsula
 - Glacier
 - Budget_Hotel_Destination
 - Hill
 - Quiet_Destination
 - Food
 - Spicy
 - Cheese
 - Drink
 - Tea
 - Coffee
 - Mineral_Water
 - Soft_Drink
 - Wine

- Sparking_Wine
 - Champagne
 - Liqueur
 - Beer
 - Chocolate
 - Hot_Chocolate
- Organic
- Finger_Food
 - Snack
 - Light_Snack
- Chicken
- Shrimp
- Vegetarian
- Mussel
- Diet_Food
- Baguette
- Fish
- Seafood
- Cookie
- Ham
- Waffle
- Rabbit
- Vegetation
 - Tree
 - Flower
 - Hyacinths
 - Tulip
 - Daffodils
- Architecture_Style
 - Medieval
 - Gothic
 - French_Rococo
 - Neo_Romanesque
 - Georgian
 - Art_Deco

2. Ontología África-Travel

owl:Thing

- View
 - Mountain_View
 - Sea_View

-
- Panoramic_Sea_View
 - Ocean_View
 - Panoramic_View
 - Accommodation
 - Shelter
 - Tent
 - Hut
 - Campground
 - Backpaker_Lodge
 - Camping_Site
 - Caravan_Site
 - Hotel
 - Resort
 - Motel
 - Inn
 - Luxury_Hotel
 - Apartment
 - Bed_and_Breakfast
 - House
 - Manor
 - Castle
 - Condominium
 - Cabin
 - Cottage
 - Chalet
 - Bungalow
 - Villa
 - Facility
 - Butler_Service
 - Golf_Course
 - Restaurant
 - Tennis_Court
 - Swimming_Pool
 - Secluded_Swimming_Pool
 - Heated_Swimming_Pool
 - Plunge_Pool
 - Internet_Access
 - Television
 - Basketball_Court
 - Bar
 - Mini_Bar
 - Snack_Bar
 - Cocktail_Bar
 - Coffee_Bar
 - Spa
 - Croquet_Court
 - Garden

- Water_Garden
 - Playroom
 - Telephone
 - Air_Conditioning
 - Room_Service
 - Library
 - Fitness_Center
 - Gym
 - Patio
 - Sun_Patio
- Vacation_Period
 - Month
 - Season
 - Summer
 - Spring
 - Winter
- Transportation
 - Ship
 - Sailboat
 - Canoe
 - Dugout_Canoe
 - Motorboat
 - Yacht
 - Ferry
 - Train
 - Aircraft
 - Light_Aircraft
 - Private_Charter
 - Helicopter
 - Micro_Light
 - Charter_Flight
 - Air_Balloon_Flight
 - Hot_Air_Balloon_Flight
 - Bus
 - Minibus
 - Car
 - Combi
 - All_Terrain_Vehicle
 - 4x4_Vehicle
 - Safari_Vehicle
 - Open_Safari_Vehicle
 - Taxi
- Activity
 - Sport
 - Kloofing
 - Horseback_Riding
 - Shooting

- Clay_Shooting
- Air_Rifle_Shooting
- Tennis
- Sailing
- Falconry
- Polo
- Golf
- Ice_Skating
- Basketball
- Paddling
- Extreme_Sport
 - Water_Skiing
 - Rock_Climbing
 - Windsurfing
 - Heli_Skiing
 - Gorge_Swing
 - Mountain_Climbing
 - Paragliding
 - Hang_Gliding
 - Parasailing
- Climbing
 - Rock_Climbing
 - Mountain_Climbing
- Skiing
 - Water_Skiing
 - Heli_Skiing
- Biking
 - Mountain_Biking
- Volleyball
- Hiking
- Archery
- Water_Sport
 - Boating
 - Jet_Boating
 - Rafting
 - White_Water_Rafting
 - Canoeing
 - Swimming
 - Surfing
 - Kayaking
 - Fishing
 - Angling
 - River_Fishing
 - Tiger_Fishing
 - Fly_Fishing
 - Deep_Sea_Fishing
 - Diving

-
- Scuba_Diving
 - Wreck_Diving
 - Snorkeling
 - Aqua_Aerobics
 - Squash
 - Football
 - Abseiling
 - Relaxation
 - Yoga
 - Pilates
 - Sailing
 - Fishing
 - Angling
 - River_Fishing
 - Tiger_Fishing
 - Fly_Fishing
 - Deep_Sea_Fishing
 - Spa_Treatment
 - Facial_Treatment
 - Body_Treatment
 - Reflexology
 - Sauna
 - Meditation
 - Massage
 - Sun_Bed
 - Lymphatic_Drainage
 - Acupressure
 - Mud_Bath
 - Hydrotherapy
 - Digitopressure
 - Steam_Bath
 - Aromatherapy
 - Jacuzzi
 - Turkish_Bath
 - Golf
 - Sunbathing
 - Entertainment
 - Theatre
 - Opera
 - Cinema
 - Aquarium
 - Singing
 - Concert
 - Casino
 - Dancing
 - Ballet
 - Shopping

- Gallery
- Safari
 - Scenic_Safari
 - Walking_Safari
 - Boat_Safari
 - Photographic_Safari
 - Elephant_Back_Safari
 - Fishing_Safari
 - Camping_Safari
 - Ballooning_Safari
- Museum
 - Natural_History_Museum
 - Art_Museum
- Nightclub
- Sightseeing
- Adventure
 - Excursion
 - Ruin
 - Scenic_Flight
 - Village_Visit
 - River_Cruise
 - Guided_Walk
 - Mining_Plant
 - Historical_Building
 - Church
 - Cathedral
 - Shipwreck
 - Pony_Trekking
 - Colonial_Building
 - Church
 - Cathedral
 - Boat_Cruise
 - Educational_Tour
 - Commercial_Building
 - Wildlife_Watching
 - Bird_Watching
 - Bungee_Jumping
 - Safari
 - Scenic_Safari
 - Walking_Safari
 - Boat_Safari
 - Photographic_Safari
 - Elephant_Back_Safari
 - Fishing_Safari
 - Camping_Safari
 - Ballooning_Safari
 - Hunting

-
- Camping
 - Mountain_Walk
 - Destination
 - Lake
 - Urban_Area
 - Country
 - Town
 - City
 - Capital
 - Archipelago
 - Beach
 - Desert
 - Waterhole
 - Rural_Area
 - Reserve
 - Rehabilitation_Camp
 - Park
 - National_Park
 - Farmland
 - Forest
 - Fall
 - Family_Destination
 - Lagoon
 - Island
 - Retiree_Destination
 - River
 - Backpacker_Destination
 - Mountain
 - Delta
 - Budget_Hotel_Destination
 - Plain
 - Quiet_Destination
 - Chef
 - Travel_Class_3
 - Food
 - Diet_Food
 - Spicy
 - Drink
 - Coffee
 - Wine
 - Beer
 - Finger_Food
 - Snack
 - Organic
 - Seafood
 - Vegetarian
 - Animal

- Invertebrate
 - Insect
 - Spider
 - Orb_Web_Spider
 - Long_Tongued_Fly
 - Butterfly
 - Centipede
 - Dung_Beetle
 - Scorpion
- Vertebrate
 - Amphibian
 - Mammal
 - Primate
 - Monkey
 - Carnivore
 - Panther
 - Jackal
 - Black_Backed_Jackal
 - Cheetah
 - Tiger
 - Leopard
 - Hyena
 - Brown_Hyena
 - Lion
 - Black_Mane_Lion
 - Herbivore
 - Giraffe
 - Boar
 - Wild_Boar
 - Mongoose
 - Wallaby
 - Warthog
 - Antelope
 - Water_Buck
 - Kafue_Lechue
 - Sable
 - Bontebok
 - Kudu
 - Roan
 - Springbok
 - Eland
 - Gembsbok
 - Duiker
 - Oryx
 - Klipspringer
 - Sheep
 - Rhino

- White_Rhino
 - Black_Rhino
 - Squirrel
 - Ground_Squirrel
 - Suricate
 - Zebra
 - Mountain_Zebra
 - Lemur
 - Porcupine
 - Elephant
 - White_Elephant
 - Impala
 - Black_Faced_Impala
 - Honey_Badger
 - Deer
 - Gnu
 - Wildebeest
 - Calf
 - Baboon
 - Hippo
 - Buffalo
 - Cape_Buffalo
- Aquatic
 - Dolphin
 - Bottlenose
 - Whale
- Fish
 - Oyster
 - Vundu
 - Tuna
 - Yellow_Tuna
 - Kapenta
 - Bream
 - Chessa
 - Shark
 - White_Shark
 - Nkupe
 - Emperor
 - Barracuda
 - Cornish_Jack
 - Tiger_Fish
 - Seahorse
 - Marlin
 - Blue_Marlin
 - Black_Marlin
- Reptile
 - Tortoise

- Karoo_Tortoise
- Lizard
- Cocodrile
- Bird
 - Flamingo
 - Egret
 - Little_Egret
 - Owl
 - Bustard
 - Stork
 - Black_Stork
 - Peacock
 - White_Peacock
 - Pigeon
 - Pink_Pigeon
 - Heron
 - Grey_Heron
 - Goose
 - Pygmy_Goose
 - Weaver
 - Skimmer
 - Penguin
 - Bird_of_Prayer
 - Jackal_Buzzard
 - Falcon
 - Pygmy_Falcon
 - Vulture
 - White_Backed_Vulture
 - Kingfisher
 - Pied_Kingfisher
 - Pelican
 - Eagle
 - Black_Eagle
 - Fish_Eagle
- Vegetation
 - Aloe
 - Fynbos
 - Tree
 - Jacaranda
 - Acacia
 - Mopane
 - Tropical_Fruit
 - Banyan
 - Quiver
 - Palm
 - Camel_Thorn
 - Flower

- Daisy

3. Consultas y Resultados de la Recuperación

| ID_consulta | Consulta |
|-------------|--|
| qE1 | Where can I visit medieval buildings? |
| qE2 | Where can I visit a modern art museum? |
| qE3 | What city has famous gardens? |
| qE4 | What city can I visit famous castles? |
| qE5 | Where can I do a guided tour? |
| qE6 | What city can I visit famous castles? |
| qE7 | What city offers Unesco world heritage excursions? |
| qE8 | What city has a cabaret? |
| qE9 | Where can I visit a famous cathedral? |
| qE10 | Where can I see famous paintings? |
| qE11 | Where can I do a train excursion? |
| qE12 | Where can I do a day trip? |

Tabla D.1. Consultas Analizadas en el Dominio Europa-Travel

| ID_consulta | Consulta |
|-------------|---|
| qA1 | Where can I find a caravan site? |
| qA2 | Where can I do a mount resort reservation? |
| qA3 | What Island can I do angling? |
| qA4 | Where can I drink wine? |
| qA5 | Where can I eat seafood in Africa? |
| qA6 | Where can I visit a museum? |
| qA7 | Where can I watch antelopes? |
| qA8 | What city has gardens? |
| qA9 | Where can I watch cheetahs? |
| qA10 | Where can I do bird watching? |
| qA11 | What river can I take rafting classes? |
| qA12 | Where can I do gorge swing? |
| qA13 | What cities offer national park excursions? |
| qA14 | Where can I visit a cathedral? |
| qA15 | Where can I do a ballooning safari? |
| qA16 | What city offers shipwreck excursions? |
| qA17 | Where can I do tiger fishing? |

Tabla D.2. Consultas Analizadas en el Dominio Africa-Travel

| q _i | d _k | | | | | | | | | | | | | | | m _j |
|----------------|----------------|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----------------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | |
| qE1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | | | | | | 6 |
| qE2 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | | | | | | 6 |
| qE3 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | | | | | | | | 7 |
| qE4 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | | | | | 8 |
| qE5 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | | | | 4 |
| qE6 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | 11 |
| qE7 | 1 | 1 | 1 | 1 | 1 | | | | | | | | | | | 5 |
| qE8 | 1 | 0 | 1 | 1 | | | | | | | | | | | | 3 |
| qE9 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | | | | | | 6 |
| qE10 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | | 7 |
| qE11 | 1 | 1 | | | | | | | | | | | | | | 2 |
| qE12 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | 12 |

Tabla D.3. Resultados Cuantitativos de la Búsqueda en el Dominio Europa-Travel

| q _j | d _k | | | | | | | | | | | | | | | | m _j |
|----------------|----------------|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----------------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | |
| qA1 | 1 | | | | | | | | | | | | | | | | 1 |
| qA2 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | | | | | | | | | 8 |
| qA3 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 16 |
| qA4 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | | | | | | | 6 |
| qA5 | 1 | | | | | | | | | | | | | | | | 1 |
| qA6 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | | | | | 8 |
| qA7 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 16 |
| qA8 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | | | | | | 7 |
| qA9 | 0 | 1 | 1 | 1 | 1 | 1 | | | | | | | | | | | 6 |
| qA10 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | | | | | 8 |
| qA11 | 0 | 0 | 1 | 1 | 1 | | | | | | | | | | | | 5 |
| qA12 | 1 | 0 | 0 | 1 | | | | | | | | | | | | | 4 |
| qA13 | 1 | 1 | | | | | | | | | | | | | | | 2 |
| qA14 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | | | | 9 |
| qA15 | 1 | | | | | | | | | | | | | | | | 1 |
| qA16 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | | | | | | 11 |
| qA17 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | | | | 13 |

Tabla D.4. Resultados Cuantitativos de la Búsqueda en el Dominio Africa-Travel

| q _j | d _k | | | | | | | | | | | | | | |
|----------------|----------------|---|---|---|---|---|---|---|---|----|----|----|----|----|----|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| qE1 | 1 | 2 | 3 | 4 | 5 | 6 | | | | | | | | | |
| qE2 | 1 | 2 | 3 | 4 | 5 | 6 | | | | | | | | | |
| qE3 | 1 | 2 | 3 | 4 | 5 | 6 | 6 | 7 | | | | | | | |
| qE4 | 1 | 2 | 3 | 3 | 3 | 3 | 4 | 5 | 6 | 7 | 8 | | | | |
| qE5 | 1 | 2 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 4 | | | |
| qE6 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | | | | |
| qE7 | 1 | 2 | 3 | 4 | 5 | | | | | | | | | | |
| qE8 | 1 | 1 | 2 | 3 | | | | | | | | | | | |
| qE9 | 1 | 2 | 3 | 4 | 5 | 6 | | | | | | | | | |
| qE10 | 1 | 2 | 2 | 2 | 3 | 4 | 4 | 4 | 4 | 4 | 4 | 5 | 6 | 7 | |
| qE11 | 1 | 2 | | | | | | | | | | | | | |
| qE12 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | | | |

Tabla D.5. Valores R_{kj} del Dominio Europa-Travel

| q _j | d _k | | | | | | | | | | | | | | | |
|----------------|----------------|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| qA1 | 1 | | | | | | | | | | | | | | | |
| qA2 | 1 | 2 | 3 | 3 | 3 | 3 | 3 | 4 | | | | | | | | |
| qA3 | 1 | 2 | 3 | 4 | 5 | 6 | 6 | 6 | 6 | 7 | 7 | 7 | 7 | 7 | 8 | |
| qA4 | 1 | 2 | 3 | 4 | 5 | 6 | | | | | | | | | | |
| qA5 | 1 | | | | | | | | | | | | | | | |
| qA6 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | | | | | | | | |
| qA7 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| qA8 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | | | | | | | | | |
| qA9 | 0 | 1 | 2 | 3 | 4 | 5 | | | | | | | | | | |
| qA10 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | | | | | | | | |
| qA11 | 0 | 0 | 1 | 2 | 3 | | | | | | | | | | | |
| qA12 | 1 | 1 | 1 | 2 | | | | | | | | | | | | |
| qA13 | 1 | 2 | | | | | | | | | | | | | | |
| qA14 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | | | | | | | |
| qA15 | 1 | | | | | | | | | | | | | | | |
| qA16 | 1 | 2 | 3 | 4 | 5 | 5 | 5 | 5 | 6 | 7 | 8 | | | | | |
| qA17 | 1 | 2 | 3 | 4 | 5 | 6 | 6 | 6 | 7 | 8 | 9 | 10 | 11 | | | |

Tabla D.6. Valores R_{kj} del Dominio África-Travel

| q_j | d_k | | | | | | | | | | | | | | |
|-------|-------|------|------|------|------|------|------|------|------|-----|------|------|------|-----|----|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| qE1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | | | | | |
| qE2 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | | | | | |
| qE3 | 1 | 1 | 1 | 1 | 1 | 1 | 0,86 | 0,88 | | | | | | | |
| qE4 | 1 | 1 | 1 | 0,75 | 0,6 | 0,5 | 0,57 | 0,63 | 0,67 | 0,7 | 0,73 | | | | |
| qE5 | 1 | 1 | 1 | 0,75 | 0,6 | 0,5 | 0,43 | 0,38 | 0,33 | 0,3 | 0,27 | 0,33 | | | |
| qE6 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | |
| qE7 | 1 | 1 | 1 | 1,00 | 1,00 | | | | | | | | | | |
| qE8 | 1 | 0,50 | 0,67 | 0,75 | | | | | | | | | | | |
| qE9 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | | | | | |
| qE10 | 1 | 1 | 0,67 | 0,5 | 0,6 | 0,67 | 0,57 | 0,5 | 0,44 | 0,4 | 0,36 | 0,42 | 0,46 | 0,5 | |
| qE11 | 1 | 1 | | | | | | | | | | | | | |
| qE12 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | |

Tabla D.7. Valores $Precision(R_{jk})$ del Dominio Europa-Travel

| q_j | d_k | | | | | | | | | | | | | | | |
|-------|-------|-----|------|------|------|------|------|------|------|-----|------|------|------|------|------|------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| qA1 | 1 | | | | | | | | | | | | | | | |
| qA2 | 1 | 1 | 1 | 0,75 | 0,60 | 0,50 | 0,43 | 0,50 | | | | | | | | |
| qA3 | 1 | 1 | 1 | 1 | 1 | 1 | 0,86 | 0,75 | 0,67 | 0,6 | 0,64 | 0,58 | 0,54 | 0,5 | 0,47 | 0,50 |
| qA4 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | | | | | | |
| qA5 | 1 | | | | | | | | | | | | | | | |
| qA6 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | | | | |
| qA7 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0,89 | 0,9 | 0,91 | 0,92 | 0,92 | 0,93 | 0,93 | 0,94 |
| qA8 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | | | | | |
| qA9 | 0 | 0,5 | 0,67 | 0,75 | 0,80 | 0,83 | | | | | | | | | | |
| qA10 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | | | | |
| qA11 | 0 | 0 | 0,33 | 0,50 | 0,6 | | | | | | | | | | | |
| qA12 | 1 | 0,5 | 0,33 | 0,50 | | | | | | | | | | | | |
| qA13 | 1 | 1 | | | | | | | | | | | | | | |
| qA14 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | | | |
| qA15 | 1 | | | | | | | | | | | | | | | |
| qA16 | 1 | 1 | 1 | 1 | 1 | 0,83 | 0,71 | 0,63 | 0,67 | 0,7 | 0,73 | | | | | |
| qA17 | 1 | 1 | 1 | 1 | 1 | 1 | 0,86 | 0,75 | 0,78 | 0,8 | 0,82 | 0,83 | 0,85 | | | |

Tabla D.8. Valores $Precision(R_{jk})$ del Dominio Africa-Travel

| q _j | d _k | | | | | | | | | | | | | | |
|----------------|----------------|------|------|------|------|------|------|------|------|------|------|------|------|----|----|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| qE1 | 0,17 | 0,33 | 0,50 | 0,67 | 0,83 | 1 | | | | | | | | | |
| qE2 | 0,17 | 0,33 | 0,50 | 0,67 | 0,83 | 1 | | | | | | | | | |
| qE3 | 0,14 | 0,29 | 0,43 | 0,57 | 0,71 | 0,86 | 0,86 | 1 | | | | | | | |
| qE4 | 0,13 | 0,25 | 0,38 | 0,38 | 0,38 | 0,38 | 0,50 | 0,63 | 0,75 | 0,88 | 1 | | | | |
| qE5 | 0,25 | 0,50 | 0,75 | 0,75 | 0,75 | 0,75 | 0,75 | 0,75 | 0,75 | 0,75 | 0,75 | 1 | | | |
| qE6 | 0,09 | 0,18 | 0,27 | 0,36 | 0,45 | 0,55 | 0,64 | 0,73 | 0,82 | 0,91 | 1 | | | | |
| qE7 | 0,20 | 0,40 | 0,60 | 0,80 | 1 | | | | | | | | | | |
| qE8 | 0,33 | 0,33 | 0,67 | 1 | | | | | | | | | | | |
| qE9 | 0,17 | 0,33 | 0,50 | 0,67 | 0,83 | 1 | | | | | | | | | |
| qE10 | 0,14 | 0,29 | 0,29 | 0,29 | 0,43 | 0,57 | 0,57 | 0,57 | 0,57 | 0,57 | 0,57 | 0,71 | 0,86 | 1 | |
| qE11 | 0,50 | 1 | | | | | | | | | | | | | |
| qE12 | 0,08 | 0,17 | 0,25 | 0,33 | 0,42 | 0,50 | 0,58 | 0,67 | 0,75 | 0,83 | 0,92 | 1 | | | |

Tabla D.9. Valores $Precision(R_{jk})$ del Dominio Europa-Travel

| q _j | d _k | | | | | | | | | | | | | | | |
|----------------|----------------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|----|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| qA1 | 1 | | | | | | | | | | | | | | | |
| qA2 | 0,25 | 0,50 | 0,75 | 0,75 | 0,75 | 0,75 | 0,75 | 1 | | | | | | | | |
| qA3 | 0,13 | 0,25 | 0,38 | 0,50 | 0,63 | 0,75 | 0,75 | 0,75 | 0,75 | 0,75 | 0,88 | 0,88 | 0,88 | 0,88 | 0,88 | 1 |
| qA4 | 0,17 | 0,33 | 0,50 | 0,67 | 0,83 | 1 | | | | | | | | | | |
| qA5 | 1 | | | | | | | | | | | | | | | |
| qA6 | 0,13 | 0,25 | 0,38 | 0,50 | 0,63 | 0,75 | 0,88 | 1 | | | | | | | | |
| qA7 | 0,07 | 0,13 | 0,20 | 0,27 | 0,33 | 0,40 | 0,47 | 0,53 | 0,53 | 0,60 | 0,67 | 0,73 | 0,80 | 0,87 | 0,93 | 1 |
| qA8 | 0,14 | 0,29 | 0,43 | 0,57 | 0,71 | 0,86 | 1,00 | | | | | | | | | |
| qA9 | 0 | 0,20 | 0,40 | 0,60 | 0,80 | 1 | | | | | | | | | | |
| qA10 | 0,13 | 0,25 | 0,38 | 0,50 | 0,63 | 0,75 | 0,88 | 1 | | | | | | | | |
| qA11 | 0,00 | 0,00 | 0,33 | 0,67 | 1 | | | | | | | | | | | |
| qA12 | 0,50 | 0,50 | 0,50 | 1,00 | | | | | | | | | | | | |
| qA13 | 0,50 | 1,00 | | | | | | | | | | | | | | |
| qA14 | 0,11 | 0,22 | 0,33 | 0,44 | 0,56 | 0,67 | 0,78 | 0,89 | 1 | | | | | | | |
| qA15 | 1 | | | | | | | | | | | | | | | |
| qA16 | 0,13 | 0,25 | 0,38 | 0,50 | 0,63 | 0,63 | 0,63 | 0,63 | 0,75 | 0,88 | 1,00 | | | | | |
| qA17 | 0,09 | 0,18 | 0,27 | 0,36 | 0,45 | 0,55 | 0,55 | 0,55 | 0,64 | 0,73 | 0,82 | 0,91 | 1 | | | |

Tabla D.10. Valores $Precision(R_{jk})$ del Dominio Africa-Travel

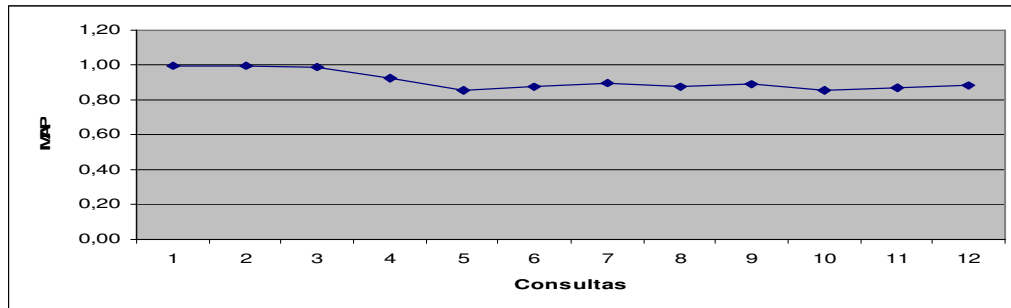


Figura D.1. Valores del MAP en función del número de consultas para Europa

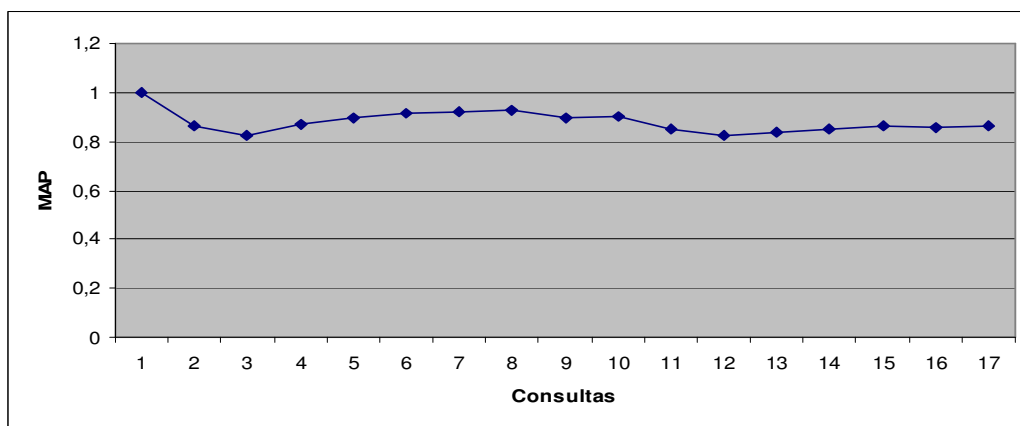


Figura D.2. Valores del MAP en función del número de consultas para África

Ontologías: Herramientas y Lenguajes

1. Herramientas de Gestión de Ontologías

En los últimos años el número de entornos y herramientas para la construcción de ontologías ha crecido significativamente. El objetivo de estas herramientas es proveer soporte para el proceso de desarrollo y subsiguiente uso de ontologías. En esta sección se tratará de proveer una visión amplia de los entornos y herramientas disponibles para el desarrollo de ontologías, ya sea desde cero o haciendo uso de otras ontologías. Se analizaron las herramientas de gestión ontológica más relevantes entre las que se encuentran *Ontolingua*, *Ontosaurus*, *WebOnto*, *WebODE*, *OntoEdit* y particularmente la utilizada en esta arquitectura: *Protégé*.

Apollo es una aplicación de desarrollo de ontologías amigable al usuario diseñada a partir de las experiencias de los creadores (Knowledge Media Institute – The Open University) en sus trabajos con socios industriales que deseaban utilizar técnicas de modelado de conocimiento, pero requerían un entorno y sintaxis fácil de usar y de comprender (Apollo, 2008).

LinKFactory es un sistema de gestión de ontologías formal desarrollado por Language and Computing, diseñado para construir y gestionar ontologías formales, independientes del lenguaje, muy grandes y complejas (Ceusters y col., 2001).

OILEd es un editor de ontologías gráfico desarrollado por la Universidad de Manchester que permite al usuario construir ontologías usando DAML+OIL (Bechhofer y col., 2001a). La tarea principal a la cual apunta OILEd es la edición de ontologías o esquemas, en oposición a la adquisición del conocimiento o la construcción de grandes bases de conocimiento de instancias. Un aspecto clave del comportamiento de OILEd es el uso del razonador FaCT para clasificar ontologías y para verificar consistencia a través de la traducción desde DAML+OIL a la lógica de descripción SHIQ (Horrocks y col., 1999). Esto permite al usuario describir las clases de su ontología y dejar que el razonador determine el lugar apropiado en la jerarquía para dicha definición.

OntoEdit es un entorno de ingeniería ontológica que da soporte al desarrollo y al mantenimiento de ontologías usando medios gráficos (Sure y col., 2002). Este paradigma soporta modelado con lenguaje de representación neutral tanto como sea posible para conceptos, relaciones y axiomas. Varias de las vistas gráficas de las estructuras contenidas en la ontología soportan el modelado de diferentes fases del ciclo de construcción de una ontología.

El *Ontolingua Server* es un conjunto de herramientas y servicios que dan soporte a la construcción de ontologías compartidas entre grupos distribuidos (Farquhar y col., 1996). Se desarrolló en el Knowledge Systems Laboratory (KSL) de la Universidad de Stanford. Inicialmente, el módulo principal del Ontolingua Server era el editor de ontología, luego se agregaron otros módulos al entorno, tales como el Webster, solucionador de ecuaciones, un servidor OKBC (Open Knowledge Based Connectivity), Chimaera (una herramienta de mezclado de ontología), traductores de lenguajes (Prolog, CORBA, IDL, CLIPS, Loom), etc (McGuinness y col., 2000). Editores remotos pueden visualizar y editar ontologías, y aplicaciones remotas o locales pueden acceder a cualquiera de las ontologías en la biblioteca de ontologías mediante el protocolo OKBC (Chaudhri y col., 1998).

Ontosaurus fue desarrollado por el Information Sciences Institute (ISI) en la Universidad de South California. OntoSaurus consiste de dos módulos: un servidor ontológico, que usa Loom como su sistema de representación de

conocimiento y un visualizador para ontologías que crea dinámicamente páginas HTML (incluyendo documentación textual y gráfica) que visualiza la jerarquía ontológica (Swartout y col., 1997). La ontología puede ser editada mediante formularios HTML, y presenta traductores desde Loom a Ontolingua, KIF, KRSS and C++. Las ontologías OntoSaurus pueden también ser accedidas con el protocolo OKBC.

KnoME es un conjunto de varias herramientas para el desarrollo colaborativo de ontologías en el lenguaje de modelado conceptual GRAIL (Rector y col., 1997). TIGGER es una parte importante de este conjunto, desarrollado para la rápida adquisición de conocimiento de expertos de dominio que no poseen conocimiento de desarrollo ontológico. Juntas, estas herramientas están disponibles libremente como OpenKnoME (Rogers y col., 2001). Fueron desarrolladas en la Universidad de Manchester sobre varios programas de ontologías médicas y farmacéuticas grandes, incluyendo GALEN (1992-1995), GALEN-IN-USE (1996-1999) y PRODIGY (1999-2001).

SymOntoX (Symbolic Ontology Manager XML) es un software prototipo para la gestión de ontologías del dominio de negocios (Missikoff y col., 2002). Ha sido desarrollado por el LEKS (Laboratory for Enterprise Knowledge and Systems) en IASI-CNR. SymOntoX soporta una interfaz de usuario gráfica basada en formularios para las funcionalidades de edición, visualización y diagramación del contenido de ontologías.

WebODE ha sido desarrollado por el Grupo de Ingeniería Ontológica (GIO) del Departamento de Inteligencia Artificial de la Facultad de Informática de la Universidad Politécnica de Madrid (Arpirez y col., 2001). WebODE es un banco de trabajo de ingeniería ontológica que provee una variedad de servicios relacionados a las ontologías, cubre y da soporte a la mayoría de las actividades involucradas en el proceso de desarrollo y uso de una ontología. Está construido sobre la base de un servidor de aplicaciones que provee una alta usabilidad y extensibilidad permitiendo una fácil adición de nuevos servicios y la utilización de los existentes. Las ontologías WebODE están representadas usando un modelo de conocimiento expresivo basado en el conjunto de referencia de

representaciones intermedias de la metodología METHONTOLOGY, el cual incluye componentes ontológicos tales como conceptos (con atributos de clases e instancias), particiones, relaciones binarias ad-hoc, relaciones predefinidas (taxonómicas y parte-todo), instancias, axiomas, reglas, constantes y referencias bibliográficas (Fernández-López y col., 1999).

WebOnto es una herramienta desarrollada por el KMI (Knowledge Media Institute) de la Open University (Domingue, 1998). Soporta la visualización colaborativa, la creación y edición de ontologías, las cuales están representadas en el lenguaje de modelado de conocimiento OCML. Sus características principales son: gestión de ontologías usando una interfaz gráfica; la generación automática de formularios de edición de instancias; soporte para el modelado de procesos y tareas, inspección de elementos tomando en cuenta la herencia de las propiedades y el chequeo de consistencia y soporte para trabajo colaborativo mediante técnicas de “broadcast” y recepción y haciendo uso de anotaciones.

1.1 Protégé

Protégé es la última herramienta en una línea establecida de herramientas desarrolladas por la Universidad de Stanford para la adquisición del conocimiento (Noy y col., 2000; 2001). Protégé tiene miles de usuarios en todo el mundo que usan el sistema para proyectos que varían desde guías para el modelado de protocolos para el cáncer hasta el modelado de plantas de energía nuclear. Protégé está libremente disponible para ser bajado bajo la licencia open-source de Mozilla.

Protégé provee un entorno de desarrollo basado en conocimiento y diseño ontológico gráfico e interactivo. Ayuda a los ingenieros de conocimiento y expertos de dominio a realizar tareas de gestión del conocimiento. Los desarrolladores de ontologías pueden acceder a información relevante rápidamente cuando la necesiten, y pueden usar manipulación directa para

navegar y gestionar una ontología. Los controles tipo árbol permiten una navegación simple y rápida a través de la jerarquía de clases. Protégé usa formularios para completar los valores en los campos.

El modelo de conocimiento de Protégé es compatible con OKBC. Incluye el soporte para clases y la jerarquía de clases con herencia múltiple, plantillas y campos propios; especificación de facetas pre-definidas y arbitrarias para campos, los cuales incluyen valores permitidos, restricciones de cardinalidad, valores por defecto, campos inversos, metaclases y jerarquías de metaclases.

En adición a una interfaz altamente usable, otras dos importantes características distinguen a Protégé de la mayoría de los entornos de edición de ontologías: es escalable y extensible.

Los desarrolladores han usado exitosamente Protégé para construir ontologías con más de 150.000 conceptos. Bases de conocimiento de soporte de miles de marcos incluyen dos componentes:

- i.* una base de datos para el almacenamiento y consulta de los datos
- ii.* y un mecanismo de “cacheo” para permitir la carga de nuevos marcos en la memoria cuando se ha excedido.

Una de las mayores ventajas de la arquitectura Protégé es que el sistema está construido de una forma modular y abierta. Su arquitectura basada en componentes permite a los constructores agregar nuevas funcionalidades creando las extensiones (plugins) apropiadas.

La mayoría de las extensiones caen en una de tres categorías:

- i.* “backends”, que le permiten a los usuarios almacenar e importar bases de conocimiento en varios formatos;

- ii. “spot widgets”, los cuales se usan para mostrar y editar valores de campos o su combinación para tareas específicas o específicas de dominio y,
- iii. “tab plugins”, las cuales son aplicaciones basadas en conocimiento usualmente muy relacionadas con las bases de conocimiento Protégé.

En conclusión, Protégé tiene varias características que lo diferencian de otras herramientas de edición de ontologías (entorno gráfico e intuitivo, plataforma escalable, extensible, modular y abierta) y que lo hacen la herramienta de edición ideal para la gestión de las ontologías de dominio en la arquitectura Onto-DOM.

2. Lenguajes de Definición de Ontologías

Además de la plataforma de desarrollo de la ontología es necesario seleccionar un lenguaje particular para definirla. En los principios de los 90 se desarrolló un conjunto de lenguajes de implementación de ontologías basados en Inteligencia Artificial. Básicamente el paradigma de Recuperación de Conocimiento subyacente en esos lenguajes era la lógica de primer orden (por ejemplo KIF), en marcos combinados con lógica de primer orden (por ejemplo, Ontolingua, OCML y FLogic) o en lógica descriptiva (por ejemplo, Loom). Se describen en esta sección los diferentes lenguajes de definición ontológica poniendo especial énfasis en OWL.

KIF es un lenguaje basado en lógica de primer orden creado en 1992 como un formato de intercambio para diversos sistemas de recuperación de conocimiento (Genesereth y Fikes, 1992). Ontolingua creado en base a KIF, fue desarrollado en 1992 por el Knowledge Systems Laboratory (KSL) de la Universidad de Stanford (Farquhar, y col., 1996; Gruber, 1992). Es uno de los

lenguajes más expresivos que se han usado para la representación de ontologías, permitiendo la representación de conceptos, taxonomías de conceptos, relaciones n-arias, funciones, axiomas, instancias y procedimientos. Su alta expresividad conduce a dificultades en la construcción de mecanismos de razonamiento. Por lo tanto, este lenguaje no provee soporte para razonamiento.

Loom fue desarrollado al mismo tiempo con Ontolingua en el Information Science Institute (ISI) en la Universidad de South California (MacGregor, 1991). Inicialmente, no tenía como objetivo la implementación de ontologías sino bases de conocimiento en general. Loom está basado en Lógica Descriptiva y reglas de producción y provee clasificación automática de conceptos.

OCML fue desarrollado en 1993 en el KMI de la Open University (Motta, 1999). Fue creado como una especie de Ontolingua “operacional”. De hecho, la mayoría de las definiciones que pueden expresarse en OCML son similares a las correspondientes definiciones en Ontolingua, y pueden definirse componentes adicionales: reglas de producción y deductivas y definiciones operacionales para funciones. OCML fue construido para el desarrollo de ontologías y modelos ejecutables en los métodos de resolución de problemas.

FLogic fue desarrollado en 1995 en la Karlsruhe University (Kifer y col., 1995). FLogic (Frame Logic) combina marcos y lógica de primer orden, permitiendo representar conceptos, taxonomías de conceptos, relaciones binarias, funciones, instancias, axiomas y reglas deductivas. Su motor de inferencias, Ontobroker, puede ser usado para chequeo de restricciones y deducción de información (Decker y col., 1999).

El auge de Internet condujo a la creación de lenguajes ontológicos que explotan las características de la Web. Tales lenguajes usualmente se denominan “lenguajes de ontologías basados en web” o “lenguajes de marcado ontológico”.

SHOE fue creado como una extensión de HTML en la Universidad de Maryland (Luke y Heflin, 2000). Usa diferentes etiquetas de las que usa la especificación de HTML permitiendo, por lo tanto, la inserción de ontologías en un documento HTML. SHOE combina marcos y reglas permitiendo la

representación de conceptos, sus taxonomías, relaciones n-arias, instancias y reglas de deducción, las cuales son usadas por su motor de inferencias para obtener nuevos elementos.

Luego fue creado *XML* ampliamente adoptado como un lenguaje Standard para el intercambio de información en la Web (Bray y col., 2000). Como consecuencia, la sintaxis de SHOE fue modificada para usar XML y luego otros lenguajes de ontologías se desarrollaron con las sintaxis de XML.

XOL fue desarrollado por el centro de Inteligencia Artificial del SRI internacional en 1999 (Karp y col., 1999). Es un lenguaje bastante restringido donde sólo pueden especificarse conceptos, taxonomías de conceptos y relaciones binarias. No incluye mecanismos de inferencia ya que fue desarrollado principalmente para el intercambio de ontologías en el dominio biomédico.

RDF fue desarrollado por la W3C (World Wide Web Consortium) como un lenguaje basado en redes semánticas para describir recursos Web (Lassila y Swick, 1999). *RDF Schema* fue creado por la W3C como una extensión de *RDF* con primitivas basadas en marcos (Lassila y Swick, 1999). La combinación de ambos (*RDF* y *RDF Schema*) normalmente se conoce como *RDF(S)*. *RDF(S)* no es muy expresivo permitiendo solamente la representación de conceptos, taxonomías de conceptos y relaciones binarias. Se han creado algunos motores de inferencia para este lenguaje, principalmente para el chequeo de restricciones. Estos lenguajes han establecido las bases para la Web Semántica. En este contexto, otros tres lenguajes se desarrollaron como extensiones de *RDF(S)*: *OIL*, *DAML+OIL* y *OWL*.

OIL fue desarrollado en el marco del proyecto On-To-Knowledge del European IST (Horrocks y col.; 2000). Adiciona a *RDF(S)* primitivas de recuperación de conocimiento basada en marcos y su semántica formal está basada en Lógica Descriptiva. En el 2000 fue creado *DAML+OIL* que agrega primitivas de recuperación de conocimiento basadas en lógica descriptiva y que permite la representación de conceptos, taxonomías, relaciones binarias,

funciones e instancias (Horrocks y van Harmelen, 2001). Finalmente, en el 2001 surgió un nuevo lenguaje de marcado ontológico para la Web Semántica, llamado OWL (Web Ontology Language).

2.1. OWL

OWL extiende RDF(S) para permitir la expresión de relaciones complejas entre diferentes clases RDF(S), y mayor precisión en las restricciones de clases y propiedades específicas (OWL, 2004) . Esto incluye por ejemplo:

- i.* los recursos para limitar las propiedades de clases con respecto a número y tipo,
- ii.* los recursos para inferir qué elementos que tienen varias propiedades son miembros de una clase en particular
- iii.* los recursos para determinar si todos los miembros de una clase tendrán una propiedad en particular, o si puede ser que sólo algunos la tengan
- iv.* los recursos para distinguir entre relaciones uno-a-uno, varios-a-uno o uno-a-varios, permitiendo que las "claves externas" de las bases de datos puedan representarse en una ontología
- v.* los recursos para expresar relaciones entre clases definidas en diferentes documentos en la Web
- vi.* los recursos para construir nuevas clases a partir de uniones, intersecciones y complementos de otras, y
- vii.* los recursos para restringir rangos y dominios para especificar combinaciones de clases y propiedades.

OWL es considerado una de las tecnologías fundamentales que subyacen a la definición ontológica y ha atraído el interés tanto del área académica como

comercial (Geroimenko y Chen, 2005). Lo anterior, aunado a que es un lenguaje estandarizado, altamente expresivo y con un soporte continuo, hacen de OWL el candidato ideal para la definición de la ontología de dominio incluida en la Memoria Organizacional.

3. Herramientas de Anotado Basadas en Ontologías

Puede definirse una herramienta de anotado basada en ontologías como una herramienta que hace uso de una ontología pre-existente para insertar un marcado en una página web u otro documento y/o son usadas para la población de bases de conocimiento a través de la mediación de marcado de documentos.

Existen dos filosofías principales detrás de las herramientas de anotado ontológicas. En primer lugar, lo que podría llamarse el enfoque de Web Semántica, ve la producción de páginas anotadas como prioridad y la población de instancias como secundaria. En segundo lugar, lo que puede llamarse el enfoque de ingeniería de conocimiento, ve la anotación guiada por ontología de documentos como un medio para la población de bases de conocimiento como también para la producción de documentos anotados. En esencia ambas filosofías producen herramientas similares. Sin embargo, la diferencia en el énfasis es importante ya que es probable que dependiendo del objetivo del usuario una y otra herramienta será la apropiada.

AeroDAML fue desarrollado como parte del proyecto UML Based Ontology Toolset (UBOT) (Kogut y Holmes, 2001). De acuerdo a Kogut y Holmes, *AeroDAML* es una herramienta de marcado de conocimiento que aplica técnicas de extracción de información del lenguaje natural para generar automáticamente anotaciones DAML a partir de páginas web. En su versión web la herramienta utiliza una ontología predefinida para el anotado. La anotación es en DAML y es visualizada como una página web. La herramienta es extremadamente simple de usar pero de aplicabilidad limitada.

COHSE (Conceptual Open Hypermedia Services Environment) fue desarrollado en el grupo de Gestión de la Información de la Universidad de Manchester y en el grupo de Inteligencia, Agentes y Multimedia de la Universidad de Southampton (Bechhofer y col., 2001b). El objetivo de este conjunto de herramientas es el uso de metadatos para dar soporte a la construcción y navegación de enlaces en la Web Semántica. Se apoya en tres tipos de servicios: servicios de razonamiento ontológicos, servicio de enlace hipertexto abierto basado en web (que provee facilidades de enlace) y la integración de ambos servicios para formar un sistema hipertexto conceptual. Esto permite que los documentos sean enlazados a través de metadatos describiendo sus conceptos. La anotación con *COHSE* es manual y automática. La herramienta primero extrae anotaciones automáticamente, y luego requiere que el usuario agregue sus propias anotaciones. El sistema es fácil de aprender. Si el usuario no sabe cómo crear expresiones conceptuales complejas usando un visualizador conceptual, sólo será capaz de crear anotaciones para conceptos simples. Si deben crearse varias anotaciones el proceso es lento.

MnM es una herramienta desarrollada por el AKT Interdisciplinary Research Collaboration (Motta y col., 2002). Su propósito es la anotación de documentos usando un marcado derivado a partir de ontologías preexistentes. La herramienta trabaja mediante el anotado de un conjunto de entrenamiento de textos o documentos HTML y usando esto para generar reglas léxicas que pueden usarse para extraer información automáticamente desde otro conjunto de documentos. Las instancias derivadas de este proceso pueden usarse para poblar la ontología usada en la anotación. Soporta tanto anotación manual como semi automática gracias a la integración de tecnologías de extracción de información. Se desarrolló un prototipo inicial en Java que usa el módulo de extracción de información Amilcare desarrollado en la Universidad de Sheffield (Ciravegna, 2001).

La herramienta fue diseñada para poblar una ontología después de una fase de entrenamiento. Definiciones de posibles nuevas instancias se ofrecen al usuario derivadas de las páginas anotadas automáticamente para su corrección o

aceptación. El sistema es relativamente fácil de usar una vez que se comprende el proceso de entrenamiento básico.

OntoAnnotate es una herramienta desarrollada como una implementación básica para el entorno de anotado CREAM (Handschub y col., 2001). Es una herramienta semi-automática que permite recolectar el conocimiento de documentos y páginas web para crear una base de documentos semánticos. La herramienta es semi-automática hasta el punto que el anotador es asistido por diferentes visiones de una ontología y hechos, con el proceso de anotación como un proceso simple de acciones de “drag and drop”. El sistema es fácil de usar y posee una interfaz satisfactoria.

El *SHOE Knowledge Annotator* fue desarrollado por el grupo Parallel Understanding Systems de la Universidad de Maryland (Heflin y Hendler, 2001). Está disponible como un “applet” o una aplicación Java “standalone” que permite a los usuarios marcar páginas web con instancias SHOE. Permite la carga de varias ontologías SHOE y el anotado de los documentos con esas ontologías. El anotado con esta herramienta es manual y se realiza únicamente sobre documentos HTML estáticos. El sistema es fácil de usar si el usuario está familiarizado con SHOE. La herramienta es, en su mayoría, un prototipo con propósitos de “demo” y no tiene aplicaciones en el mundo real.

Teniendo en cuenta el análisis anterior, era evidente la necesidad de desarrollar un estrategia de anotado propia que se adecuase a la arquitectura de Memoria Organizacional propuesta.

Referencias

Abidi y col. (2005)

Abidi, S.; Cheah, Y. y Curran, J. "A Knowledge Creation Info-Structure to Acquire and Crystallize the Tacit Knowledge of Health-Care Experts". *IEEE Transactions on Information Technology in Biomedicine*. 2005, vol. 9, núm. 2, p.193-204.

Abdou y Savoy (2008)

Abdou, S. y Savoy, J. "Searching in MEDLINE: Query expansion and manual indexing evaluation". *Information Processing and Management*. Núm. 44 p. 781-789.

Ackerman (1994)

Ackerman, M.S. "Definitional and Contextual Issues in Organizational and Group Memories" en *Proceedings of Twenty-seventh IEEE Hawaii International Conference of System Sciences*. 1994, p.191-200.

Ackerman y Halverson (1998)

Ackerman M.S. y C. Halverson. "Considering an Organization's Memory" en *Proceedings of the ACM 1998 Conference on Computer Supported Cooperative Work*. 1998, p.39-48. Seattle, WA: ACM Press.

Ackoff (1989)

Ackoff, R. "From Data to Wisdom". *Journal of Applied Systems Analysis*. 1989, vol. 16, p.3-9.

Ackoff (1996)

Ackoff, R. "On Learning and Systems that Facilitate It". *Center for Quality of Management Journal*. 1996, vol. 5, núm. 2, p.27-35.

Agresti (2003)

Agresti, W.W. "Tailoring IT support to communities of practice". *IT Professional*. 2003, vol. 5, núm. 6, p.24- 28.

Ahlers y Weimer (2002)

Ahlers, J. y Weimer H. "Challenges in Interactive Visualization for Knowledge Management" en *Proceedings of the Sixth International Conference on Information Visualization*. 2002, p.367-371.

Alavi (1997)

Alavi, M. *KPMG Peat Marwick U.S.: One Giant Brain*. Harvard Business School (case), 9-297-108; 1997.

Alavi (2000)

Alavi, M.; *Managing Organizational Knowledge*; En Framing the Domains of IT Management; R.W. Zmud; Cincinnati; Ohio; Pinnaflex Educational Resources, Inc, 2000.

Ale y col. (2004a)

Ale, M.A.; Chiotti, O. y Galli, M.R.; "Agent-Supported Ontology-Based Knowledge Management System" en *Proceedings of ASIS 2004 (Simposio Argentino de Sistemas de Información) and JAIIO 2004 (33° Jornadas Argentinas de Informática e Investigación Operativa)*. 2004, p.11.

Ale y col. (2004b)

Ale, M.A.; Chiotti, O. y Galli, M.R. "A Distributed Knowledge Management Conceptual Model For Knowledge Organizations" en *Proceedings de las Jornadas Iberoamericanas de Ingeniería de Software e Ingeniería del Conocimiento (JIISIC 2004)*. 2004, Madrid, España. ISBN 84-688-8957-1.

Ale y col. (2005a)

Ale, M.A.; Chiotti, O. y Galli, M.R. "A Distributed Knowledge Management Conceptual Model for Knowledge Organizations". *ICFAI Journal of Knowledge Management*. ICFAI University Press. 2005, p.27-39.

Ale y col. (2005b)

Ale, M.A.; Chiotti, O. y Galli, M.R. "Semantic Document Indexing In Ontology - Driven Organizational Memories" en *Proceedings del XI Congreso Argentino de Ciencias de la Computación (CACIC)*. 2005, Concordia, Entre Rios, Argentina.

Ale y col. (2005c)

Ale, M.A.; Chiotti, O. y Galli, M.R. "A Strategy For Semantic Document Classification In An Ontology Driven Knowledge Management System" en *Proceedings de la IADIS International Conference*. 2005, Porto, Portugal, p.410-414. ISBN 972-8924-06-2.

Ale y col. (2006a)

Ale, M.A.; Chiotti, O. y Galli, M.R. "Semantic Document Representation in an Ontology-Driven Distributed Organizational Memory" en *Proceedings de JIISIC 2006*. 2006, Puebla, México, p.191-198. ISBN 970-94770-0-5.

Ale y col. (2006b)

Ale, M.A.; Chiotti, O. y Galli, M.R. "Gestión Del Conocimiento Empresarial En Organizaciones Emergentes: Un Enfoque Basado En Ontologías" en *Proceedings del 4° Simposio de Sociedad de la Información (JAIIO 2006)*. 2006 Mendoza, Argentina.

Ale y col. (2007a)

Ale, M.A.; Gerarduzzi, C.; Chiotti, O. y Galli, M.R. "Onto-Dom: A Question-Answering Ontology-Based Strategy For Heterogeneous Knowledge Sources" en *Proceedings de las VI Jornadas Iberoamericanas de Ingeniería de Software e Ingeniería del Conocimiento (JIISIC 2007)*. 2007, Lima, Perú, p.79-86. ISBN 978-9972-2885-1-7.

Ale y col. (2007b)

Ale, M.A.; Gerarduzzi, C.; Chiotti, O. y Galli, M.R. "Onto-Dom: A Question-Answering Ontology-Based Strategy For Heterogeneous Knowledge Sources".

ICFAI Journal of Knowledge Management, ICFAI University Press. 2007, p.54-68. ISSN 0972-9216.

Ale y col. (2008a)

Ale, M.A.; Chiotti, O. y Galli, M.R. Capítulo 8: "Enterprise Knowledge Management for Emergent Organizations: An Ontology Driven Approach" en *"Knowledge Management Strategies: A Handbook Of Applied Technologies"*, Editores: Lytras, M. D., Russ, M., Maier, R., Naeve, A.; Ed. IGI Publishing. 2008. ISBN 978-1-59904-603-7.

Ale y col. (2008b)

Ale, M.A.; Gerarduzzi, C.; Chiotti, O. y Galli, M.R. "Onto-DOM: Organizational Knowledge Sources Integration through an Ontology-Based Approach" en *Proceedings de la XXXIV Conferencia Latinoamericana de Informática*. 2008. p.569-578. ISBN 978-950-9770-02-7.

Ale y col. (2008c)

Ale, M.A.; Gerarduzzi, C.; Chiotti, O. y Galli, M.R. "Capítulo: Organizational Knowledge Sources Integration through an Ontology-Based Approach: The Onto-DOM Architecture" en *Emerging Technologies and Information Systems for the Knowledge Society*; Lecture Notes in Computer Science; M.D. Lytras et al. (Eds.). 2008. LNAI 5288, p.441-450; ISBN 978-3-540-87780-6.

Allee (2000)

Allee, V. "Knowledge Networks and Communities of Practice". *OD Practitioner*. 2000. Vol. 32.

Andersen (1996)

Andersen Arthur y The American Productivity and Quality Center; *The Knowledge Management Assessment Tool: External Benchmarking Version*. 1996.

Aparicio y col. (2005)

Aparicio, A.; Farias, O. y dos Santos, N. "Applying Ontologies in the Integration of Heterogeneous Relational Databases" [en línea] en *Conferences in Research and Practice in Information Technology*. 2005. [consultado en 2008]. Disponible en <<http://crpit.com/confpapers/CRPITV58Aparicio.pdf>>.

Apollo (2008)

Apollo User Guide [en línea]. 2008. [consultado en 2008]. Disponible en <<http://apollo.open.ac.uk>>.

Arpírez y col. (2001)

Arpírez, J.C.; Corcho, O.; Fernández-López M. y Gómez-Perez, A. "WebODE: a Scalable Ontological Engineering Workbench" en *Proceedings of the First International Conference on Knowledge Capture (KCAP'01)*. 2001. ACM Press, Victoria, Canada, p.6-13.

Badaracco (1991)

Badaracco, J.L. "The Knowledge Link". *Harvard Business School Press*. 1991. Boston, Massachussets.

Baeza-Yates y Ribeiro-Neto (1999)

Baeza-Yates, R. y Ribeiro-Neto, B. *Modern Information Retrieval*. Addison-Wesley, Wokingham, UK; 1999.

Barthès (1996)

Barthès J.P. "ISMICK and Knowledge Management" en *Knowledge Management: Organization, Competence and Methodology, Proceedings of ISMICK'96*, J. F. Schreinemakers (Ed.). Rotterdam, the Netherlands, Wurzburg:Ergon Verlag. 1996. p.9-13.

Bartsch (2004)

Bartsch, S. "Knowledge Representation Language and Conceptualization". *Knowledge Representation Handouts* [en línea]. 2004 [consultado en 2008]. p.1-3. Disponible en http://schwicky.net/projects/2004/language_and_conceptualization/20040208-Knowledge_Representation-Handouts.pdf.

Bateson (2002)

Bateson, G. *Mind and Nature: a Necessary Unity*; Hampton Press; New Ed edition; 2002.

Bechhofer y col. (2001a)

Bechhofer, S.; Horrocks, I.; Goble, C. y Stevens, R. "OILED: a Reason-able Ontology Editor for the Semantic Web" en *Proceedings of KI2001, Joint German/Austrian Conference on Artificial Intelligence*. 2001. Vienna, Springer-Verlag LNAI. Vol. 2174, p.396-408.

Bechhofer y col. (2001b)

Bechhofer, S.; Carr, L.; Goble, C. y Hall, W. "Conceptual Open Hypermedia = The Semantic Web?" [en línea]. 2001. [consultado en 2008]. Disponible en <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.25.9002>.

Bellinger y col. (2004)

Bellinger, G.; Castro, D. y Mills, A. "Data, Information, Knowledge and Wisdom" [en línea]. 2004. [consultado en 2008]. Disponible en <http://www.systems-thinking.org/dikw/dikw.htm>.

Bernaras y col. (1996)

Bernaras, A.; Laresgoiti, I. y Corera, J. "Building and Reusing Ontologies for Electrical Networks Applications" en *Proceedings of European Conference on Artificial Intelligence (ECAI 96)*. 1996. Budapest, Hungary, p.298-302.

Bodner y Song (1996)

Bodner, R.C. y Song, F. "Knowledge-based Approaches to Query Expansion in Information Retrieval" en *Proceedings of Canadian Conference on AI*. 1996. p.146-158.

Bolloju y col. (2002)

Bolloju, N.; Khalifa, M. y Turban, E. "Integrating Knowledge Management into Enterprise Environments for the Next Generation Decision Support". *Decision Support Systems*. 2002. Vol. 33, núm. 2, p.163-176.

Bonifacio y col. (2002)

Bonifacio, M.; Bouquet, P. y Cuel, R. "The Role Of Classification(s) In Distributed Knowledge Management" en *Proceedings of 6th International Conference on Knowledge-Based Intelligent Information Engineering Systems & Allied Technologies Special Session on Classification*. 2002.

Borst (1997)

Borst, W.N. *Construction of Engineering Ontologies*. PhD Thesis. University of Twente, Enschede, NL – Centre for Telematica and Information Technology. 1997.

Bray y col. (2000)

Bray, T.; Paoli, J.; Sperberg-McQueen, CM. y Maler, E. "Extensible Markup Language (XML) 1.0" [en línea] W3C Recommendation Segunda edición. 2000. [consultado en 2008]. Disponible en <<http://www.w3.org/TR/REC-xml>>.

Bresman y col. (1999)

Bresman, H.; Birkinshaw, J. y Nobel, R. "Knowledge Transfer in International Acquisitions". *Journal of International Business Studies*. 1999. Vol. 30, núm. 3, p.439-462.

Brewster y col. (2004)

Brewster, C.; O'Hara, K.; Fuller, S.; Wilks, Y.; Franconi, E.; Musen, M.; Ellman, J. y Buckingham Shum, S. "Knowledge Representation with Ontologies: The Present and Future". *IEEE Intelligent Systems*. 2004. Vol. 19, núm. 1, p.72-81.

Brusa y col. (2008)

Brusa, G.; Caliusco, M.L. y Chiotti, O. "Towards Ontological Engineering: A Process for Building a Domain Ontology from Scratch in Public Administration". *Expert Systems, The Journal of Knowledge Engineering*. 2008. Vol. 25, núm. 5, p. 484-503.

Ceusters y col. (2001)

Ceusters, W.; Martens, P.; Dhaen, C. y Terzic, B. "LinkFactory: an Advanced Formal Ontology Management System" en *Proceedings of Interactive Tools for Knowledge Capture, KCAP-2001*. 2001.

Chaudhri y col. (1998)

Chaudhri, H.; Farquhar, A.; Fikes, R.; Karp, P.D. y Rice, J.P. "Open Knowledge Base Connectivity 2.0.3" Technical Report [en línea]. 1998. [consultado en 2008]. Disponible en <<http://www.ai.sri.com>>.

Choo (1996)

Choo, C. "An Integrated Information Model of the Organizations: The Knowing Organization". [en línea]. 1996. [consultado en 2008]. Disponible en <<http://www.fis.utoronto.ca/people/faculty/choo/FIS/KO/KO.html>>

Ciravegna (2001)

Ciravegna, F. "Adaptive Information Extraction from Text by Rule Induction and Generalisation" en *Proceedings of the 17th International Joint Conference on Artificial Intelligence (IJCAI 2001)*. 2001. Seattle.

Cohen y Levinthal (1990)

Cohen, W.M. y Levinthal D.A. "Absorptive Capacity: A New Perspective on Learning and Innovation". *Administrative Science Quarterly*. 1990. p.128-152.

Cook y Brown (1999)

Cook, S.D.N. y Brown J.S. "Bridging Epistemologies: The Generative Dance between Organizational Knowledge and Organizational Knowing". *Organization Science*. 1999. Vol. 10, p.381-400.

Corcho y col. (2003)

Corcho, O.; Fernández-López, M. y Gómez-Pérez, A. "Methodologies, tools and languages for building ontologies. Where is their meeting point?". *Data & Knowledge Engineering*. 2003. Vol. 46, núm 1, p.41-64.

Cumming y Teng (2003)

Cumming, J. y Teng, B. "Transferring R&D Knowledge: The Key Factors Affecting Knowledge Transfer Success". *Journal of Engineering and Technology Management*. 2003. Núm. 20, p.39-68.

Davenport (1996)

Davenport T.H. "Some Principles of Knowledge Management". *Knowledge Management Server Graduate School of Business University of Texas at Austin* [en línea]. 1996. [consultado en 2008]. Disponible en <<http://www.bus.utexas.edu/kman>>.

Davenport (1997)

Davenport, T.H. *Information Ecology: Mastering the Information and Knowledge Environment*. Oxford University Press, New York. 1997.

Davenport y Grover (2001)

Davenport, T.H. y Grover, V. "Special Nro.: Knowledge Management". *Journal of Management Information System*. 2001. Vol. 18, p.3-4.

Davenport y Prusak (1998)

Davenport, T.H. y Prusak, L. *Working Knowledge: How Organizations Manage What they Know*. Harvard Business School Press. Boston, MA. 1998.

Decker y col. (1999)

Decker, S.; Erdmann, M.; Fensel, D. y Studer, R; "Ontobroker: Ontology based Access to Distributed and Semi-structured Information" en *Semantic Issues in*

Multimedia Systems (DS8). 1999. Kluwer Academic Publisher, Boston, p.351-369.

Despres y Chauvel (2002)

Despres, C. y Chauvel, D. "Knowledge Management Models: A State of the Art". *KnowledgeBoard*. 2002.

Desouza (2003)

Desouza, K. "Knowledge Management Barriers: Why the Technology Imperative Seldom Works". *Business Horizons*. 2003. Vol. 46, núm. 1, p.25-29.

Dieng y col. (1999)

Dieng, R.; Corby, O.; Giboin, A. y Ribière, M. "Methods and Tools for Corporate Knowledge Management". *International Journal of Human-Computer Studies*. 1999. (Eds.) S. Decker and F. Maurer. Special Nro. on Knowledge Management. Vol. 51, p.567-598.

Domingue (1998)

Domingue, J. "Tadzebao y Webonto: Discussing, Browsing and Editing Ontologies on the Web" en *Proceedings of the 11th Knowledge Acquisition Workshop (KAW98)*, Banff. 1998.

Drucker (1988)

Drucker, P.F. *The Coming of the New Organization*. Harvard Business Review. 1988. Reimpresión de 1998.

Du y col. (2009)

Du, T.C., Li, F. y King, I. "Managing Knowledge On The Web – Extracting Ontology From Html Web". *Decision Support Systems*. 2009.

Duffy (2001)

Duffy, J. "The Tools and Technologies Needed for Knowledge Management". *Information Management Journal*. 2001. Vol. 35, núm. 1, p.64-67.

Dzbor y col. (2000)

Dzbor, M.; Paralic, J. y Paralic, M. "Knowledge Management in a Distributed Organisation" en *Proceedings of the BASYS'2000 - 4th IEEE/IFIP International Conference on Information Technology for Balanced Automation Systems in Manufacturing*. 2000. Kluwer Academic Publishers, London, p.339-348. ISBN 0-7923-7958-6.

Earl (1994)

Earl, M.J. "Knowledge as Strategy: Reflections on Skandia International and Shorko Films" en *Strategic Information Systems: A European Perspective*. 1994. (Eds.) C. Ciborra and T. Jelassi, John Wiley & Sons Ltd. Chichester, p.53-69.

Ermine (2000)

Ermine, J.L. "Challenges and Approaches for Knowledge Management in Companies" en *Workshop in Knowledge Management: Theory and*

Applications, Fourth European Conference on Principles and Practice of Knowledge Discovery in Databases. 2000. Lyon, France.

Euzenat (1996)

Euzenat, J. "Corporate Memory through Cooperative Creation of Knowledge Bases and Hyper-documents" en *Proceeding of the KAW'96*. 1996. (Eds.) B. Gaines, M. Musen. Banff, Canada, p.36-1 36-18.

Farquhar y col. (1996)

Farquhar, A.; Fikes, R. y Rice, J. "The Ontolingua Server: A Tool for Collaborative Ontology Construction" en *Proceedings of the 10th Knowledge Acquisition for Knowledge-Based Systems Workshop (KAW96)*. 1996. Banff, Canada, p.44.1-44.19.

Fernandes (2000)

Fernandes, A. "Combining Inductive and Deductive Inference in Knowledge Management Tasks" en *Proceedings of the 11th International Workshop on Database and Expert Systems Applications, First International Workshop on Theory and Applications of Knowledge Management*. 2000. p.1109-1114.

Fernández-López y col. (1999)

Fernández-López, M.; Gómez-Pérez, A.; Pazos, A. y Pazos, J. "Building a Chemical Ontology Using Methontology and the Ontology Design Environment". *IEEE Intelligent Systems and their Applications*. 1999. Vol. 4, núm. 1, p.37-46.

Fleming (1996)

Fleming, N.; *Coping with a revolution: will the internet change learning?*; Lincoln University: Canterbury, New Zealand; 1996.

Fukuda (1995)

Fukuda, Y. "Variations of Knowledge in Information Society" en *Proceedings of ISMICK 95*. 1995. p.3-8.

Genesereth y Fikes (1992)

Genesereth, M. y Fikes, R. "Knowledge Interchange Format". *Technical Report Logic-92-1*. 1992. Computer Science Department, Stanford University.

Geroimenko y Chen (2005)

Geroimenko, V. y Chen, C. (Eds.); "SVG and X3D in the Context of the XML Family and the Semantic Web"; en *Visualizing Information Using SVG and X3D. XML-based Technologies for the XML-based Web*; p. 298, 2005; ISBN: 978-1-85233-790-2

Gongla y Rizzuto (2001)

Gongla, P. y Rizzuto, C.R. "Evolving Communities of Practice: IBM Global Services experience". *IBM Systems Journal*. 2001. Vol. 40, p.843-862.

Grant (1996)

Grant, R.M.; *Toward a Knowledge-based Theory of the Firm*; Strategic Management Journal, Vol. 17, Página(s) 109-122, 1996.

Grant y Grant (2008)

Grant, K.A. y Grant, C.T. "Developing a Model of Next Generation Knowledge Management". *Issues in Informing Science and Information Technology*. 2008. Volumen 5.

Gruber (1992)

Gruber, T.R. "ONTOLINGUA: A Mechanism to Support Portable Ontologies". *Technical Support*. 1992. Knowledge Systems Laboratory, Stanford University.

Gruber (1993a)

Gruber, T.R. "A Translation Approach to Portable Ontologies". *Knowledge Acquisition*. 1993. Vol. 5, núm. 2, p.199-220.

Gruber (1993b)

Gruber, T.R. "Toward Principles for the Design of Ontologies Used for Knowledge Sharing" en *Formal Ontology in Conceptual Analysis and Knowledge Representation*. 1993. (Ed.) Nicola Guarino y Roberto Poli, Kluwer Academic Publishers.

Grunstein y Barthès (1996)

Grunstein, M. y Barthès J.P. "An Industrial View of the Process of Capitalizing Knowledge" en *Knowledge Management: Organization, Competence and Methodology, Proceedings of the ISMICK'96*. 1996. (Ed.) J. F. Schreinemakers. Rotterdam, the Netherlands, Wurzburg:Ergon Verlag, Advances in Knowledge Management. Vol. 1, p.258-264.

Guarino y col. (1995)

Guarino, N.; Carrara, M. y Giaretta, P. "Ontologies and Knowledge Bases: Towards a Terminological Clarification" en *Towards Very Large Knowledge Bases, Knowledge Building and Knowledge Sharing*. 1995. (Ed.) N. Mars, IOS Press, Amsterdam. p.25-32.

Guarino y col. (1999)

Guarino, N.; Masolo, C. y Vetere, G. "Ontoseek: Content-based Access to the Web". *IEEE Intelligent Systems*. 1999. Vol. 14, núm. 3, p.70-80.

Hackbarth y Grover (1999)

Hackbarth G. y Grover, V. "The Knowledge Repository: Organizational Memory Information Systems". *Information Systems Management*. 1999. Vol. 16, núm. 3, p.21-30.

Hall (1998)

Hall, R. "The Management of Intellectual Assets: A New Corporate Perspective" en *The Knowledge Economy*. 1998. (Ed.) D. Neef. Woburn, MA: Butterworth-Heinemann. p.119-132.

Hall y col. (2001)

Hall, D.; Paradice, D. y Courtney, J. "Creating Feedback Loops to Support Organizational Learning and Knowledge Management in Inquiring

Organizations” en *Proceedings of the 34th Hawaii International Conference on Systems Sciences*. 2001.

Handschub y col. (2001)

Handschub, S.; Staab, S. y Maedche, A. “CREAM – Creating Relational Metadata with a Component-Based, Ontology-Driven Framework” en *Proceedings of the First International Conference on Knowledge Capture*. 2001. Victoria, Canada.

Harris (2008)

Harris, A. “Leading Innovation and Change: knowledge creation by schools for schools”. *European Journal of Education*. 2008. Vol. 43, núm 2, p.219-228.

van Heijst y col. (1996)

Van Heijst, G.; Van der Spek, R. y Kruizinga, E. “Organizing Corporate Memories” en *Proceedings of the 10th Banff Knowledge Acquisition for Knowledge-Based Systems Workshop (KAW'96)*. 1996. (Ed.) N. Mars. Banff, Canada. p.42-1 42-17.

Heflin y Hendler (2000)

Heflin, J. y Hendler, J. “Searching the Web with SHOE” en *Artificial Intelligence for Web Search*. 2000. Papers from the AAAI workshop, AAAI Press: Menlo Park.

Heflin y Hendler (2001)

Heflin, J. y Hendler, J. “A Portrait of the Semantic Web in Action”. *IEEE Intelligent Systems*. 2001. Vol. 16, núm. 2.

Hlupic y col. (2002)

Hlupic, V.; Pouloudi, A. y Rzevski, G. “Towards an Integrated Approach to Knowledge Management: ‘Hard’, ‘Soft’, and ‘Abstract’ Issues”. *Knowledge and Process Management*. 2002. Vol. 9, núm 2, p.90-102.

Holzner y Marx (1979)

Holzner, B. y Marx, J. *Knowledge Application: The Knowledge System in Society*. 1979. Allyn-Bacon, Boston.

Horrocks y col. (1999)

Horrocks, I.; Sattler, U. y Tobies, S. “Practical Reasoning for Expressive Description Logics” en *Proceedings of 6th International Conference on Logic for Programming and Automated Reasoning (LPAR99)*. 1999. LNAI Springer-Verlag. p.161-180.

Horrocks y col. (2000)

Horrocks, I.; Fensel, D.; Harmelen, F.; Decker, S.; Erdmann, M. y Klein, M. “OIL in a Nutshell” en *Proceedings of ECAI00 Workshop on Application of Ontologies and PSMs*. 2000. Berlín.

Horrocks y van Harmelen (2001)

Horrocks, I. y van Harmelen, F. "Reference Description of the DAML+OIL Ontology Markup Language". [en línea] Technical Report. [consultado en 2008]. Disponible en <<http://www.daml.org/2001/03/reference.html>>.

Hussain y Wahba (2002)

Hussain, A. y Wahba, K. "The Readiness of Information and Decision Support Center in Egypt to Adopt Knowledge Management" en *Issues and Trends of IT Management in Contemporary Organizations*. 2002. Idea Group Inc (IGI), Seattle, Washington, USA.

Joshi (2001)

Joshi, K. "A Framework to Study Knowledge Management Behaviors During Decision Making" en *Proceedings of the 34th Hawaii International Conference on System Sciences (HICSS-34)*. 2001. Vol. 4, p.4024.

Junarkar (1997)

Junarkar, B. "Leveraging Collective Intellect by Building Organizational Capabilities". *Expert Systems with Applications*. 1997. Vol. 13, núm. 1, p.29-40.

Karp y col. (1999)

Karp, R.; Chaudhri, V. y Thomere, J. "XOL: An XML-Based Ontology Exchange Language" [en línea] 1999. Technical Report. [consultado en 2008]. Disponible en <<http://www.ai.sri.com>>5

Khan y McLeod (2000)

Khan, L. y McLeod, D. "Effective Retrieval of Audio Information from Annotated Text using Ontologies" en *Proceedings of Workshop on multimedia data mining (MDM/KDD 2000)*. 2000. (Eds.) S.J. Simoff y O.R. Zaiane, Boston, MA.

Kifer y col. (1995)

Kifer, M.; Lausen, G. y Wu, J. "Logical Foundation of Object-Oriented and Frame-Based Languages". *Journal of the ACM*. 1995. Vol. 42, núm. 4, p.741-843.

Kimble y col. (2001)

Kimble, C.; Hildreth, P. y Wright, P. "Communities of Practice: Going Virtual". *Chapter 13 in Knowledge Management and Business Model Innovation*. Idea Group Publishing, Hershey (USA)/London (UK). 2001.p.220 – 234. ISBN 1 878289 98 5.

King (1999)

King, W.R. "Integrating Knowledge Management into IS strategy". *Information Systems Management*. 1999. Vol. 16, núm. 4, p.70-72.

King y col. (2002)

King, W.R.; Marks, P.V. y McCoy, S. "The Most Important Issues in Knowledge Management". *Communications of the ACM*. 2002. Vol. 45, núm. 9, p.93-97.

Koenig y Srikantaiah (2000)

Koenig, M.E.D. y Srikantaiah, T.K. "The Evolution of Knowledge Management" en *Knowledge Management for the Information Professional*. 2000. (Eds.) Srikantaiah T.K. y Koenig M.E.D., Medford, New Jersey: Information Today. p. 23-36.

Kogut y Holmes (2001)

Kogut, P. y Holmes, W.; *AeroDAML: Applying Information Extraction to Generate DAML Annotations from Web Pages*; First International Conference on Knowledge Capture (K-CAP 2001); Workshop on Knowledge Markup and Semantic Annotation; Victoria, B.C.; 2001.

Kogut y Zander (1992)

Kogut, B. y Zander, U. "Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology". *Organization Science*. 1992. Vol. 3, p.383-397.

Kogut y Zander (1996)

Kogut, B. y Zander, U. "What Firms do? Coordination, Identity, and Learning". *Organization Science*. 1996. Vol. 7, p.502-518.

von Krogh (1998)

Von Krogh, G. "Care in Knowledge Creation: Special Nro. of Knowledge and the Firm". *California Management Review*. 1998. Vol. 40, núm. 3, p.133. ISSN: 0008-1256.

von Krogh y col. (2000)

Von Krogh, G; Ichijo, K. y Nonaka, I. *Enabling Knowledge Creation*. Oxford University Press, Oxford, 2000.

Kuhn (1996)

Kuhn, T.S. *The Structure of Scientific Revolutions*. The University of Chicago Press. 3ra edición. Chicago, 1996.

Kuhn y Abecker (1997)

Kuhn, O. y Abecker, A. "Corporate Memories for Knowledge Management in Industrial Practice: Prospects and Challenges". *Journal of Universal Computer Science*. 1997. Vol. 3, núm. 8, p. 929-954.

Kwan y Balasubramanian (2003)

Kwan, M. y Balasubramanian, P. "KnowledgeScope: Managing Knowledge in Context". *Decision Support Systems*. 2003. Vol. 35, núm. 4, p.467-486.

Lassila y Swick (1999)

Lassila, O. y Swick, R. "Resource Description Framework (RDF) model and syntax specification" [en línea]. 1999. W3C Recommendation. [consultado en 2008]. Disponible en <<http://www.w3.org/TR/REC-rdf-syntax>>.

Lee y col. (2001)

Lee, J.; Kim, Y. y Yu, S. "Stage Model for Knowledge Management" en *Proceeding of the 34th Hawaii International Conference on System Sciences*. 2001.

Lei y col. (2006)

Lei, Z.; Bin, G. y Yanning, X. "A Knowledge Management Model For Product Design Process". *International Conference on Computer-Aided Industrial Design and Conceptual Design*. 2006. p.1 – 6.

Leonard-Barton (1995)

Leonard-Barton, D. *Wellsprings of Knowledge*. Boston: Harvard Business School Press, 1995.

Lesser y Storck (2001)

Lesser, E.L. y Storck, J. "Communities of Practice and Organizational Performance". *IBM Systems Journal*. 2001. Vol. 40, p.831-841.

Levine (2001)

Levine, L. "Integrating Knowledge and Processes in a Learning Organization". *Information System Management*. 2001. p.21-32.

Li y col. (2008)

Li, Z.; Xi, Y. y Ge, J. "Study On Knowledge Management Model Based On Business Process". *Conference on Fuzzy Systems and Knowledge Discovery*. 2008. Volumen 5, p. 562 – 569.

Liao y col. (1999)

Liao, M.; Hinkelmann, K.; Abecker, A. y Sintek, M. "A Competence Knowledge Base System for the Organizational Memory" en Frank Puppe (ed.) XPS-99 / 5. 1999. Deutsche Tagung Wissensbasierte Systeme, Würzburg, Springer Verlag, LNAI 1570.

Lin y col. (2007)

Lin, C.; Yen, D. y Tarn, D. "An Industry-Level Knowledge Management Model—A Study Of Information-Related Industry In Taiwan". *Information & Management*. 2007. Volumen 44, Número 1, p. 22-39.

de Long y Seeman (2000)

de Long, D. y Seeman, P. "Confronting Conceptual Confusion and Conflict in Knowledge Management". *Organizational Dynamics*. 2000. Vol. 29, núm. 1, p.33-44.

Luke y Heflin (2000)

Luke, S. y Heflin, J. "SHOE 1.01 Proposed Specification". [en línea] 2000. SHOE Project Technical Report, University of Maryland. [consultado en 2008]. Disponible en <<http://www.cs.umd.edu/projects/plus/SHOE/spec1.01.htm>>.

Maier y Remus (2001)

Maier, R. y Remus, U. "Towards a Framework for Knowledge Management Strategies: Process Orientation as Strategic Starting Point" en *Proceedings of the 34th Hawaii International Conference on System Sciences*. 2001.

Malone (2002)

Malone, D. "Knowledge Management: A Model for Organizational Learning". *International Journal of Accounting Information Systems*. 2002. Vol. 3, núm. 02, p.111-123.

Manning y col. (2008)

Manning, C.; Raghavan, P. y Schütze, H. *Introduction to Information Retrieval*. Cambridge University Press. 2008. ISBN: 0521865719.

Mansingh y col. (2009)

Mansingh, G.; Osei-Bryson, K y Reichgelt, H. "Issues In Knowledge Access, Retrieval And Sharing – Case Studies In A Caribbean Health Sector". *Expert Systems with Applications*. 2009. Volumen 36, Núm. 2, Parte 2, p. 2853-2863.

Marchand (1998)

Marchand, D.A. *Competing with Intellectual Capital. Knowing in Firms: Understanding, Managing and Measuring Knowledge*. Ed. Georg vonKrogh, Johan Roos and Dirk Kleine. 1998. London: SAGE. p.253-268.

MacGregor (1991)

MacGregor, R. *Inside the LOOM Classifier*. 1991. SIGART Bulletin 2 núm. 3, p.70-76.

Massa y Testa (2009)

Massa, S. y Testa, S. "A Knowledge Management Approach To Organizational Competitive Advantage: Evidence From The Food Sector". *European Management Journal*. 2009. Volumen 27, Núm. 2, p.129-141.

McDermott (1999)

McDermott, R. "Why Information Technology Inspired but cannot Deliver Knowledge Management". *California Management Review*. 1999. Vol. 41, núm. 4, p.103-117.

McGuinness (1998)

McGuinness, D.L. "Ontological Issues for Knowledge-enhanced Search" en *Proceedings of FOIS98*. 1998. N. Guarino (Ed.), Formal Ontology in Information Systems, Trento, Italy, IOS Press.

McGuinness y col. (2000)

McGuinness, D.L.; Fikes, R.; Rice, J. y Wilder, S. "The Chimaera Ontology Environment" en *Proceedings of the 17th National Conference on Artificial Intelligence (AAAI00)*. 2000. Austin.

Menville y Foot (1996)

Menville, B. y Foot, N. "Harvest your Workers' Knowledge". *Datamation*. 1996. Vol. 42, p.78-83.

Missikoff y col. (2002)

Missikoff, M.; Velardi, P. y Navigli, R. "The Usable Ontology: an Environment for Building and Assessing a Domain Ontology" en *Proceedings of the International Semantic Web Conference (ISWC2002)*. 2002. Sardinia, Italy.

Moody (2005)

Moody, D. "Theoretical and practical issues in evaluating the quality of conceptual models: current state and future directions". *Data & Knowledge Engineering*. 2005. Núm. 55 p. 243-276.

Motta (1999)

Motta, E. *Reusable Components for Knowledge Modelling*, IOS Press, Amsterdam, 1999.

Motta y col. (2002)

Motta, E.; Vargas-Vera, M.; Domingue, J.; Lanzoni, M. y Ciravegna, F. "MnM: Ontology Driven Semi-Automatic and Automatic Support for Semantic Markup". *Semantic Authoring, Annotation & Knowledge Markup Workshop*. 2002. Lyon, France; ECAI.

Murray (2001)

Murray, M.; *Beyond the myths and magic of mentoring: how to facilitate an effective mentoring process*; 2da. Edición ilustrada; Publicado por John Wiley and Sons, 2001.

Nagendra Prasad y Plaza (1996)

Nagendra Prasad M.V.N. y Plaza E. "Corporate Memories as Distributed Case Libraries" en *Proceedings of the 10th Banff Knowledge Acquisition for Knowledge-Based Systems Workshop (KAW'96)*. 1996. (Eds.) Dans B. Gaines, M. Musen. Banff, Alberta, p.40-1 a 40-19.

Neches y col. (1991)

Neches, R.; Fikes, R.E.; Finin, T.; Gruber, T.R.; Senador, T. y Swartout, W.R. "Enabling Technology for Knowledge Sharing". *AI Magazine*. Vol. 12, núm. 3, p.36-56.

Neef (1998)

Neef, D. *The Knowledge Economy* (Ed.), Woburn, MA: Butterworth-Heinemann, 1998.

Nemati y col. (2002)

Nemati, N.; Steiger, D.; Iyer, L. y Herschel, R. "Knowledge Warehouse: an Architectural Integration of Knowledge Management Decision Support, Artificial Intelligence and Data Warehousing". *Decision Support Systems*. 2002. Vol. 33, núm. 2, p.143-161.

Newell (1982)

Newell A. "The Knowledge Level". *Artificial Intelligence*. 1982. Vol. 18, núm. 1, p.87-127.

Nonaka (1994)

Nonaka, I. "Dynamic Theory of Organizational Knowledge Creation". *Organizational Science*. 1994. Vol. 5, núm. 1, p.14-37.

Nonaka y Takeuchi (1995)

Nonaka I. y Takeuchi H. *Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*, Oxford University Press, 1995.

Noy y col. (2000)

Noy, N.F.; Ferguson, R.W. y Musen M.A. "The Knowledge Model of Protégé-2000: Combining Interoperability and Flexibility" en *Proceedings of the 12th International Conference in Knowledge Engineering and Knowledge Management (EKAW00)*. 2000. Lecture Notes in Artificial Intelligence, Springer, Berlin. Vol. 1937, p.17-32.

Noy y col. (2001)

Noy, N.F.; Sintek, M.; Decker, S.; Crubezy, M.; Ferguson, R.W. y Musen, M.A. "Creating Semantic Web Contents with Protégé-2000". *IEEE Intelligent Systems*. 2001. Vol. 16, núm. 2; p.60-71.

Nunes y col. (2009)

Nunes, V.T.; Santoro, F.M. y Borges, M.R.S.; "A Context-Based Model For Knowledge Management Embodied In Work Processes". *Information Sciences*. 2009. Volumen 179, Núm 15, p. 2538-2554.

O'Leary (1998)

O'Leary, D. "Enterprise Knowledge Management". *IEEE Computer*. 1998. Vol. 31, núm. 3, p.54-61.

O'Leary (2002)

O'Leary, D. *Chapter 13: Knowledge Management in Accounting and Professional Services*. En: V. Arnold y S. Sutton (eds.), *Researching Accounting as an Information Systems Discipline*, Sarasota, FL, American Accounting Association, 2002.

OWL (2004)

OWL Web Ontology Language [en línea]. 2004. [consultado en 2008]. Disponible en <<http://www.w3.org/TR/owl-features/>>.

Patel y Hlupic (2002)

Patel, N. y Hlupic, V. "A Methodology for the Selection of Knowledge Management (KM) Tools" en *Proceedings of the 24th International Conference on Information Technology Interfaces (ITI 2002)*. 2002. Cavtat, Croatia. p.369-374.

Petrash (1996)

Petrash, G. "Dow's Journey to a Knowledge Value Management Culture". *European Management Journal*. 1996. Vol. 14, núm. 4, p.365-373.

Pitt (2001)

Pitt, J.; *What Engineers Know*. [en línea]. 2001. [consultado en 2008]. Disponible en <<http://scholar.lib.vt.edu/ejournals/SPT/v5n3/pdf/pitt.pdf>>.

Polanyi (1997)

Polanyi, M. *Personal Knowledge: Towards a Post-Critical Philosophy*. Routledge, London, 1997, Primera Edición. 1958

Ponelis y Fair-Wessels (1998)

Ponelis S. y Fair-Wessels, F. “Knowledge Management: A Literature Overview”. *South Africa Journal of Library Information Science*. 1998. Vol. 66, núm. 1, p.1-10.

Probst y col. (2000)

Probst, G.; Steffen, R. y Kai, R. *Managing Knowledge: Building Blocks for Success*. Chichester: John Wiley & Sons. 2000.

Prusak (1998)

Prusak, L. “Introduction to Series: Why Knowledge, Why Now?” en *The Knowledge Economy*. 1998. (Ed.) D. Neef: Butterworth-Heinemann. p.9-10.

Qu y Ren (2008)

Qu, Z. y Ren, Z. “The Frame Of Enterprise Knowledge Management Model Based On Semantic Web”. *9th International Conference on Signal Processing*. 2008. p. 2944 – 2947.

Quinn y col. (1996)

Quinn, J.B.; Anderson, P. y Finkelstein, S. “Leveraging Intellect”. *Academy of Management Executive*. 1996. Vol. 10, núm. 3, p.7-27.

Rabarijaona y col. (1999)

Rabarijaona, A.; Dieng, R. y Corby, O. “Building a XML-based Corporate Memory” en *Proceedings of the IJCAI'99 Workshop on Knowledge Management and Organizational Memories*. 1999. (Eds.) John Debenham, Stefan Decker, Rose Dieng, Ann Macintosh, Nada Matta and Ulrich Reimer eds., Stockholm, Sweden.

Raghu y Vinze (2007)

Raghu, T.S. y Vinze, A. “A Business Process Context For Knowledge Management”. *Decision Support Systems*. 2007. Volumen 43, Núm. 3, p. 1062-1079.

Recker (2005)

Recker, J. “Conceptual Model Evaluation. Towards more Paradigmatic Rigor” en *Proceedings of the Workshop on Evaluating Modeling Methods for Systems Analysis and Design (EMMSAD'05)*, 2005, Porto, Portugal, ISBN 9727520774.

Rector y col. (1997)

Rector, A.L.; Bechhofer, S.K.; Goble, C.A.; Horrocks, I.; Nowlan, W.A. y Solomon, W.D. "The GRAIL Concept Modelling Language for Medical Terminology" *Artificial Intelligence in Medicine*. 1997. Vol. 9.

Rodriguez y Egenhofer (2003)

Rodriguez, M.A. y Egenhofer, M.J. "Determining Semantic Similarity among Entity Classes from Different Ontologies" *IEEE Transactions on Knowledge and Data Engineering*. 2003. Vol. 15, núm. 2.

Rogers y col. (2001)

Rogers, J.E.; Roberts, A.; Solomon, W.D.; van der Haring, E.; Wroe, C.J.; Zanstra, P.E. y Rector, A.L. "GALEN Ten Years On: Tasks and Supporting Tools" en *Proceedings of MENDINFO2001*. 2001. (Eds.) Patel, V. y col.; IOS Press. p. 256-260.

Sangüesa Sol y Pujol Serra (2001)

Sangüesa Sol, R. y Pujol Serra J.M. "NetExpert: A Multiagent System for Expertise Location" en *Proceedings IJCAI'01 Workshop on Knowledge Management and Organizational Memories (KM/OM)*. 2001. p.85-93.

Schreiber y col. (1995)

Schreiber, A.; Wielinga, B. y Jansweijer, W. "The KACTUS view on the "O" word". *Technical Report*, ESPRIT Project 8145 KACTUS. 1995. University of Amsterdam, The Netherlands.

Seely Brown y Duguid (1998)

Seely Brown, J. y Duguid, P. "Organizing Knowledge". *California Management Review*. 1998. Vol. 40, núm. 3, p.90-111.

Senge (2006)

Senge, P.M. *The Fifth Discipline: The Art and Practice of the Learning Organization*, New York: Currency/DoubleDay, 2006, Primera Edición en 1994.

Seufert y col. (1999)

Seufert, A.; von Krough, G. y Black, A. "Towards Knowledge Networking". *Journal of Knowledge Management*. 1999, núm 3.

Simon (1996)

Simon, G. "Knowledge Acquisition and Modeling for Corporate Memory: Lessons Learnt from Experience" en *Proceedings of KAW'96*. 1996. (Eds.) B. Gaines, M. Musen. Banff, Canada, p.41-1 41-18.

Simon y Grandbastien (1995)

Simon, G. y Grandbastien, M. "Corporate knowledge: a Case Study in the Detection of Metallurgical Flaws" en *Proceedings of the Third International Symposium on the Management of Industrial and Corporate Knowledge (ISMICK'95)*. 1995. (Ed.) J. P. Barthès. Compiègne, France. p.43-52.

Simonin (1999)

Simonin, B. "Transfer of Marketing Know-How in International Strategic Alliances: An Empirical Investigation of the Role and Antecedents of Knowledge Ambiguity". *Journal of International Business Studies*. 1999. Vol. 30, núm. 3, p.463-490.

Sintek y col. (2000)

Sintek, M.; Tschaitshian, B.; Abecker, A.; Bernardi, A. y Müller, H.J. "Using Ontologies for Advanced Information Access" en *Proceedings of the Third International Conference and Exhibition on the Practical Application of Knowledge Management*. 2000. (Ed.) J. Domingue. Manchester, UK.

Skyrme (2003)

Skyrme, D.; "Knowledge Management: Making Sense of an Oxymoron". *Management Insight*. [en línea]. 2003. [consultado en 2008]. Disponible en <http://www.skyrme.com/insights/22km.htm>.

Slade y Bokma (2001)

Slade, A. y Bokma, A. "Conceptual Approaches for Personal and Corporate Information and Knowledge Management". *IEEE Proceedings of the 34th Hawaii International Conference on System Sciences*. 2001.

Smith (1776)

Smith, A. *La Riqueza de las Naciones*; Ed. Longseller, 2004 Argentina. Primera Edición en 1776

Smith y Farquhar (2000)

Smith, R.G. y Farquhar, A. "The Road Ahead for Knowledge Management: An AI Perspective". *AI Magazine*. 2000. p.17-40.

Snyder y Wilson (2002)

Snyder C.A. y Wilson L.T. "Implementing Knowledge Management: Issues for Managers". *Knowledge Mapping and Management*. 2002. (Ed.) D. White. Hershey, Pennsylvania: Idea Group. p.154-165.

Sowa (1984)

Sowa, J.F. *Conceptual Structures Information Processing in Mind and Machine*; Reading, Massachusetts: Addison-Wesley, 1984.

van der Spek y Spijkervet (1997)

Spek, R.V. y Spijkervet, A. "Knowledge Management: Dealing Intelligently with Knowledge". *Knowledge Management and Its Integrative Elements*. 1997. (Ed.) Wilcox Liebowitz J., L.C. CRC Press.

Spiegler (2003)

Spiegler, I. "Technology and Knowledge: Bridging a "Generating" Gap". *Information & Management*. 2003. Vol. 40, núm. 6, p.533-539.

Staab (2004)

Staab, S. "Knowledge Representation with Ontologies: The Present and and Future". *IEEE Intelligent Systems*. 2004. p.72-81.

Staab y col. (2000)

Staab, S.; Angele, J.; Decker, S.; Erdmann, M.; Hotho, A.; Maedche, A.; Schnurr, H.P.; Studer, R. y Sure, Y.; "AI for the web – Ontology-based community portals" en *Proceedings de AAAI/IAAI*. 2000. p.1034-1039.

Studer y col. (1998)

Studer, R.; Benjamins, V.R. y Fensel, D. "Knowledge Engineering: Principles and Methods". *Data and Knowledge Engineering*. 1998. núm. 25, p.161-197.

Struts (2000)

Struts [en línea]. 2000. [consultado en 2008]. Disponible en <<http://struts.apache.org/>>.

Sure y col. (2002)

Sure, Y.; Erdmann, M.; Angele, J.; Staab, S.; Studer, R. y Wenke, D. "OntoEdit: Collaborative Ontology Engineering for the Semantic Web" en *Proceedings of the First International Semantic Web Conference (ISWC02)*. 2002. Lecture Notes in Computer Science, Springer, Berlin. Vol. 2342, p.221-235.

Sveiby (1997a)

Sveiby, K.E. *The New Organizational Wealth: Managing and Measuring Knowledge-based Assets*. Berret-Koehler Publishers, Inc. San Francisco. 1997.

Sveiby (1997b)

Sveiby, K.; *The New Organizational Wealth*. San Francisco: Berret-Koehler; 1997.

Swan y Scarborough (2002)

Swan, J. y Scarborough, H. "The paradox of "knowledge management"". *Informatik Informatique*. 2002, Núm 1, p. 10-13.

Swartout y col. (1997)

Swartout, B.; Ramesh, P.; Knight, K. y Russ, T. "Toward Distributed Use of Large-Scale Ontologies" en *Proceedings of the AAAI Symposium on Ontological Engineering*. 1997. Stanford, California.

Szulanski (1996)

Szulanski, G. "Exploring Internal Stickiness: Impediments to the Transfer of Best Practice within the Firm". *Strategic Management Journal*. 1996. p.27-43.

Teece (2001)

Teece, D.J. "Strategies for Managing Knowledge Assets: the Role of Firm Structure and Industrial Context" en *Managing Industrial Knowledge: Creation, Transfer and Utilization*. 2001. (Eds.) I. Nonaka and D.J. Teece: Sage Publications. p.125-144.

Thomas y col. (2001)

Thomas, J.C.; Kellogg, W.A. y Erickson, T. "The Knowledge Management Puzzle: Human and Social Factors in Knowledge Management". *IBM Systems Journal*. 2001. Vol. 40, núm. 4, p.863-884.

Tiwana (2001)

Tiwana, A. "Integrating Knowledge on the Web". *IEEE Internet Computing*. 2001. p.32-39.

Tuomi (1999)

Tuomi, I. "Data is more than knowledge: implications of the reversed knowledge hierarchy for knowledge management and organizational memory". *Journal of Management Information Systems*. 1999. Vol. 16, núm. 3, p.103 – 117.

Tuomi (2002)

Tuomi, I. "The Future of Knowledge Management". *Lifelong Learning in Europe (LLinE)*. 2002. Vol. VII, Núm. 2, pp. 69-79.

Turban y Frenzel (1992)

Turban, E. y Frenzel L.E. Jr. *Expert Systems and Applied Artificial Intelligence*. New York: Macmillan Publishing. 1992.

Turban y col. (2002)

Turban, E.; McLean, E. y Wetherbe, J. *Information Technology for Management: Transforming Business in the Digital Economy*. 3ra. Edición. New York: John Wiley & Sons. 2002.

Uschold y Gruninger (1996)

Uschold M. y Gruninger M. "Ontologies: Principles, Methods and Applications". *Knowledge Engineering Review*. 1996. Vol. 11, núm. 2, p.93-136.

Uschold y Jasper (1999)

Uschold, M. y Jasper, R. "A Framework for Understanding and Classifying Ontology Applications" en *Proceedings of the IJCAI99 Workshop on Ontologies and Problem-Solving Methods*. 1999. Stockholm.

Vasconcelos y col. (2002)

Vasconcelos, J.; Gouveia, F. y Kimble, C. "An Organizational Memory Information System using Ontologies" en *Proceedings of the 3rd Conference of the Associação Portuguesa de Sistemas de Informação*. 2002. University of Coimbra, Portugal.

Voss y Schäfer (2003)

Voss, A. y Schäfer, A. "Discourse Knowledge Management in Communities of Practice" en *Proceedings of the 14th Internacional Workshop on Database and Expert Systems Applications (DEXA'03)*. 2003.

Weggeman (1996)

Weggeman, M. "Knowledge Management: The Modus Operandi for a Learning Organization" en *Proceedings of ISMICK'96*. 1996. (Ed.) J. F. Schreinemakers ed, Knowledge Management: Organization, Competence and Methodology. Rotterdam, the Netherlands, Wurzburg:Ergon Verlag, Advances in Knowledge Management. Vol. 1, p.175-187.

Wenger y Zinder (2000)

Wenger, E.C. y Zinder, W.M. "Communities of Practice: the Organizational Frontier". *Harvard Business Review*. 2000. p.139-145.

Wiig (1993a)

Wiig, K.M. *Knowledge Management Foundations: Thinking About Thinking – How People and Organizations Create, Represent and Use Knowledge*. Schema Press, Arlington, TX. 1993.

Wiig (1993b)

Wiig, K.M. *Knowledge Management Foundations*, Arlington, TC.: Schema Press. 1993.

Wijnhoven (1998)

Wijnhoven, F. "Designing Organizational Memories: Concept and Method". *Journal of Organizational Computing and Electronic Commerce*. 1998. Vol. 8, núm. 1, p.29-55.

Woolf (1990)

Woolf, H. eds.; *Webster's New World Dictionary of the American Language*, G. and C. Merriam, 1990.

WordNet (2006)

WordNet, *WordNet Documentation Index*. [en línea]. 2006. [consultado en 2008]. Disponible en <<http://wordnet.princeton.edu/doc>>.

Xie y col. (2006)

Xie, X.; Zhang, W. y Xu, L. "A Description Model To Support Knowledge Management". *International Multi-Symposiums on Computer and Computational Sciences*. 2006. Volumen 2, p. 433 – 436.

Yang y Ho (2007)

Yang, H. y Ho, H. "Emergent Standard Of Knowledge Management: Hybrid Peer-To-Peer Knowledge Management". *Computer Standards & Interfaces*. 2007. Volumen 29, Núm. 4, p. 413-422.

Yang (1999)

Yang, Y. "An Evaluation of Statistical Approaches to Text Categorization". *Information Retrieval*. 1999. Vol. 1, p.69-90.

Zack (1999)

Zack, M. "Managing Codified Knowledge". *Sloan Management Review*. 1999. Vol. 40, núm. 4, p.45-58.

Zander y Kogut (1994)

Zander, U. y Kogut, B. "Knowledge and the Speed of the Transfer and Imitation of Organizational Capabilities: An Empirical Test". *Organization Science*. 1994. Vol. 6, núm. 1, p.76-92

Zhang y Faerman (2002)

Zhang, J. y Faerman, S. "Divergent Approaches and Converging Views: Drawing Sensible Linkages between Knowledge Management and Organizational Learning" en *Proceedings of the 36th Hawaii International Conference on System Sciences (HICSS'03)*. 2002. IEEE.

van Zolingen y col. (2001)

Van Zolingen, S.J.; Streumer, J.N. y Stooker, N. "Problems in Knowledge Management: A Case Study of a Knowledge-Intensive Company". *International Journal of Training and Development*. 2001. Vol. 5, núm. 3, p.168-184.