

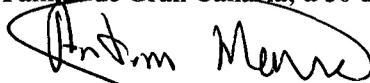
121/2002-03

**UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA
UNIDAD DE TERCER CICLO Y POSTGRADO**

Reunido el día de la fecha, el Tribunal nombrado por el Excmo. Sr. Rector Magfco. de esta Universidad, el/a aspirante expuso esta **TESIS DOCTORAL**.

Terminada la lectura y contestadas por el/a Doctorando/a las objeciones formuladas por los señores miembros del Tribunal, éste calificó dicho trabajo con la nota de **SOBRESALIENTE CUM LAUDE POR UNANIMIDAD**

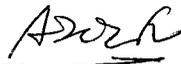
Las Palmas de Gran Canaria, a 30 de septiembre de 2003.



El/la Presidente/a: Dr. D. Antonio Marrero Hernández



El/la Secretario/a: Dr. D. Jorge Rodríguez Díaz



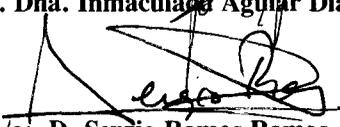
El/la Vocal: Dr. D. Adolfo Blanco Martínez



El/la Vocal: Dra. Dña. Mª Rosario Ferriz Marcén



El/la Vocal: Dra. Dña. Inmaculada Aguiar Díaz



El/la Doctorando/a: D. Sergio Ramos Ramos

UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA
Doctorado en Ciencias Económicas y Empresariales



Facultad de Ciencias Económicas y Empresariales
Departamento de Economía y Dirección de Empresas

TESIS

**SISTEMA AVANZADO DE GESTIÓN DEL CONOCIMIENTO
APLICADO A LA COMPETITIVIDAD INTERNACIONAL**

Sergio Ramos Ramos

2003

UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA
Doctorado en Ciencias Económicas y Empresariales



Departamento de Economía y Dirección de Empresas
Programa de Doctorado: Economía y Dirección de Empresas

Título de la Tesis

**SISTEMA AVANZADO DE GESTIÓN DEL CONOCIMIENTO
APLICADO A LA COMPETITIVIDAD INTERNACIONAL**

Tesis Doctoral presentada por D. Sergio Ramos Ramos
Dirigida por el Dr. D. Juan Manuel García Falcón

El Director

A highly stylized, dense signature in black ink, consisting of many overlapping loops and vertical strokes, positioned above a horizontal line.

El Doctorando

A handwritten signature in black ink, appearing to read 'Sergio Ramos', with a large, sweeping flourish at the end.

Las Palmas de Gran Canaria, a 9 de Septiembre de 2003



UNIVERSIDAD
DE
LAS PALMAS DE GRAN CANARIA



D. GONZALO DÍAZ MENESES, SECRETARIO DEL DEPARTAMENTO DE ECONOMÍA Y DIRECCIÓN DE EMPRESAS DE LA UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA,

CERTIFICA,

Que, el Consejo de Doctores del Departamento en su sesión de fecha 10 de julio de dos mil tres, tomó el acuerdo de dar el consentimiento para su tramitación, a la tesis doctoral titulada “Sistema Avanzado de Gestión del Conocimiento aplicado a la Competitividad Internacional” presentada por el doctorando D. Sergio Ramos Ramos y dirigida por el Doctor D. Juan Manuel García Falcón.

Y para que así conste, y a efectos de lo previsto en el Artº 73.2 del Reglamento de Estudios de esta Universidad, firmo la presente en Las Palmas de Gran Canaria, a 14 de julio de dos mil tres.



**SISTEMA AVANZADO DE GESTIÓN DEL CONOCIMIENTO
APLICADO A LA COMPETITIVIDAD INTERNACIONAL**

*A Leocadia Ramos
mi origen*

y

*A Iris Ramos
mi destino*

Esta investigación ha sido posible gracias al trabajo desinteresado de muchas personas y organizaciones que hacen posible el desarrollo tecnológico y dibujan, con su trabajo, el futuro de nuestra sociedad. En particular, quisiera agradecer la colaboración prestada por Ernest Friedman-Hill del Sandia National Laboratories en Canada, el Stanford Medical Informatics del Stanford University School of Medicine, al equipo de desarrollo de Yale en la Universidad de Dourmund, Alemania; a los miembros del Weka Machine Learning Project de la Universidad de Waikato, Nueva Zelanda; y al equipo internacional de desarrollo de JADE. Y también, a todos los compañeros del Departamento de Economía y Dirección de Empresas por sus muestras de apoyo y motivación. Públicamente, quisiera también agradecer el apoyo constante y continuo de Rosario, cuyos amplios y valiosísimos conocimientos han permitido diseñar y mantener de forma tan exhaustiva la arquitectura y el contenido del sistema avanzado de gestión del conocimiento aplicado a la competitividad internacional, propuesto en esta tesis. Gracias a mi director, Juan Manuel García Falcón, por su orientación científica, y por la confianza que desde un principio depositó en esta investigación.

Las fuentes de estímulo para llevar a cabo esta investigación se las debo especialmente a las personas que comparten mi vida personal y que han vivido de cerca esta investigación. Me refiero a Carmelo, Chari, Leocadio, Araceli, Milagrosa, German y a nuestra madre, así como, a los que hoy no pueden estar presentes. Mención especial a Susa y a Iris que reconfortan y facilitan afectivamente, la ardua tarea de este investigador.

ÍNDICE DE CONTENIDOS

INTRODUCCIÓN	1
 CAPÍTULO 1. EL ESTUDIO DE LA COMPETITIVIDAD INTERNACIONAL	
1.1. IMPORTANCIA DEL ESTUDIO DE LA COMPETITIVIDAD DE LAS NACIONES	11
1.2. MODELOS TEORICOS DE LA COMPETITIVIDAD DE LAS NACIONES	15
1.2.1. Modelos del comercio internacional	15
1.2.1.1. El mercantilismo	16
1.2.1.2. La ventaja absoluta.....	16
1.2.1.3. La ventaja comparativa	17
1.2.1.4. La dotación de factores	18
1.2.1.5. La Paradoja de Leontief	19
1.2.1.6. El ciclo del producto.....	21
1.2.1.7. La similitud del país	22
1.2.1.8. Las economías de escala.....	22
1.2.2. Modelos de la ventaja competitiva de las naciones.....	23
1.2.2.1. El diamante de la ventaja competitiva nacional.....	25
1.2.2.2. El doble diamante generalizado de Moon, Rugman y Verbeke	32
1.2.2.3. Los nueve factores de Cho.....	35
1.3. MODELOS EMPÍRICOS DE LA COMPETITIVIDAD DE LAS NACIONES	35
1.3.1. El modelo del <i>World Economic Forum</i>	36
1.3.2. El modelo del <i>International Institute for Management Development</i>	42
1.3.3. El modelo de la <i>Universidad de Las Palmas de Gran Canaria</i>	48

CAPÍTULO 2. TECNOLOGÍAS APLICADAS A LA GESTION DEL CONOCIMIENTO

2.1.	LA GESTION DEL CONOCIMIENTO Y LAS TECNOLOGÍAS DE LA INFORMACIÓN	59
2.2.	LOS SISTEMAS DE INFORMACIÓN E INTERNET.....	63
2.2.1.	Semántica y Ontología.....	65
2.2.2.	Agentes Inteligentes.....	67
2.2.3.	Introducción al Conocimiento Explícito	74
2.2.4.	Representación del Conocimiento.....	76
2.3.	REPRESENTACIÓN DEL CONOCIMIENTO BASADO EN SÍMBOLOS.....	78
2.3.1.	Representación Lógica.....	81
2.3.2.	Representación Procedural.....	87
2.3.3.	Representación en Redes.....	89
2.3.4.	Representaciones Estructuradas	91
2.3.5.	Representación Simbólica Híbrida	95
2.4.	REPRESENTACIÓN NO-SIMBÓLICA DEL CONOCIMIENTO.....	96
2.4.1.	Redes Neuronales.....	97
2.4.2.	Sistemas Evolutivos	102
2.4.3.	Representación No-Simbólica Híbrida	105
2.5.	INFERENCIA DEL CONOCIMIENTO.....	106
2.5.1.	Sistemas de Búsqueda	107
2.5.2.	Inferencia Distribuida.....	108
2.5.3.	Otros tipos de Inferencias	109
2.6.	DESCUBRIR CONOCIMIENTO	109
2.6.1.	Metodologías de aplicación.....	112
2.6.2.	Fuentes de datos.....	113
2.7.	FUNCIONES BÁSICAS DE LA MINERÍA DE DATOS	118
2.7.1.	Análisis de Asociación.....	119
2.7.2.	Clasificación o predicción.....	121
2.7.3.	Análisis de cluster	130

2.8.	TENDENCIAS TECNOLÓGICAS.....	136
2.8.1.	La Web Semántica	137
2.8.2.	E-Services	140
2.8.3.	Comunicaciones y datos.....	141
2.8.4.	Informática Móvil.....	142
2.8.5.	Informática Ubicua.....	143

CAPÍTULO 3. SISTEMA AVANZADO DE GESTION DEL CONOCIMIENTO APLICADO A LA COMPETIVIDAD INTERNACIONAL

3.1.	FUNDAMENTOS BÁSICOS DEL SISTEMA.....	151
3.1.1.	Orígenes del sistema.....	151
3.1.2.	Estructura del contenido del informe base	154
3.1.3.	Ventajas y desventajas del informe base	157
3.2.	METODOLOGÍA DE DESARROLLO DEL SISTEMA	161
3.2.1.	Introducción a la metodología	162
3.2.2.	Metodologías para la fase de implementación	167
3.3.	PROPUESTA TECNOLÓGICA DEL SISTEMA	169
3.3.1.	Componentes básicos del sistema.....	172
3.3.2.	Protége-2000	172
3.3.3.	Jess	176
3.3.4.	JADE.....	178
3.3.5.	Yale-Weka	182
3.3.6.	Métodos de integración	186
3.4.	SIMULACIÓN PRÁCTICA DEL SISTEMA	187
3.4.1.	Proceso de conversión de datos a modelo ontológico.	188
3.4.2.	Administración del modelo ontológico	190
3.4.3.	Integración del motor de inferencia	192
3.4.4.	Vinculación de las <i>facts</i> al modelo ontológico	193
3.4.5.	Producción de código ontológico de agentes	196
3.4.6.	Descubrimiento de un árbol de decisión.....	198
3.4.7.	Conversión de árbol de decisión a regla de decisión	204

3.5. ACCESOS Y USUARIOS DEL SISTEMA	207
3.5.1. Modalidades de acceso	207
3.5.2. Tipología de usuarios del sistema.....	208
3.5.3. Enfoques aplicados al diseño ontológico	210
CONCLUSIONES Y RECOMENDACIONES	215
BIBLIOGRAFÍA.....	227

ANEXOS

Anexo I. Relación de países utilizados en el estudio

Anexo II. Relación de variables utilizadas en el estudio

ÍNDICE DE CUADROS

Cuadro 1.1	Ejemplo de la ventaja absoluta.....	16
Cuadro 1.2	Ejemplo de la ventaja comparativa.....	17
Cuadro 1.3	Requerimientos de capital y trabajo.....	20
Cuadro 2.1	Lista de productos para el desarrollo de multi-agentes.....	73
Cuadro 2.2	Técnicas de representación simbólica.....	81
Cuadro 2.3	Tabla de Verdad.....	82
Cuadro 2.4	Posibles interpretaciones de operadores modales.....	87
Cuadro 2.5	Representación del grafo ejemplo en KIF.....	91
Cuadro 2.6	Ejemplo de Script de un restaurante japonés.....	94
Cuadro 2.7	Técnicas de representación no simbólica.....	97
Cuadro 2.8	Método evolutivo generalizado.....	104
Cuadro 2.9	Representación de los registros de varias transacciones.....	117
Cuadro 2.10	Tipos de métodos de clasificación.....	122
Cuadro 2.11	Método general de ID3 para generación árbol de decisión.....	124
Cuadro 2.12	Método general de aprendizaje de una red neuronal.....	128
Cuadro 2.13	Tipos de métodos de clustering.....	133
Cuadro 2.14	Proceso general particionado.....	134
Cuadro 3.1	Clasificación de criterios por factores y subfactores.....	153
Cuadro 3.2	Clasificación de criterios por factores.....	153
Cuadro 3.3	Ficha técnica de la investigación de naturaleza primaria.....	154
Cuadro 3.4	Clasificación de criterios según metodología IMD.....	155
Cuadro 3.5	Clasificación de criterios según metodología WEF.....	156
Cuadro 3.6	Requerimientos del Sistema propuesto.....	166
Cuadro 3.7	Ejemplo de una <i>fact</i> expresada en LISP.....	177
Cuadro 3.8	Métodos disponibles en el entorno Yale.....	185
Cuadro 3.9	Métodos disponibles en el entorno del sistema.....	186
Cuadro 3.10	Tabla de datos para generar el árbol de decisión.....	199
Cuadro 3.11	Proceso 7 – Reglas de decisión transcritas a JessTab.....	205
Cuadro 3.12	Relación de enfoques en el diseño ontológico.....	212

ÍNDICE DE FIGURAS

Figura 1.1	Modelos de evaluación de la competitividad de las naciones.....	14
Figura 1.2	Teorema de Heckscher-Ohlin.....	19
Figura 1.3	Comercio sin economías de escala.....	22
Figura 1.4	Comercio con economías de escala.....	23
Figura 1.5	Los determinantes de la ventaja competitiva nacional.....	24
Figura 1.6.	El diamante como sistema.....	29

Figura 1.7.	El cluster italiano de la moda y el calzado.....	30
Figura 1.8.	Fuentes de la ventaja competitiva.....	31
Figura 1.9	El doble diamante generalizado.....	32
Figura 1.10	El modelo de los nueve factores.....	34
Figura 1.11	Determinantes de la productividad.....	38
Figura 1.12	El cubo competitivo del IMD.....	43
Figura 1.13	Evolución del sistema de valores.....	47
Figura 1.14	Fuentes de información, objetivos y análisis.....	52
Figura 2.1	Sistema de pago por teléfono mobipay.....	65
Figura 2.2	Portal comunitario.....	71
Figura 2.3	Estructura jerárquica de la semántica basada en ontología.....	79
Figura 2.4	Representación gráfica del operador lógico Intersección.....	83
Figura 2.5	Representación gráfica del operador lógico Unión.....	84
Figura 2.6	Representación gráfica del operador lógico Complemento.....	84
Figura 2.7	Forma gráfica de un ejemplo de Grafo Conceptual.....	90
Figura 2.8	Representación gráfica de Red Neuronal Artificial.....	98
Figura 2.9	Diagrama esquemático y ecuaciones de la neurona.....	99
Figura 2.10	Representación gráfica del espacio de soluciones.....	103
Figura 2.11	La evolución de la tecnología de la base de Datos.....	111
Figura 2.12	Representación gráfica simulación de un cubo 4 dimensiones....	112
Figura 2.13	La minería de datos como paso hacia descubrir conocimiento ...	113
Figura 2.14	Representación gráfica del modelo de Bases de Datos Relac....	115
Figura 2.15	Diagrama simplificado de relaciones de entidades.....	116
Figura 2.16	Ejemplo de árbol de decisión.....	123
Figura 2.17	Esquema de red neuronal multicapa de alimentación.....	127
Figura 2.18	Diagrama de representación de tendencias.....	137
Figura 2.19	Diagrama de representación de tendencias en lenguajes.....	138
Figura 2.20	Evolución del uso de los sistemas informáticos.....	144
Figura 3.1	Fases del modelo de desarrollo en cascada.....	163
Figura 3.2	Propuesta del Sistema.....	170
Figura 3.3	Protégé-2000 en su primera forma de acceso.....	173
Figura 3.4	Protégé-2000 con el plugin Ontology Bean Generador.....	174
Figura 3.5	Protégé-2000 en su segunda forma de acceso.....	174
Figura 3.6	Representación de Protégé-2000 en su tercera forma.....	175
Figura 3.7	Cuadro Diálogo de entrada de Protégé-2000.....	176
Figura 3.8	Ventana de la consola de Jess en el entorno de Protégé-2000...	178
Figura 3.9	Interfaz gráfico de usuario de JADE.....	179
Figura 3.10	Interfaz gráfico de usuario de Yale.....	182
Figura 3.11	Simple ejemplo de una cadena de operadores.....	184
Figura 3.12	Proceso 1 – Conversión de datos a modelo ontológico.....	190
Figura 3.13	Proceso 2a – Modificación del modelo ontológico.....	191
Figura 3.14	Proceso 2b.....	192
Figura 3.15	Proceso 3 – Integración de Jess a Protégé-2000.....	193
Figura 3.16	Proceso 4a – Portabilidad desde el Protégé-2000 a Jess.....	194
Figura 3.17	Proceso 4b.....	194

Figura 3.18	Proceso 4c – Modificación del nombre USA a United Status.....	195
Figura 3.19	Confirmación del cambio automático del nombre en Jess	195
Figura 3.20	Proceso 5a – Generación código JAVA para JADE.....	197
Figura 3.21	Proceso 5b – Código JAVA generado desde Protégé-2000	197
Figura 3.22	Proceso 6a – Árbol de la cadena de operador del experimento...	200
Figura 3.23	Proceso 6b – Fichero XML descriptor del experimento.....	201
Figura 3.24	Proceso 6c – Representación gráfica del experimento	202
Figura 3.25	Proceso 6d – Árbol de decisión resultante del experimento.....	203
Figura 3.26	Proceso 7 – Resultados de la inferencia en JessTab.....	206
Figura 3.27	Áreas de influencia de la Ingeniería Ontológica	209

INTRODUCCIÓN

En un mundo cada vez más globalizado, interconectado y basado en el conocimiento, la competitividad parece haberse convertido en un factor clave del crecimiento económico y sostenible de los territorios. El conocimiento sobre qué factores influyen en las capacidades competitivas de estos territorios es fundamental para llegar a entender y gestionar dicha competitividad, aplicar políticas públicas y privadas, y entender el comportamiento de los competidores. Esta apreciación es compartida tanto por los investigadores y académicos, como por las organizaciones públicas y privadas. Como resultado de ello, la gestión y la generación de conocimiento sobre competitividad internacional de los territorios se ha convertido en un aspecto clave, en estos momentos, donde se impone un mercado global, dinámico, y sumergido en la incertidumbre, y donde el avance y el desarrollo de los territorios están directamente relacionados con los factores determinantes que mejoran su competitividad. Los estudios empíricos de competitividad internacional existentes en la actualidad, requieren de la evaluación y el análisis de un número cada vez mayor de criterios cuantificables (*hard*) y no cuantificables (*soft*). Tales criterios, como puede ser la evolución del producto interior bruto, son actualmente utilizados para analizar los factores determinantes de la competitividad de dichos territorios. Entre estos factores determinantes podemos encontrar: la economía, el gobierno, la infraestructura, la tecnología, la educación y el medio ambiente, entre otros.

Los trabajos dirigidos a la evaluación de la competitividad de los territorios, como los del *World Economic Forum* (WEF), el *International Institute For Management Development* (IMD), o el de la *Universidad de Las Palmas de Gran Canaria* (ULPGC), utilizan grandes volúmenes de datos que, en algún caso, supera los 300 criterios o variables, por cada unidad territorial y año. Si el objetivo es la optimización del rendimiento de este importante volumen de datos, y la gestión eficiente del conocimiento resultante, se hace imprescindible la utilización de tecnologías avanzadas de información. En la actualidad, el *state of arts* de estas tecnologías

facilitaría y extendería la gestión y generación de conocimiento sobre competitividad, más allá del rendimiento que se obtiene a través de los estudios de competitividad actuales. Esta extensión puede orientarse hacia dos direcciones bien diferenciadas, por un lado, hacia la gestión de los datos y la información, y, por otro, hacia la gestión de los modelos de conocimiento. De esta forma, se puede convertir un documento físico como el informe de competitividad del *World Economic Forum*, en un sistema avanzado de gestión del conocimiento sobre competitividad internacional de los territorios.

Tecnologías como el lenguaje XML (eXtensible Markup Language), que en la actualidad están configurando la futura *World Wide Web* (WWW), también llamada Web Semántica, y muchas aplicaciones informáticas, han surgido como solución al problema del exceso de información publicado en la WWW. De esta forma, los navegadores y los programas informáticos, como los sistemas de búsqueda de información, pasan de gestionar información a gestionar contenido. En este sentido, el exceso de información se contraresta con la gestión de contenidos haciendo, por ejemplo, las tareas de búsqueda de información, que todavía se basan en búsquedas sintácticas, en búsquedas semánticas basadas en contenidos. Para que este tipo de operaciones puedan tener éxito, el usuario y el sistema de búsqueda deben compartir la definición de los conceptos, es decir, deben compartir el mismo modelo ontológico. No obstante, y en menor medida, este exceso de información no sólo se encuentra en la Web, también se puede encontrar en informes de competitividad como los que tratamos en esta tesis. Por lo tanto, esta tecnología puede tener un ámbito de aplicación que traspasa la propia Web. Sin embargo, existe también la posibilidad tecnológica de descubrir nuevo conocimiento inmerso en el mar de datos de este tipo de informes de competitividad, a través de la inferencia del conocimiento sistematizado ó mediante tecnologías como la minería de datos. Y por último, no se debe obviar la automatización como método para extender las capacidades humanas en el desarrollo de tareas. El conocimiento que se posee en la actualidad, sobre los niveles de automatización posibles a través de la tecnología, es bastante significativo y sobrepasa la capacidad imaginativa de cualquier usuario.

En esta misma dirección, no se tiene constancia de que los desarrollos tecnológicos existentes para la gestión y creación de conocimiento estén siendo aplicados al área

de investigación objeto de estudio, la competitividad internacional de los territorios, tan necesaria y de tanto interés global. Este vacío, se hace de igual forma patente, en muchas publicaciones y trabajos de investigación enmarcados dentro del área de Gestión del Conocimiento. Un alto porcentaje de estos trabajos, que estudian la Gestión del Conocimiento en las organizaciones, eluden o minimizan la importancia que tienen las Tecnologías de la Información y las Comunicaciones (TICs) en dicha gestión. A su vez, y desde hace muchos años, ha quedado patente, que las capacidades humanas están estrechamente ligadas a las posibilidades que ofrecen las herramientas y las máquinas existentes en cada momento. Por lo tanto, resulta difícil conocer o describir las capacidades de la gestión del conocimiento en las organizaciones, sin considerar las herramientas tecnológicas existentes o posibles, que permitan definir los límites y capacidades de dicha gestión, y llevar a cabo su implementación. Esto exige combinar el conocimiento existente sobre el ser humano, sus herramientas o máquinas y el área de investigación que se analiza en esta investigación como es la competitividad internacional de los territorios.

Es por ello, que la creciente necesidad de conocer y modelar el conocimiento sobre la competitividad de los territorios, y las posibilidades que presentan los avances tecnológicos de última generación, aplicados a la gestión del conocimiento, han sido las fuentes de inspiración que han despertado el interés por llevar a cabo esta investigación. Además, la combinación estratégica de estas dos áreas de conocimiento abre un abanico de posibilidades que permite avanzar en el conocimiento sobre la competitividad internacional, pero también avanza en el conocimiento sobre las posibilidades tecnológicas de las potentes herramientas utilizadas. Hoy en día, un alto porcentaje de los gastos en tecnologías de información y comunicaciones, que precisan las organizaciones actuales, están dirigidos hacia la integración tecnológica. Y este porcentaje del gasto tiende a crecer si no se crean estándares compartidos por los diferentes sistemas de información a nivel global. Con todo ello, las tecnologías se convierten en elementos básicos para la extensión de nuestras capacidades cognitivas. En este sentido, esta extensión disponible a través de las tecnologías no niega sino extiende las capacidades humanas.

Objetivo de la investigación

En el desarrollo de nuestro trabajo, una vez revisados los modelos teóricos y empíricos de la competitividad internacional de los territorios en el primer capítulo, así como las tecnologías avanzadas aplicables a la gestión del conocimiento en el capítulo dos, en el capítulo tres se propone el siguiente objetivo de investigación:

“Definir y desarrollar desde un punto de vista teórico-conceptual, con una perspectiva tecnológica, un sistema software basado en los modelos de competitividad internacional, que cubra y extienda sus fortalezas, que elimine sus debilidades, y que funcione como plataforma modular, autónoma y abierta para el desarrollo de bases de conocimiento actualizable, y que permita, a su vez, el descubrimiento automático de nuevo conocimiento para la competitividad internacional de los territorios”

Importancia de la investigación

Este trabajo de investigación, crea un puente entre dos áreas de conocimiento, la competitividad internacional y las tecnologías de información de vanguardia aplicadas a la gestión del conocimiento. A su vez, propone un sistema, único en su naturaleza, puesto que combina una serie de componentes tecnológicos, de última generación, que no han sido antes propuestos. Esta interconexión resulta importante porque muestra una aplicación empírica de los métodos tecnológicos, que facilitan la gestión de conocimiento sobre competitividad a través de las tecnologías de la información de última generación. Asimismo, en los últimos años, se ha experimentado un rápido incremento del número de tecnologías abiertas, orientadas a la gestión del conocimiento, lo que produce una mayor integración.

Estructura de la Tesis

Para alcanzar el objetivo propuesto en esta investigación, se ha estructurado la tesis en los siguientes capítulos:

En el capítulo 1, *El estudio de la competitividad internacional*, después de revisar el concepto de competitividad de las naciones se revisan las principales teorías económicas relacionadas con el comercio internacional: el mercantilismo, la ventaja absoluta, la ventaja comparativa, la dotación de factores, la paradoja de Leontief, el ciclo del producto y las economías de escala. A continuación, se analizan las principales teorías sobre la ventaja competitiva de las naciones: el modelo del diamante de la ventaja nacional de Michael Porter, el doble diamante generalizado de Moon, Rugman y Verbeke y el modelo de nueve factores de Cho. Asimismo, en este capítulo se analizan los modelos empíricos de competitividad internacional del *World Economic Forum*, *International Institute for Management Development* y *Universidad de Las Palmas de Gran Canaria*.

En el capítulo 2, *Tecnologías aplicadas a la gestión del conocimiento*, se hace una revisión teórica de las tecnologías aplicadas a la gestión del conocimiento, así como de la importancia de la semántica, la ontología, los agentes inteligentes y el conocimiento en las organizaciones. Asimismo, se realiza una revisión de las técnicas de representación de conocimiento explícito más importantes, haciendo una distinción entre las técnicas de representación del conocimiento basado en símbolos, y las técnicas de representación del conocimiento no-simbólico, se revisan a su vez los métodos de inferencias. A continuación, se revisan las técnicas de descubrimiento de nuevo conocimiento, y de las funciones básicas de la minería de datos. Para terminar este capítulo se realiza una revisión de las tendencias tecnológicas en las áreas como Web Semántica, *e-services*, comunicaciones y datos y, por último, informática móvil y ubicua.

En el capítulo 3, *Sistema de gestión del conocimiento aplicado a la competitividad internacional*, se realiza, en una primera parte, una revisión de los fundamentos del sistema basados en los informes de competitividad internacional. En una segunda parte, se estudia la metodología aplicada al desarrollo del modelo y, seguidamente, se propone una solución tecnológica así como las herramientas o componentes que la conforman y los diferentes métodos de integración que permite. Posteriormente, se realiza una simulación práctica del sistema avanzado de gestión del conocimiento aplicado a la competitividad internacional de los territorios. Finalmente, se proponen los niveles de usuarios que pueden hacer usos aplicados al sistema.

Por último, se presentan las conclusiones finales de nuestro trabajo donde se describen los principales resultados obtenidos en el desarrollo de esta tesis, así como las principales implicaciones académicas y prácticas, las limitaciones de la investigación, y se relacionan las principales recomendaciones para futuras investigaciones. Posteriormente, se incluyen las fuentes bibliográficas utilizadas, así como los anexos referenciados en la tesis.

Capítulo 1

El estudio de la competitividad internacional

“El equilibrio se ha convertido en un tipo de santo sacramento de las economías y ha desviado seriamente la atención del mundo real del flujo Heracliteniano....El sistema económico es una estructura en el espacio-tiempo. Por consiguiente, es algo evolutivo y cambia irreversiblemente”

(Adaptado de Kenneth Boulding)

1.1. Importancia del estudio de la competitividad de las naciones

En la actualidad, la importancia del estudio de la competitividad de las naciones a nivel mundial nunca ha sido tan señalada. En efecto, los niveles de competitividad en las economías basadas en el conocimiento parecen constituir hoy un factor clave del crecimiento económico sostenible. Así, numerosos académicos e investigadores comparten la opinión de que la competitividad se ha convertido en una de las principales preocupaciones de gobiernos e industrias de cada nación (Porter, 1990; Ezeala-Harrison, 1999; Toh y Tan, 1998; Lawton, 1999; Lloyd-Reason y Wall, 2000; Cho y Moon, 2000). Asimismo, el afán de las economías mundiales por alcanzar el éxito económico y mantenerse competitivas, ha resultado en que la competitividad se haya incrementado en casi todo el mundo, obligando a los países, con independencia de su nivel de desarrollo, a tener un mayor entendimiento y control de la competitividad (Toh y Tan, 1998). Además, un aspecto distintivo de este siglo parece ser el efecto de la integración regional en la sociedad basada en el conocimiento, donde el desarrollo económico y social de una región depende esencialmente de la producción, adquisición y gestión del conocimiento (COM, 2000).

Por otra parte, la estructura de la competencia unida a la influencia de la cultura asiática en América y en Europa, ha provocado un mayor dinamismo en la economía internacional (Porter, Takeuchi y Sakakibara, 2000). De hecho, técnicas como la gestión de la calidad total, la mejora continua y el *just-in-time*, procedentes de la forma de actuar de las empresas japonesas, han sido ya difundidas por todo el mundo y utilizadas por múltiples empresas, en su sed por mantenerse competitivas.

Asimismo, esta actitud compartida tanto de académicos e investigadores, como de las instituciones públicas y empresas sobre la competitividad, ha resultado del impacto que han tenido en la economía mundial tres grandes temas: las eliminaciones de barreras comerciales que se han estado produciendo en todo el mundo, que sin lugar a dudas contribuye a reducir el *gap* de desarrollo económico y tecnológico entre las economías territoriales; Internet y la creación de empresas y organizaciones virtuales, que implica indiscutiblemente una nueva forma de competencia; y la frenética

evolución de las Tecnologías de la Información y las Comunicaciones (TICs), que están transformado, como consecuencia, la forma en que comunicamos la información y nos mantenemos competitivos.

Por otra parte, el surgimiento de los sistemas avanzados de TICs y su aplicación a casi cada aspecto de la economía, ha permitido crear una competencia mucho más dinámica, innovadora y eficiente; y menos estática y dependiente de los recursos naturales. En este sentido, la disponibilidad de sistemas de telecomunicaciones a un menor coste y cada vez más eficientes, las conexiones a Internet con mayores anchos de banda, y el desarrollo de la telefonía móvil son algunas de las nuevas prioridades tecnológicas de las naciones que quieren ser competitivas, tales como África del Sur, Méjico o Polonia (Garelli, 2000).

El interés por el estudio de la competitividad nacional se remonta al periodo mercantilista (XVI-XVIII), y muchos académicos e investigadores han ofrecido razonamientos analíticos y teóricos para definir el concepto (Porter, 1987, 1990, 1996; Krugman, 1990; Zysman & Tyson, 1983; Cohen & Zysman, 1987; Tyson, 1992; Toh & Tan, 1998; Lloyd-Reason & Wall, 2000). Sin embargo, resulta difícil analizar la competitividad de un país o de una región sin hacer referencia a otras partes del mundo. Como argumenta Alonso (1992), la competitividad es un término comparativo, no cabe hablar de competitividad de un país si no es en relación con el comportamiento de sus competidores. A pesar de que la experiencia de un país con sus antecedentes históricos y culturales, no pueda ser transferida fácilmente a otros, las comparaciones se convierten en una herramienta útil para las discusiones de políticas macroeconómicas y estructurales de una nación (Lloyd-Reason y Wall, 2000).

Por consiguiente, para poder determinar el nivel de competitividad de una nación necesitamos traducir la teoría a la práctica. En la actualidad, existen dos estudios líderes a escala mundial que analizan y cuantifican los factores que determinan los niveles de competitividad de las naciones. El *Global Competitiveness Report (World Economic Forum)* y el *World Competitiveness Yearbook (International Institute for Management Development)*, ambos producidos en Suiza. Además, en diciembre de 2001, fruto de una tesis doctoral en la Universidad de Las Palmas, surge el primer estudio conocido de competitividad regional a escala internacional, con el objeto de

medir la competitividad internacional de Canarias. Un análisis detallado de estos informes de competitividad ha confirmado el impacto que están teniendo estos indicadores de competitividad en las economías del mundo. Tanto gobiernos, como empresas, académicos, todos utilizan los indicadores de competitividad para formular sus políticas y desarrollar sus estrategias. El impacto que estos indicadores de competitividad tienen en la práctica, va mucho más lejos de lo que nos podríamos imaginar.

La figura 1.1 representa los modelos existentes de análisis de la competitividad, tanto desde el punto de vista teórico como empírico. Dentro de los teóricos, incluimos los principales modelos propuestos sobre el comercio internacional, desde la época mercantilista y Adam Smith, hasta Krugman y Lancaster, con su modelo de economías de escala. Así como, el modelo del diamante de la ventaja nacional propuesto por Michael Porter en su libro *The Competitive Advantage of Nations*¹, y otros modelos que han aparecido con posterioridad y que versan sobre este mismo, nos referimos a, -el doble diamante generalizado de Moon, Rugman y Verbeke (1995) y el modelo de nueve factores de Cho (1994). Con referencia a los modelos empíricos analizaremos el de *World Economic Forum* y el *International Institute for Management Development*, así como el modelo de la Universidad de Las Palmas de Gran Canaria.

¹ Traducido al castellano con el título "La ventaja competitiva de las naciones", Plaza y Janés, Barcelona 1992.

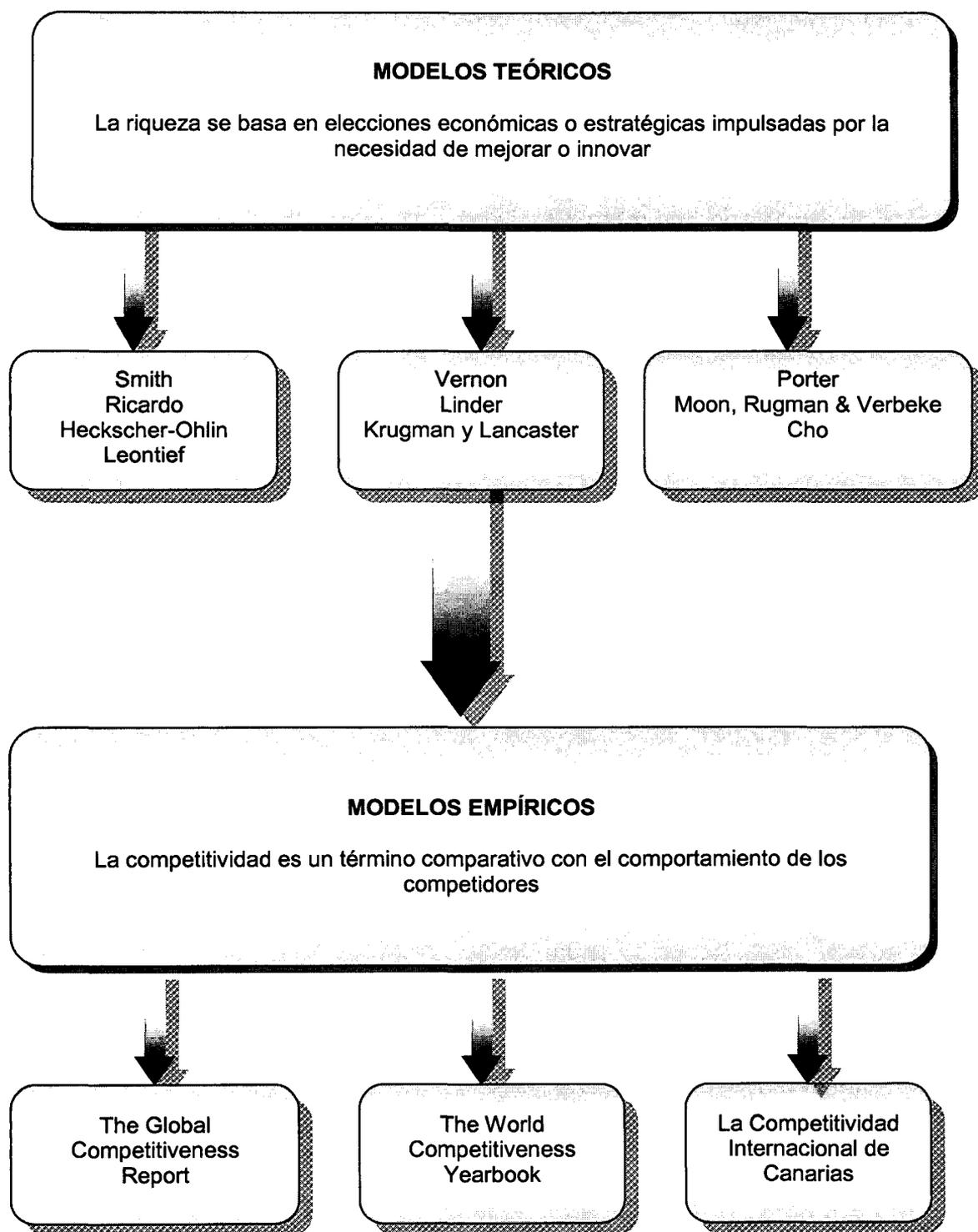


Figura 1.1: Modelos de evaluación de la competitividad de las naciones

1.2. MODELOS TEÓRICOS DE LA COMPETITIVIDAD DE LAS NACIONES:

Iniciaremos este apartado con los principales modelos del comercio internacional para luego presentar posteriormente aquellos basados en la ventaja competitiva de las naciones.

1.2.1. Modelos del comercio internacional

El comercio internacional ocupa un aspecto central de la actividad económica, el crecimiento y los procesos de desarrollo de la mayoría de las sociedades modernas. De hecho, su importancia en el crecimiento económico y en el desarrollo podría haber sido reconocida durante la era mercantilista del pensamiento económico (Ezeala-Harrison, 1999). El *mercantilismo* explicaba que en el comercio no todos los países resultan ganadores. En contraste, la teoría que apareció a continuación, denominada *ventaja absoluta*, consideraba al comercio como un juego de suma positivo donde no existirían perdedores si los países se especializaran en producir aquello que hacen mejor o más barato. Este marco teórico fue ampliado por Ricardo (1971) al plantear la posibilidad de obtener una *ventaja comparativa* en aquellos productos que al país le resultara comparativamente más beneficioso producir para vender a sus socios comerciales, aunque no disfrutase de ventaja absoluta. A su vez, Heckscher y Ohlin (1919,1933), en un intento de superar las limitaciones existentes en las teorías precedentes, consideran que la ventaja comparativa surge de las diferencias en las *dotaciones de factores*. Aunque aparentemente ello parecía evidente, la *Paradoja de Leontief* sugiere que la teoría de dotación de factores no funcionaba bien en el mundo real. Otros economistas han desarrollado teorías alternativas para complementar los modelos precedentes y de esta forma ayudar a entender y a evaluar el debate existente sobre la competitividad. Nos referimos a la teoría del *ciclo del producto*, la *similitud del país*, y el comercio basado en *las economías de escala*. A continuación pasamos a detallar las principales teorías existentes sobre el comercio internacional y su importancia para la competitividad de una nación: el mercantilismo, la ventaja absoluta, la ventaja comparativa, la dotación de factores, la Paradoja de Leontief, el ciclo del producto, la similitud del país y las economías de escala.

1.2.1.1. El mercantilismo

Durante más de dos décadas y hasta el siglo XVIII, perduró un pensamiento económico dominante denominado *mercantilismo*², cuya teoría económica se basaba en promocionar la producción y el comercio de empresarios privados, los cuales alcanzaban un beneficio económico a la vez que contribuían a la consolidación, prosperidad y poder de una nación (Allen, 1988). Para el mercantilismo, el comercio, considerado como un juego de *suma cero*, tenía como finalidad básica el fortalecimiento de la economía doméstica para hacerla más productora de bienes y servicios demandados en el extranjero, a la vez que se incitaba a los ciudadanos del país a ser menos dependientes de los productos fabricados fuera y más demandantes de los productos locales, reduciendo de esta forma el nivel de las importaciones.

1.2.1.2. La ventaja absoluta

Adam Smith (1937), proponía que el comercio era un juego de *suma suma*, donde el superávit comercial de un país no tenía por qué suponer un déficit comercial de otro. Concretamente, los socios comerciales podían beneficiarse del comercio si ambos países se especializan en la producción de aquel bien en el que asumirían un menor coste, es decir, tendrían una ventaja absoluta. En el cuadro 1.1 se ilustra la ventaja absoluta de Adam Smith en dos países, Inglaterra y Portugal, que producen ropa y vino, respectivamente.

BIENES	COSTE DE PRODUCCIÓN	
	En Inglaterra	En Portugal
Ropa	4	6
Vino	8	3

Cuadro 1.1: Ejemplo de la ventaja absoluta

² A los autores del mercantilismo se les relaciona con las ideas de empresarios y profesionales en Inglaterra y algunos altos funcionarios públicos.

En la comercialización de ropa y vino entre Inglaterra y Portugal se puede observar que Inglaterra necesita 4 hombres para fabricar ropa mientras que Portugal 6. En el caso de comercializar vino, a Inglaterra le supone 8 hombres y a Portugal sólo 3. En la teoría de la ventaja absoluta de Adam Smith (1937), Inglaterra fabricará ropa porque le es el bien que le otorga la ventaja absoluta, mientras que el vino será producido por Portugal. Así, mientras que un país tiene ventaja absoluta en la producción de un bien, el otro país tiene ventaja absoluta en la producción de otro. Por lo que, el comercio internacional beneficiaría a ambos países.

1.2.1.3. La ventaja comparativa

Esta nueva teoría denominada la *teoría de la ventaja comparativa de David Ricardo* (1971), consideraba que en el comercio internacional el país superior se especializaría en la producción de aquel bien donde tenía la mayor ventaja absoluta y el país inferior en el bien donde tenía la menor ventaja absoluta. Así, el comercio podía seguir siendo un juego de *suma suma*, aunque se permitía que los países se especializaran en la producción de más de un bien. Como apunta Gandolfo (1998), el argumento de Ricardo de que el comercio internacional tendría lugar incluso si un país era superior a otro en la producción de los dos bienes, fue la contribución más importante del autor a la teoría del comercio internacional. En la siguiente ilustración, se describe la situación de comercio internacional en la que los países pueden beneficiarse de poseer ventaja comparativa.

BIENES	COSTE DE PRODUCCIÓN	
	En Inglaterra	En Portugal
Ropa	4	6
Vino	8	10

Cuadro 1.2: Ejemplo de la ventaja comparativa

Como se puede comprobar, Inglaterra tiene una mayor ventaja relativa (comparativa) en la producción de ropa, ya que su coste unitario es menor que en Portugal en un 33% (2/6); mientras que el coste unitario de vino es menor también en un 20% (2/10). Asimismo, Portugal tiene una pequeña desventaja en la producción de vino: su coste unitario, es mayor que en Inglaterra en un 25% (2/8), mientras que el coste unitario de producir ropa es también mayor en Portugal que en Inglaterra en un 50% (2/4). Por lo tanto, suponiendo que los términos del comercio son mayores que 0,5 y menores que 0,6, la ropa inglesa se intercambiará por el vino portugués para el beneficio de ambos países. Con este ejemplo, Ricardo en su libro *On the Principles of Political Economy and Taxation*³, demostró que la importancia del comercio se basaba simplemente en tener una diferencia relativa y no absoluta en los costes de producción de bienes.

1.2.1.4. La dotación de factores

Este modelo de Heckscher-Ohlin (HO) fue desarrollado a principios del siglo XX por los economistas suecos, Eli Heckscher y Bertil Ohlin⁴, quienes argumentaron que las ventajas comparativas entre países surgen de las diferencias en las dotaciones de los factores. El teorema expone que la ventaja comparativa de un país dependerá de que exporte aquel bien cuya producción es relativamente intensa en el factor con el que está relativamente bien dotado y por el que asume un menor coste.

En una situación anterior al comercio, con una ratio determinada e idéntica de A a B en ambos países, (e.g., la representada por la recta OR), el país 1 estará en el punto H_1 y el país 2 en el punto H_2 de sus respectivas curvas de transformación. La tasa marginal de transformación es mayor en el país 1 que en el país 2 (calculado en H_1 y H_2 , respectivamente). El producto A será relativa. Es decir, el coste de oportunidad de A, en términos de B, es menor en el país 1, el cual es abundante en capital, y tiene

³ Traducido al castellano con el título *Principios de la Economía Política y la Tributación*, Ayuso, Madrid, 1973.

⁴ El trabajo original de Eli Heckscher, publicado en 1919 en sueco, recibió una escasa atención hasta su traducción al inglés en 1949. Por su parte, Bertil Ohlin, un estudiante de Heckscher, elaboró su tesis doctoral en 1924, fundamentándose en las ideas de Heckscher, trabajo que fue también escrito en sueco, y no recibió mucha atención hasta su traducción al inglés en 1933 (Jones, 1988).

una predisposición a producir el producto que es abundante en capital mientras que el país 2, abundante en trabajo, tiene predisposición a producir bienes intensivos en trabajo. En este sentido, cada país puede expandir su producción en el producto que es intensivo en su factor abundante, a un coste de oportunidad menor que el otro.

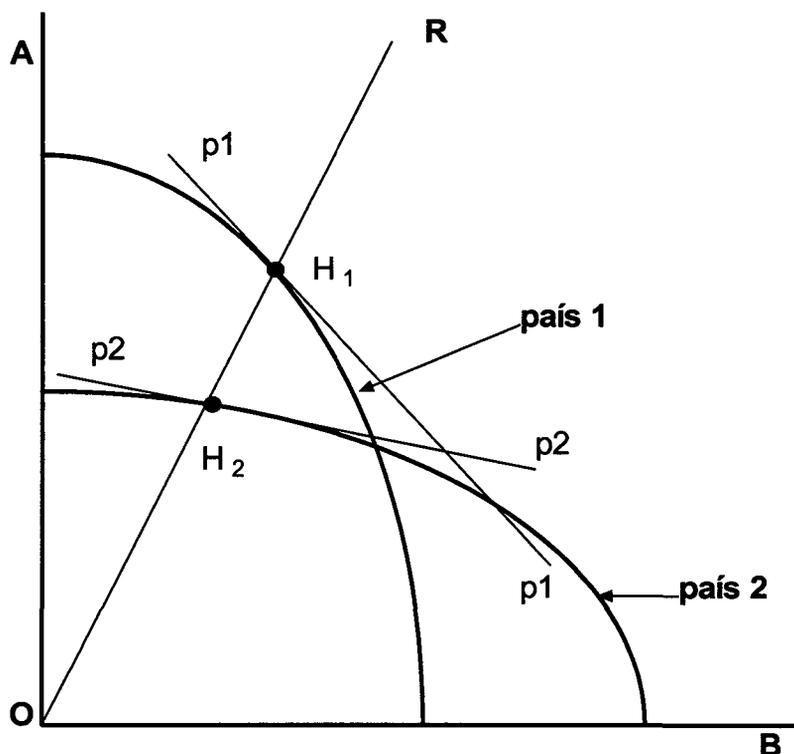


Figura 1.2: Teorema de Heckscher-Ohlin
(Fuente: Gandolfo, 1998)

En consecuencia, parece evidente que cada país debería exportar el producto que utiliza intensamente el factor más abundante. El modelo de HO ha sido ampliado por tres teoremas importantes: 1) Teorema de Igualdad del Precio del Factor, 2) Teorema de Stolper-Samuelson y 3) Teorema de Rybczynski.

1.2.1.5. La Paradoja de Leontief

El estudio empírico más conocido del modelo de HO fue realizado por Leontief (1953), a quien se le otorgó el premio Nóbel en 1973. Este autor previó que EE.UU., el país

del mundo con mayor abundancia de capital, debería exportar los bienes intensivos en capital e importar los bienes intensivos en trabajo, pero descubrió que los bienes competitivos en la importación de EE.UU. requerían un 30% más de capital por trabajador que los bienes que exportaba EE.UU. De acuerdo con sus cálculos, la ratio trabajo-capital era de aproximadamente 14,000 dólares de EE.UU. por trabajador y año en los bienes de exportación, y en torno a 18,000 dólares de EE.UU. por trabajador y año en los bienes que compiten en las importaciones. Este descubrimiento, que resultó ser lo opuesto a lo que predijo el modelo de HO, pasó a considerarse en la literatura como la *Paradoja de Leontief*.

Aplicando su análisis de *input-output* a la economía de EE.UU., Leontief calculó las demandas totales de *input* (directas e indirectas) de capital y trabajo por unidad de combinación del producto “exportación de EE.UU.” y del producto “sustitución de importaciones competitivas de EE.UU.”⁵; en ambos casos, la unidad tenía el valor de un millón de dólares de productos en precios y composición de 1947. El cuadro 1.3 resume los principales descubrimientos de este análisis.

	CAPITAL (K) (dólares, en precios de 1947)	TRABAJO (L) (hombres-años)	K/L
EXPORTACIONES	2.550.780	182.313	13.991
IMPORTACIONES	3.091.339	170.004	18.184

Cuadro 1.3: Requerimientos de capital y trabajo doméstico por millón de dólares de exportación de EE.UU. y de sustitución de importaciones competitivas (1947).

(Fuente: Gandolfo, 1998: 87).

⁵ Por el concepto de *importaciones competitivas* Leontief se refiere a las importaciones de productos que pueden ser y son, al menos en parte, producidas actualmente por las industrias domésticas.

1.2.1.6. El ciclo del producto

Raymond Vernon (1966) argumentó que muchos de los productos fabricados pasan por un ciclo integrado de cuatro etapas sucesivas: introducción, crecimiento, madurez y declive. Así, en la etapa de introducción, la demanda de los nuevos productos tiende a estar basada en factores no relacionados con el precio y las empresas suelen ofertar sus nuevos productos a precios relativamente elevados. En esta etapa inicial del ciclo del producto, la producción suele estar localizada en el país donde la empresa innovadora y donde se servirá al mercado local.

Cuando el nuevo producto ha llegado a tener una continuación en el mercado local, el productor comenzará a entrar en mercados extranjeros, inicialmente exportando productos a esos mercados. Mientras madura el mercado, el producto se convierte en el más estándar y el precio en el método competitivo más relevante. En esta fase de maduración los motivos que subyacen a la localización inicial desaparecen y la empresa empieza a examinar la mejor forma de servir a la demanda extranjera. Por un lado, la empresa puede continuar produciendo los output en casa y exportando la cantidad demandada al extranjero. Por otro lado, la empresa puede dar licencia a productores extranjeros o implicarse directamente en la producción del producto en plantas localizadas en países extranjeros con demanda; en esta fase, los países afectados normalmente son países avanzados. Como consecuencia, el lugar de producción se mueve hacia otros países desarrollados, aunque en fases posteriores del modelo del ciclo del producto, lo hará hacia países menos desarrollados.

El modelo del ciclo de producto resulta útil para reconciliar la Paradoja de Leontief. En la etapa de introducción, la producción de bienes de exportación es intensiva en trabajo, ya que los productores necesitan conocer cómo fabricar el producto de la manera más eficiente y cómo reacciona el mercado ante el nuevo producto. Sin embargo, en la etapa de madurez, las exportaciones son intensivas en capital ya que los productores están más familiarizados con la ingeniería eficiente y reciben retroalimentación del mercado.

1.2.1.7. La similitud del país

Hasta la actualidad, las teorías del comercio han analizado la oferta, siendo la teoría de similitud del país propuesta por Staffan Linder (1961) la que se centra en el estudio de la demanda a partir de dos hipótesis. Por un lado, se plantea que un país exporta aquellos productos fabricados para los que existe un mercado local significativo que es conocido por los productores y, como consecuencia, introduce nuevos productos. Además, la producción para el mercado local debe ser lo suficientemente grande como para que las empresas alcancen economías de escala y puedan reducir sus costes. Así, cada país exporta sus productos a otros países con gustos y niveles de renta similares. Cada país producirá principalmente para sus mercados locales, si bien parte del *output* lo exportará a otros países similares.

1.2.1.8. Las economías de escala

A finales de los años 70, Krugman y Lancaster (1979) desarrollaron, de forma independiente, modelos de comercio para productos diferenciados. Su conclusión es que las economías de escala y el comercio internacional hacen posible que cada país pueda producir bienes de forma eficiente sin sacrificar la variedad de bienes.

Si la industria no ofrece productos diferenciados, el país doméstico es abundante en capital y el país extranjero abundante en trabajo, entonces el país doméstico exportará productos industriales e importará alimentos. La figura 1.3 muestra que en casa se exportará productos industriales igual en valor a los alimentos que importará. La longitud de las líneas indica el valor del comercio en cada dirección.

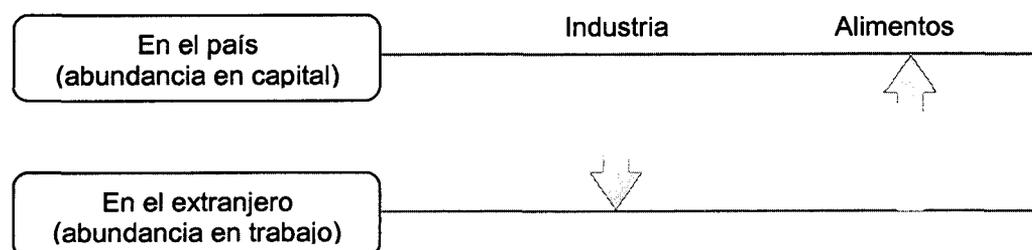


Figura 1.3: Comercio sin economías de escala

(Fuente: Krugman y Obstfeld, 2000: 136)

Si por el contrario, asumimos que la industria comprende sectores competitivos monopolísticos (los productos de cada empresa se diferencian entre sí), el país doméstico seguirá siendo un exportador neto de productos industriales y un importador de alimentos. Sin embargo, las empresas extranjeras en el sector industrial producirán productos diferentes de aquellos que se producen en casa.

La figura 1.4 muestra los modelos del comercio con industrias competitivas monopolísticas. Así, el comercio mundial en un modelo de competencia monopolística podría tener dos partes. Una parte la ocupará el comercio de dos vías en cuanto a la industria; este intercambio de productos industriales se denominará comercio intraindustrial. Y otra parte consistirá en un intercambio de productos industriales por alimentos, denominándose comercio interindustrial.

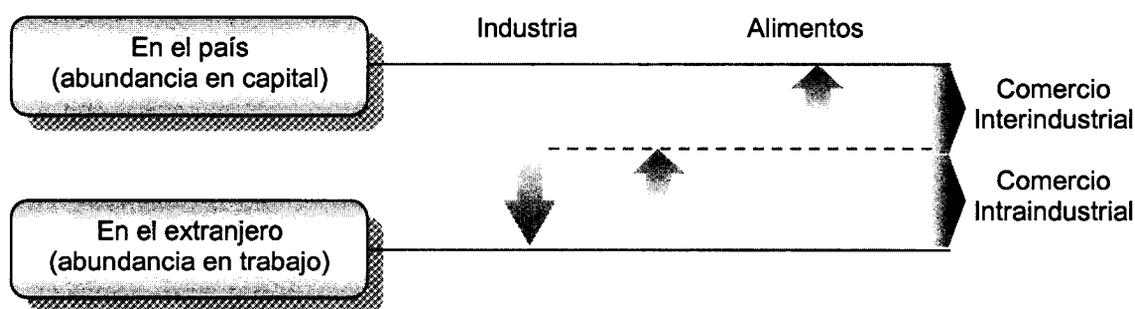


Figura 1.4: Comercio con economías de escala
(Fuente: Krugman y Obstfeld, 2000: 137)

A continuación se presentan los modelos teóricos basados en la ventaja competitiva de las naciones: el modelo del diamante de Porter, el modelo de nueve factores de Cho, y el doble diamante generalizado de Moon, Rugman y Verbeke.

1.2.2. Modelos de la ventaja competitiva de las naciones

Los modelos de la ventaja competitiva de las naciones surgieron para dar respuesta a una economía cada vez más globalizada, donde existe un mayor grado de concentración empresarial, un mayor número de multinacionales que operan en diversos países, una creciente y cada vez más variada demanda y una oferta más diferenciada. El modelo del diamante de Porter de la ventaja competitividad nacional, se presenta a continuación.

1.2.2.1. El diamante de la ventaja competitiva nacional

La teoría de Porter (1990a) de la ventaja competitiva nacional está basada en un análisis de las características del entorno nacional que identifica cuatro grupos de variables que influyen en la capacidad de las empresas para establecer y mantener una ventaja competitiva en los mercados internacionales: condiciones de los factores, condiciones de la demanda, sectores afines y auxiliares, estrategia y estructura y rivalidad en las empresas. El diamante es un sistema que se refuerza mutuamente, de forma que el efecto de un determinante es contingente del estado de otros (véase figura 1.5).

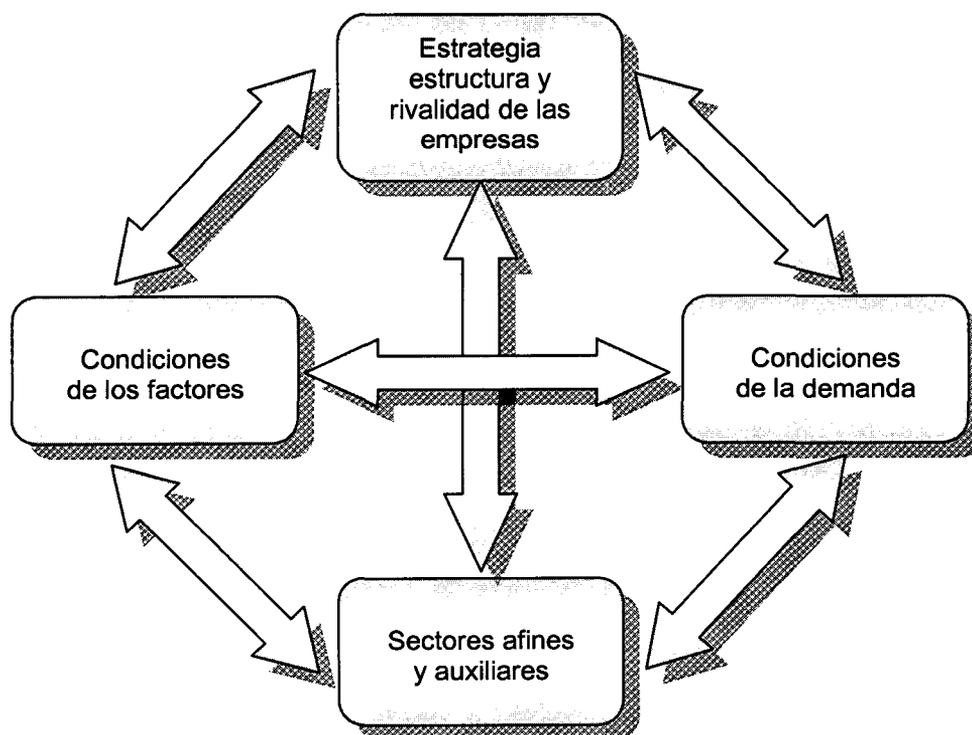


Figura 1.5: Los determinantes de la ventaja competitiva nacional (Fuente: Porter, 1990a:72)

A continuación se describe brevemente los cuatro factores o elementos del diamante.

Las condiciones de los factores

El autor distingue entre dos tipos de factores: los factores básicos y los factores avanzados. Los primeros están más relacionados con el término tradicional e incluyen factores tales como los recursos naturales, el clima, la localización y la demografía. Los segundos, hacen referencia a los factores especializados como las infraestructuras de comunicaciones, la investigación y las capacidades avanzadas, entre otros. Para Porter, los factores avanzados son los más importantes para la ventaja competitiva, puesto que al ser especializados son más difíciles de imitar y surgen de las diferentes inversiones llevadas a cabo por individuos, empresas y gobiernos.

Concretamente, los factores de producción recogidos en el diamante son agrupados en cinco categorías básicas:

- 1 *Recursos humanos.* Representa la cantidad, la capacidad y el coste del personal, teniendo en cuenta las horas normales de trabajo y la ética del mismo.
- 2 *Recursos físicos.* Constituye la abundancia, la calidad, la accesibilidad y el coste de la tierra, el agua...
- 3 *Recursos del conocimiento.* Integra el *stock* de conocimiento científico, técnico y de mercado sobre bienes y servicios, que tiene una nación. Estos recursos proceden de las universidades y centros de investigación públicos o privados, instituciones estadísticas públicas, asociaciones comerciales, bases de datos...
- 4 *Recursos de capital.* Representa la cantidad y el coste de capital disponibles para realizar inversiones en la industria.
- 5 *Infraestructura.* Abarca el tipo, la calidad y el coste de uso de la infraestructura disponible que afecta a la competencia, incluyendo el sistema de transportes, el sistema de comunicaciones, el correo, el

envío de paquetes postales, los pagos o transferencias de fondos, los servicios sanitarios y otros.

Las condiciones de la demanda

Porter asigna una gran importancia a este segundo vértice del diamante, ya que la demanda interior presiona a las empresas a innovar y mejorar. Aunque, el que las empresas de un país consigan o no ventajas competitivas va a depender, no de que exista una gran demanda interior, sino de que esa demanda sea la más refinada y exigente del mundo. La sofisticación de los compradores obliga a las empresas a responder a retos difíciles de afrontar, aunque también les ayuda a prever tendencias mundiales.

Porter (1990) describe la demanda local a través de tres amplios atributos:

1. *La composición de la demanda.* Las empresas que pertenecen a naciones pequeñas centran su estrategia en ofertar sus productos a segmentos de demanda significativos en el mercado mundial. Además, la naturaleza de la demanda local con relación a la sofisticación y exigencia de los compradores juega un papel primordial en la ventaja de las empresas.
2. *El tamaño de la demanda y el patrón de crecimiento.* El tamaño del mercado local puede ocasionar ventajas competitivas en aquellas industrias donde existan economías de escala o de aprendizaje, motivando a las empresas locales a invertir de forma agresiva en productos a gran escala o desarrollo tecnológico.
3. *La internacionalización de la demanda.* Cuanto más internacionalizada esté la demanda, más ventajas otorgará a las empresas locales. Así, si los compradores son grandes multinacionales que operan en varios países, la demanda local se convierte también en demanda extranjera.

Los sectores afines y auxiliares

La presencia en la nación de sectores afines cuyas empresas coordinan o comparten sus actividades de marketing, distribución, servicios o tecnología en la cadena de valor, o de sectores de productos complementarios (eg., ordenadores), conlleva a menudo a la aparición de nuevos sectores competitivos. Por otro lado, Porter argumenta que resulta contraproducente para una empresa o un país el crear proveedores “cautivos”, es decir, totalmente dependientes de la industria nacional e imposibilitados de acceder a satisfacer una posible demanda exterior.

La estrategia de la empresa, la estructura y la rivalidad

Este vértice del modelo predice que las circunstancias nacionales y el contexto originan fuertes tendencias en el modo en que se crean, organizan y gestionan las empresas, así como en definir la naturaleza de la competencia interna. En términos generales, ningún sistema de gestión es universalmente apropiado sino que la competitividad de un determinado sector es consecuencia de la convergencia de los modos de dirección y de organización prevaleciente en cada país y de las fuentes de ventaja competitiva existentes en cada sector. Así, las estrategias de las empresas deben responder y estar basadas en los intereses de la demanda local o extranjera. Además, los sectores locales que son líderes en el ámbito internacional son los que deben conformar las estrategias de las empresas locales.

Por otro lado, la presencia de rivales nacionales fuertes es un importante estímulo para la ventaja competitiva, impulsando a las empresas a innovar y mejorar. La sabiduría convencional argumentaba que la competencia local era muy dañina ya que motivaba la duplicación de esfuerzos e impedía a las empresas obtener economías de escala. Pero, aquellos líderes nacionales que producen grandes volúmenes y reducen sus costes por unidad de producción no son automáticamente competitivos, sino que en realidad crean una eficacia estática ya que el volumen de producción implica poco dinamismo y flexibilidad. Además, la rivalidad interior crea presión en las empresas para que inviertan, reduzcan costes, mejoren la calidad y los servicios e innoven en productos y procesos.

A su vez, a diferencia de la rivalidad extranjera, que tiende a ser analítica y distante, la rivalidad nacional suele ser más personal, por cuanto la competencia no sólo se basa en cuota de mercado, sino que las empresas compiten también por atraer recursos humanos de la competencia, por excelencia técnica, la calidad superior del servicio y la captación de clientela, entre otros. La concentración geográfica magnifica el poder de la rivalidad interior, ya que cuanto más localizada esté la rivalidad, ésta será más intensa y beneficiosa para las empresas que se ven forzadas a innovar y perciben una presión hacia la mejora constante de las fuentes de ventaja competitiva. Al contrario de la sabiduría convencional, la presencia de competidores interiores anula automáticamente los tipos de ventaja que proceden simplemente de estar en una nación determinada (economías de aprendizaje, coste de los factores, posesión de estándares relacionados con el diseño y la tecnología... etc.). Además, la rivalidad local presiona a las empresas para abordar mercados mundiales y las fortalece para tener éxito en los mismos.

El diamante como sistema

El efecto de un determinante de la competitividad incluido en el diamante es contingente del estado de los demás. Por ejemplo, los compradores más exigentes no generarán productos avanzados con sus demandas a menos que la calidad de los recursos humanos permita a las empresas satisfacer las necesidades de estos compradores. Por otro lado, las desventajas de ciertos factores de producción no motivarán la innovación a menos que la rivalidad sea vigorosa y los objetivos de las empresas respalden una inversión continuada.

La figura 1.8 muestra el modelo del diamante como un sistema donde se puede observar que junto a los cuatro atributos o determinantes básicos de la competitividad, se encuentra también el factor del gobierno y el factor del azar.

Los clusters y la competitividad

Los *clusters* representan una manifestación de las relaciones existentes entre los cuatro vértices del diamante. Estas agrupaciones de empresas afectan a la

competencia en tres aspectos: incrementan la productividad de las empresas o sectores que los integran; incrementan su capacidad para innovar y, con ello, su capacidad para aumentar la productividad; y estimulan la creación de nuevas empresas, lo cual apoya la innovación y, por consiguiente, la expansión del *cluster*.

El tercer vértice del diamante –industrias afines y auxiliares– no indica que las industrias locales competitivas en el entorno internacional están en realidad concentradas geográficamente, por ello, resulta necesario comentar este tipo de concentración que el autor denomina *clusters*. Por su dimensión geográfica, un *cluster* puede ser urbano, regional, nacional o incluso internacional. La presencia de los *clusters* deja entrever que buena parte de la ventaja competitiva se encuentra fuera de la empresa, incluso fuera del sector, en las ubicaciones de sus unidades de explotación. Los *clusters* son un motor de las exportaciones y un imán de la inversión extranjera. Constituyen un foro en el cual se pueden establecer nuevas formas de diálogo entre las empresas, los organismos públicos y las instituciones (escuelas, universidades, empresas de suministros públicos).

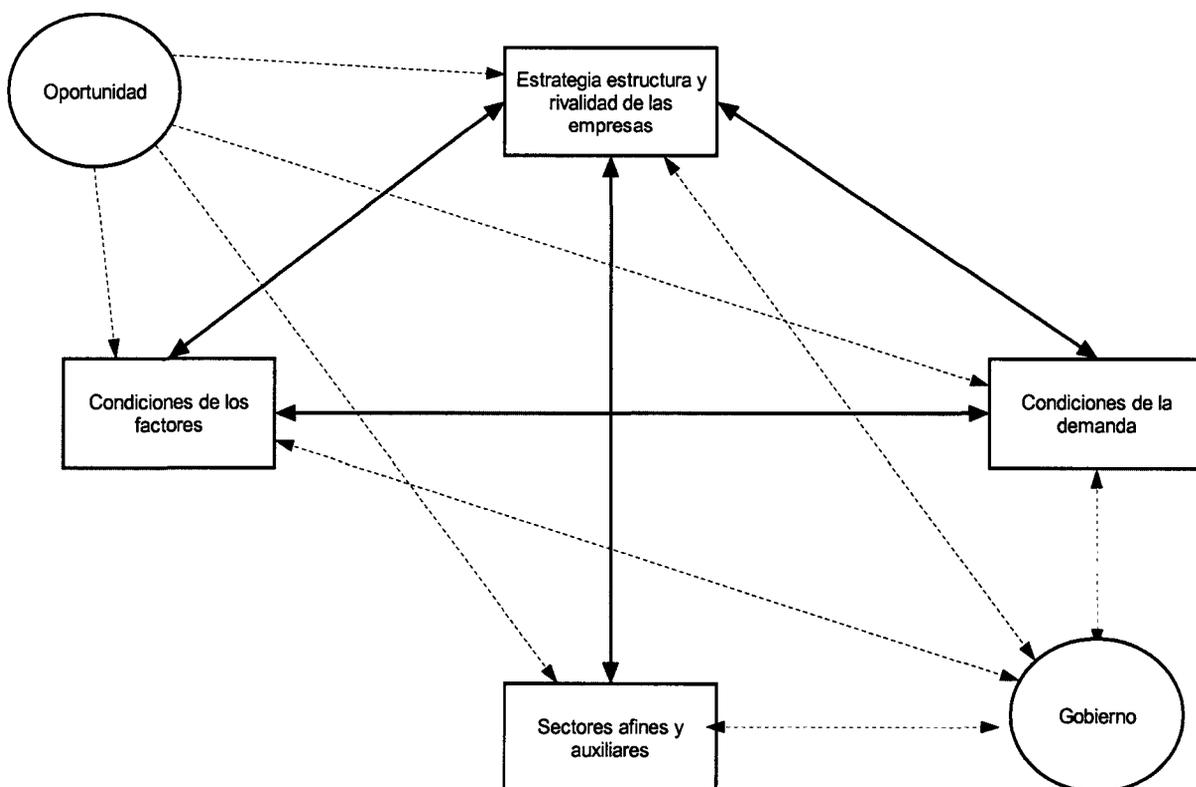


Figura 1.6: El diamante como sistema

(Fuente: Porter, 1990a: 127).

La figura 1.7 presenta un diagrama esquemático del *cluster* italiano de la moda y el calzado de cuero. En él se muestran las diversas cadenas de sectores afines que participan en el *cluster* de la moda y el calzado de cuero de Italia, entre los que se encuentran: los relacionados con los diferentes tipos de productos de cuero (productos complementarios, materiales comunes, tecnologías similares), diferentes tipos de calzados (canales comunes, materiales y tecnologías similares) y diferentes tipos de productos de moda (productos complementarios).

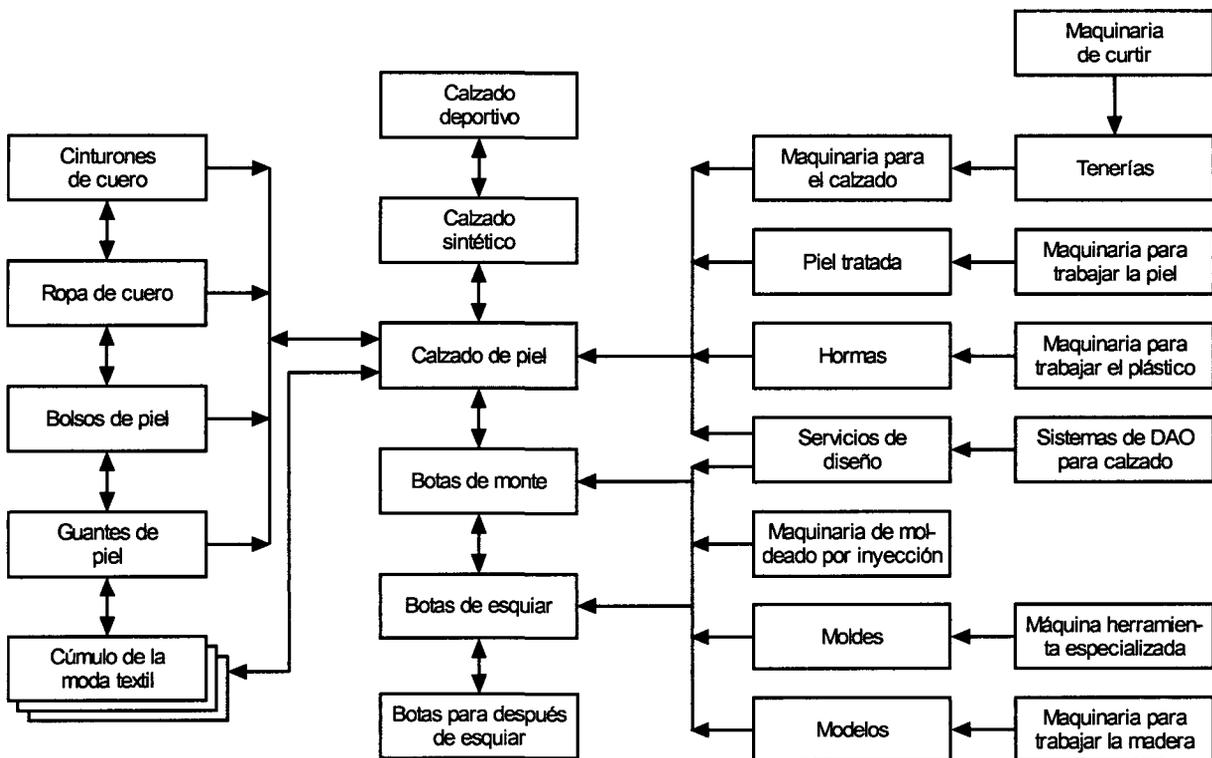


Figura 1.7: El cluster italiano de la moda y el calzado

(Fuente: Porter, 1999: 206).

En la figura 1.8, Porter representa gráficamente un modelo del efecto que tiene la ubicación en la competencia, considerando cuatro elementos interrelacionados. Los factores representados incluyen los activos tangibles (como la infraestructura material), la información, el sistema jurídico y los institutos de investigación universitaria a los

que recurren las empresas para competir. Para incrementar la productividad, los factores han de mejorar en eficiencia, calidad y, en última instancia, especialización para determinadas áreas del *cluster*.

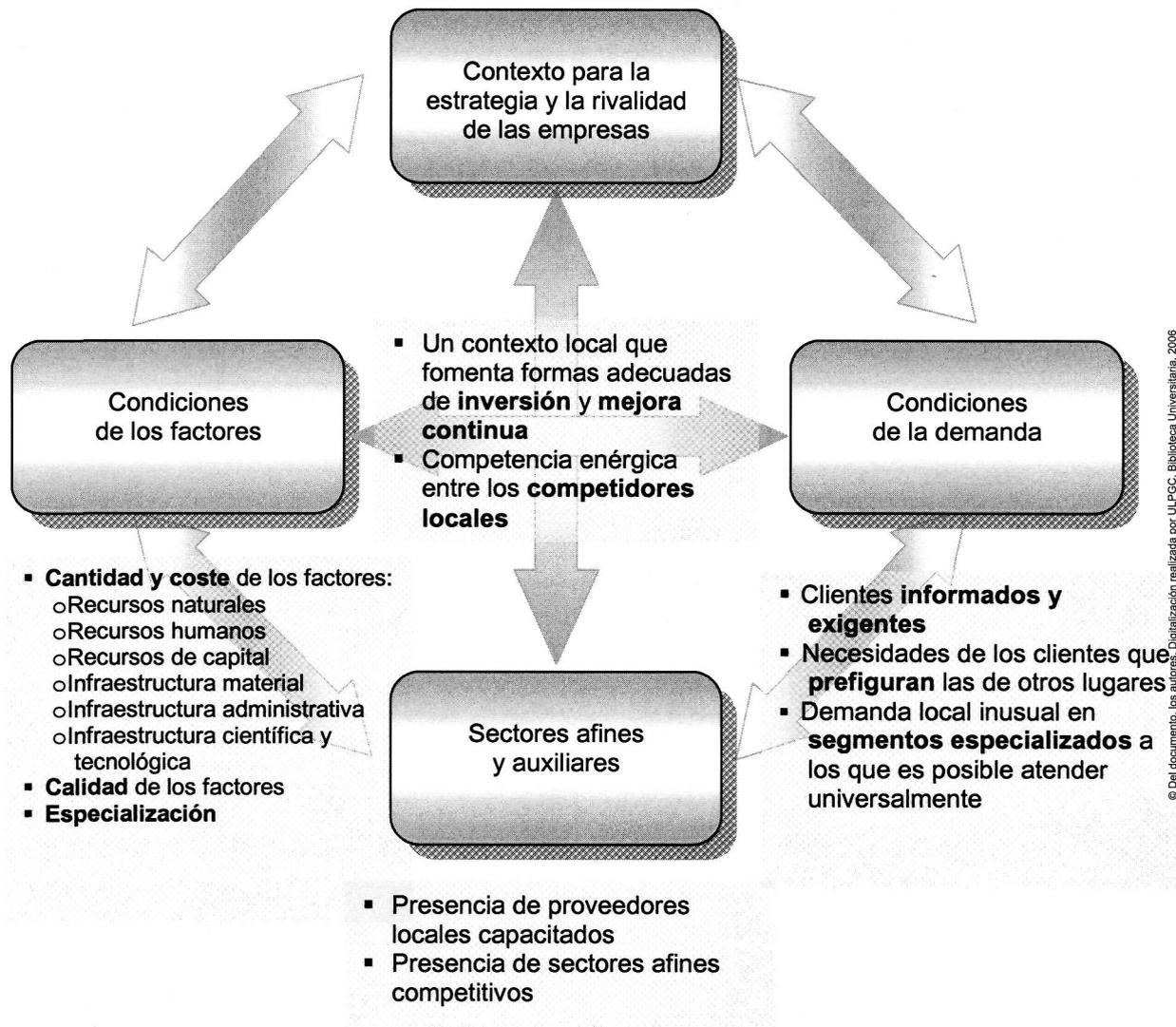


Figura 1.8: Fuentes de la ventaja competitiva de una ubicación

Fuente: Porter (1999: 217).

1.2.2.2. El doble diamante generalizado de Moon, Rugman y Verbeke

Este modelo fue desarrollado por Moon, Rugman y Verbeke (1995) para cubrir algunas debilidades del diamante de Porter e incorporar la actividad multinacional y el gobierno dentro del modelo, y no como parámetros exógenos al mismo⁶. El doble diamante generalizado considera que la competitividad de una nación depende en parte del diamante local y en parte del diamante extranjero con el que se relacionan sus empresas.

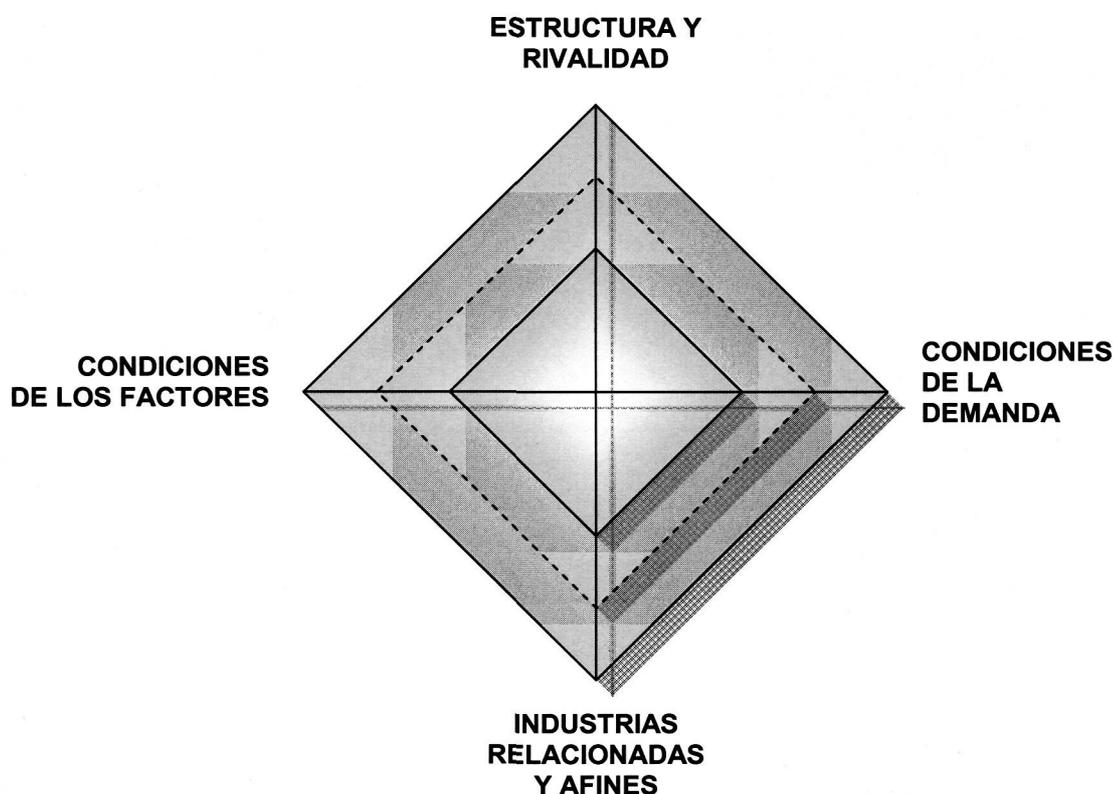


Figura 1.9: El doble diamante generalizado
(Fuente: Moon, Rugman y Verbeke, 1995:98).

⁶ El doble diamante generalizado es un desarrollo del *doble diamante* introducido por Rugman y D´Cruz (1991), cuyo propósito consistía en hacer funcionar el modelo del doble diamante para analizar cualquier economía pequeña.

La figura 1.9 muestra el doble diamante generalizado propuesto por Moon, Rugman y Verbeke (1995), donde la parte externa representa al diamante global y la parte interna al diamante local. El tamaño del diamante global es fijo dentro de un período predecible, pero el tamaño del diamante local varía de acuerdo al tamaño del país y a su competitividad. El diamante de línea de puntos entre los dos diamantes citados, es un diamante internacional que simboliza la competitividad de una nación, determinada tanto por parámetros locales como internacionales. La diferencia entre el diamante internacional y el local representa, por consiguiente, las actividades internacionales o multinacionales. Las multinacionales incluyen tanto la inversión directa extranjera en el país local como la inversión directa en el país extranjero.

1.2.2.3. Los nueve factores de Cho

El modelo de nueve factores fue desarrollado por Cho (1994), como otra extensión del modelo de Porter. En este modelo se hace una división entre factores humanos y factores físicos, y se incluye además el parámetro oportunidad no como algo exógeno al modelo sino interno al mismo. Con relación a los factores del modelo, la primera clasificación representa a los trabajadores, los políticos, los empresarios y los profesionales, y la segunda, a los recursos heredados, la demanda local, las industrias relacionadas y afines, y otros entornos empresariales. Adicionalmente, otra de las nuevas ideas que incorpora a la literatura este modelo es el hecho de que un importante elemento en la competitividad de una nación es la posición competitiva relativa entre países similares en una etapa determinada de desarrollo económico, y no entre todos los países del mundo.

La figura 1.10 muestra un análisis detallado de los nueve factores de la competitividad internacional. En ella se puede ver gráficamente el modelo de nueve factores en el que son cuatro los determinantes físicos de la competitividad internacional, nos referimos a la dotación de recursos, el entorno empresarial, las industrias relacionadas y afines y la demanda local; otros cuatro hacen referencia a los factores humanos, trabajadores, políticos, burócratas, empresarios y directivos profesionales; el noveno factor consiste en los eventos oportunos.

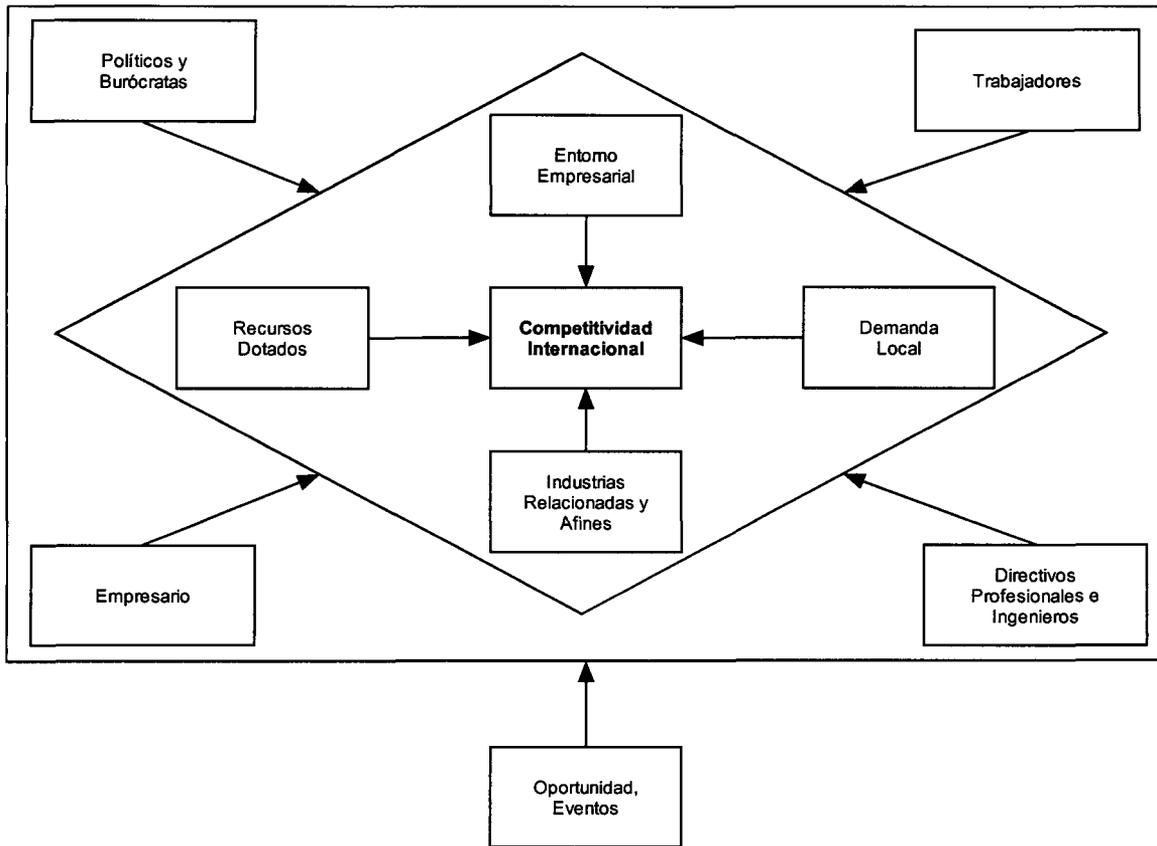


Figura 1.10: El modelo de los nueve factores
Fuente: Cho (1994: 18).

Según Cho, los nueve factores varían en importancia mientras se mueven de una etapa de desarrollo económico bajo a otra más desarrollada. En la primera, los empresarios empiezan a depender menos del gobierno, a invertir a pesar de los riesgos asociados y a buscar economías de escala. Estos se convierten en los recursos humanos que forman la fuente principal de competitividad internacional. El modelo empresarial es generalmente de competencia oligopolística y las empresas tienden a diversificarse en nuevas áreas, resultando en nuevos desarrollos de industrias relacionadas y afines. Y en la etapa de gran desarrollo, las conexiones entre industrias relacionadas horizontal o verticalmente y las de apoyo, se expanden y los bienes y servicios producidos pasan a competir en términos de igualdad con los de países avanzados.

1.3. MODELOS EMPÍRICOS DE LA COMPETITIVIDAD DE LAS NACIONES:

Actualmente, existen dos estudios líderes a escala mundial que analizan y cuantifican los factores que determinan la competitividad de las naciones: el *Global Competitiveness Report* y el *World Competitiveness Yearbook*. El primero lo realiza el *World Economic Forum* (WEF), y el segundo el *International Institute for Management Development* (IMD), ambos tienen su sede en Suiza. Estos dos institutos vienen realizando, desde hace más de una década, estudios empíricos de competitividad de diferentes economías del mundo. Tomando como base estos dos estudios, surge el estudio de la competitividad regional de Canarias en el concierto internacional, realizado por la Universidad de Las Palmas de Gran Canaria.

Todos los anteriores estudios han estado utilizando factores de competitividad similares para analizar y jerarquizar las economías de los países que incluyen en sus informes de competitividad. Los factores de competitividad utilizados e independiente del nivel de desglose que tengan, se engloban en la actividad económica de la región, la eficiencia de la actividad gubernamental y empresarial y la calidad de las infraestructuras. Asimismo, estos estudios definen la competitividad por las condiciones económicas de la región y un entorno que favorezca la competitividad con éxito.

Por otra parte, se argumenta que los factores utilizados por estos estudios, como determinantes de la competitividad internacional se obtienen de la literatura y se basan en análisis empíricos. Así lo confirman Barro y Martín (1995) cuando argumentan que los factores estructurales que se consideran críticos para la competitividad internacional, se basan en los resultados de investigaciones teóricas y empíricas de la teoría moderna del crecimiento económico. Por ejemplo, la importancia de la apertura al comercio internacional y a las finanzas para incrementar la competitividad internacional ha sido apoyada por Sachs y Warner (1995), Dollar (1992), Frankel y Romer (1996). La importancia de la estabilidad política, el tamaño del gobierno y el rol de la educación han sido empíricamente respaldados por Barro (1991). La importancia de la eficiencia de las instituciones financieras ha sido confirmada por King y Levine (1993). El apoyo empírico sobre la importancia de la contribución de la calidad de la

infraestructura para la competitividad internacional y el crecimiento puede ser hallado en Munnell (1992). Evidencia de la importancia de la tecnología, investigación y desarrollo y la difusión tecnológica se puede encontrar en Coe, Helpman y Hoffmaister (1995). La importancia de la calidad de las instituciones políticas y legales para la competitividad internacional y el desarrollo sostenible puede ser encontrada en Knack y Keefer (1995). Evidencia sobre la importancia de la flexibilidad del mercado laboral se ilustra en Abowd, Kramarz, Lemieux y Mangolis (1997) y en Coe y Snower (1997). Consecuentemente, estamos analizando factores que han sido cuidadosamente elegidos y contrastados con análisis empíricos, llevados a cabo por una gran diversidad de estudios realizados por una amplia relación de académicos e investigadores. A continuación, explicaremos con más detalle los fundamentos teóricos y metodológicos en los que estos estudios se basan para analizar y evaluar la competitividad de las naciones.

1.3.1.- EL MODELO DEL *WORLD ECONOMIC FORUM* (WEF)

El informe WEF de competitividad global es parte de una serie de informes de competitividad que ha estado publicando el *World Economic Forum* desde 1979⁷. A principios de 1996, su colaboración con *Harvard University* le permitió analizar estadísticamente los fundamentos microeconómicos de la riqueza de un amplio número de países, basándose en las condiciones que apoyan un nivel alto de productividad, medido por el Producto Interior Bruto (PIB) *per capita*. Los fundamentos microeconómicos de desarrollo económico a que se refiere el estudio, surgen de las estrategias de las empresas, sus prácticas operativas, así como de los inputs empresariales, las infraestructuras, las instituciones y las políticas que constituyen el entorno en el que compiten las empresas de una nación.

Como explican Sachs y Warner (2000), en los últimos años, las teorías macroeconómicas que pretendían explicar fluctuaciones económicas a corto plazo, están siendo seriamente cuestionadas, sobre todo porque no llegan a explicar las crisis que han estado ocurriendo en mercados emergentes o países en desarrollo. De

⁷ Para más información se puede contactar con la página web del WEF: <http://www.weforum.org>

hecho, cuando han existido fuertes fluctuaciones económicas en algunas partes del mundo (eg., México 1994-1995; Asia Oriental 1997-1998), éstas han sorprendido a economistas y pronosticadores, que intentan predecir ciclos económicos futuros. Adicionalmente, el estudio del WEF parte de la base de que, aunque existe un mayor entendimiento sobre los aspectos macroeconómicos de la competitividad y del desarrollo económico, el hecho de tener un contexto político estable y políticas macroeconómicas adecuadas no es suficiente para asegurar la prosperidad de una nación. Además, los aspectos macroeconómicos por sí solos no representan la respuesta a los países más pobres, tan importantes o más son los fundamentos microeconómicos del desarrollo económico, que muchas veces suponen la causa o el resultado de la competitividad de las naciones analizadas. Asimismo, los colaboradores del estudio explican que las mejoras en las condiciones microeconómicas no son automáticas, sino que se ven afectadas por las acciones del sector público y del sector privado. Así, la disposición de recursos humanos poco capacitados dificultaría la labor de aquellas empresas que se enfrentan a una demanda exigente. Las políticas macroeconómicas que fomentan un incremento en las tasas de inversión de capital no se traducirán en un incremento en la productividad si las formas de inversión no son apropiadas, si no existen capacidades o industrias de apoyo para hacer que estas inversiones sean eficientes, o si no existen presiones competitivas que motiven a las empresas a innovar. Asimismo, altas tasas de inversión pública en capital humano no darán resultados si las circunstancias microeconómicas no fomentan la demanda de personal cualificado en las empresas. De esta forma, este modelo nos transmite que para que las políticas a nivel macro se traduzcan en una economía cada vez más productiva, deben existir al mismo tiempo ciertas condiciones microeconómicas que lo permitan.

La figura 1.11 muestra los determinantes de la productividad y del crecimiento de la productividad. El modelo del WEF se basa en la idea de que el estándar de vida viene determinado por la productividad de la economía de una nación, medida por el valor de los bienes y servicios producidos por cada unidad de sus recursos naturales, capitales y humanos. Por ello, resulta fundamental para el desarrollo económico crear las condiciones para que ocurra un crecimiento rápido y sostenido en términos de productividad.

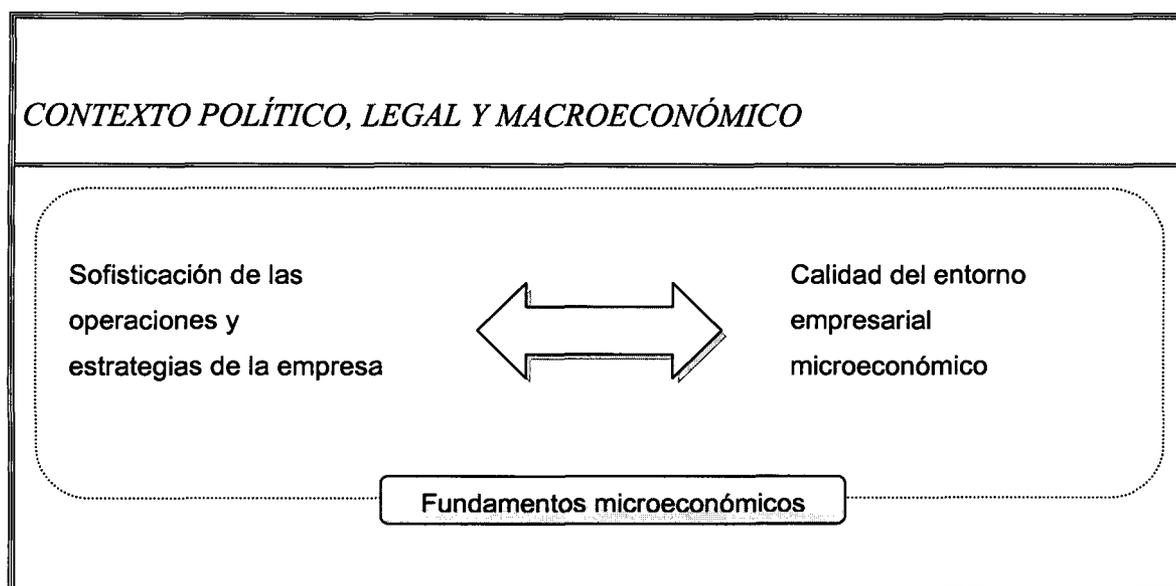


Figura 1.11: Determinantes de la productividad y del crecimiento de la productividad
(Fuente: Porter 1999)

Los fundamentos microeconómicos de la productividad se apoyan en dos áreas interrelacionadas: (1) la sofisticación con la que compiten las empresas y (2) la calidad del entorno empresarial microeconómico. La sofisticación de las empresas viene dada por la calidad del entorno empresarial nacional, que puede ser entendido en términos de cuatro influencias interrelacionadas que conforman el diamante de Porter –condiciones de los factores, condiciones de la demanda, estrategia y rivalidad de las empresas e industrias relacionadas y afines–, que ya hemos presentado anteriormente en nuestra investigación, concretamente en la figura 1.5 con el título “Los determinantes de la ventaja nacional”, del capítulo 1 de nuestro trabajo. El desarrollo económico de las naciones es un proceso de mejora continua en el que el entorno empresarial de una nación evoluciona para apoyar e impulsar formas de competir más sofisticadas y productivas. Asimismo, el nivel de desarrollo de las naciones determinará el tipo de desafío al que se pueda enfrentar.

En su análisis de la competitividad y el desarrollo económico de muchos países del mundo, el WEF realiza dos índices diferentes, uno de Competitividad Actual y otro del

Crecimiento de la Competitividad. El primero de ellos, pretende identificar los factores que fundamentan una alta competitividad actual en las naciones y por lo tanto una actividad económica productiva, medida por el nivel del PIB *per capita*. Los factores evaluados explicarían por qué algunos países pueden mantener un nivel más alto de competitividad que otros. El segundo, supone una revisión del índice de competitividad de los informes que este mismo estudio ha estado realizando en colaboración con la Universidad de Harvard desde 1996. Con el índice del crecimiento de la competitividad, se pretende medir los factores que contribuyen al crecimiento futuro de una economía evaluado por una tasa de cambio del PIB *per capita*. Estos factores permitirían explicar por qué algunos países están mejorando sus niveles de competitividad mucho más rápido que otros. Claramente, las dos dimensiones de competitividad mencionadas están muy relacionadas, debido a que ambas se centran en las condiciones que apoyan la productividad, aunque con diferentes horizontes de tiempo. Adicionalmente, comprender los determinantes del progreso de una economía en un momento determinado, es tan importante como comprender los factores determinantes del crecimiento que permite incrementar su nivel de vida. Como consecuencia de ello, los dos índices proporcionan una imagen mucho más reveladora de la competitividad de las naciones, que si se evaluara sólo uno de ellos.

Las teorías del crecimiento económico distinguen entre aquellas circunstancias que contribuyen al nivel de renta *per capita* y aquellas que contribuyen a cambiar la renta *per capita* o crecimiento. En las teorías más sencillas, el nivel de renta *per capita* (y) depende de la cantidad de capital por persona (k), a veces denominado como la *intensidad de capital* de la economía. Aunque en las investigaciones empíricas el capital ha sido a menudo medido de forma restringida, el *capital* debería ser tratado más ampliamente para incluir al *capital* físico y al *capital* humano. El capital humano abarca no sólo los niveles de educación sino también la experiencia laboral de la mano de obra y el know-how directivo.

En los modelos más sencillos, el producto nacional bruto *per capita* es proporcional a la cantidad de capital por persona:

$$y = Ak, \quad (1)$$

donde (A) representa el nivel de tecnología resumido como un número simple que mide la productividad media de una unidad de capital. El nivel de renta, entonces, es determinado por el *stock* de capital y por el nivel de tecnología.

En los modelos de crecimiento, frecuentemente se asume que se ahorra una proporción fija de renta:

$$\Delta k = s y,$$

donde (y) es el PIB o renta, (s) es la proporción de renta ahorrada, y (Δk) es el cambio en el *stock* de capital.

En el caso más simple en el que (A) sea fijo, la tasa proporcional de crecimiento de la economía ($\Delta k / y$) es igual a ($\Delta k / k$), que a su vez es igual a (s) x (A). El crecimiento de la renta en un mundo de tecnología fija está determinado por la tasa de ahorro multiplicado por el “nivel de tecnología” (A). Por supuesto, (A) no es fijo en las economías actuales. De aquí que el crecimiento económico tenga dos componentes importantes: el cambio tecnológico y el incremento de capital:

$$\Delta y / y = \Delta A / A + s A$$

crecimiento = cambio tecnológico + incremento de capital

Está claro que de este marco conceptual podemos construir dos índices distintivos o *ranking*, uno para explicar el nivel de renta en la economía, y otro para explicar la tasa de crecimiento de renta en la economía. El índice de competitividad actual medirá el capital (k) y el nivel actual de tecnología (A). Mientras que el índice de crecimiento de

la competitividad medirá la tasa de ahorro (s), el nivel actual de tecnología (A) y la tasa de mejora en tecnología ($\Delta A / A$). Sin embargo, ambos índices tendrían algún tipo de solapamiento debido a que dependen de A , pero también diferirán en el sentido de que el índice de competitividad actual también depende de (k), mientras que el índice de crecimiento depende de (s) y de ($\Delta A / A$). Como consecuencia, nos podemos encontrar con cuatro tipos de economías: a) rica y con un rápido crecimiento; b) rica y bajo crecimiento; c) pobre y de crecimiento rápido; d) pobre y de crecimiento lento.

En las economías actuales, las variables (k), (A), (s), y ($\Delta A / A$) son multidimensionales y muy complejas. El *stock* de capital de una economía incluye no simplemente el capital fijo acumulado en maquinaria, estructuras e infraestructuras físicas (carreteras, puertos y telecomunicaciones), sino también el nivel de educación, las capacidades y actitudes de la fuerza de trabajo y el talento directivo. También, como parte del *stock* de "capital" de una economía se encuentran, el conjunto de instituciones legales y prácticas reguladoras que gobiernan los negocios. El capital social (niveles de confianza, buenas costumbres y la presencia de redes) también contribuye a la calidad del *stock* de capital global. El nivel de tecnología en una economía es igualmente multidimensional. Abarca no sólo el conocimiento tecnológico que se encuentra en las instituciones técnicas y científicas de una nación, sino también en la tecnología enraizada en las empresas. La tecnología está incorporada en cualquier actividad que realiza una empresa así como en las estrategias que las empresas usan para competir.

Las condiciones que conducen a un rápido crecimiento económico incluyen no sólo las tasas de ahorro o de inversión agregadas en una economía sino también la combinación de instituciones públicas y privadas que apoyan la innovación (tales como laboratorios nacionales, centros académicos, e institutos de investigación del sector privado); la difusión de ideas entre los sectores; y la incorporación de ideas de economías extranjeras en la economía doméstica. Por ejemplo, un sistema impositivo favorable a la creación de nuevas empresas y las alianzas estratégicas internacionales también juegan un papel importante.

En la práctica, algunas de las mismas instituciones, regulaciones, atributos y prácticas afectan a ambos índices, aunque a veces a través de diferentes mecanismos. Por

ejemplo, la intensidad de la rivalidad en una economía y la sofisticación de los clientes locales conducen a la productividad actual pero también fomentan el crecimiento de la productividad. La presencia de suministrados locales capaces beneficia la eficiencia actual, pero también apoya la innovación. En la práctica entonces, las influencias sobre el índice de competitividad actual y el índice de crecimiento de la competitividad, serán diferentes aunque se solapen. Destacar también que el nivel de PIB *per capita* puede llegar a no estar alineado con la competitividad actual si un entorno de crecimiento favorable (ejemplo altos ahorros o incorporaciones de capital) enmascaran la debilidad con relación a la competitividad actual. O el progreso de una economía a lo largo de diferentes dimensiones de la competitividad actual podría ser poco uniforme, con algunas áreas llegando a convertirse en restricciones que, en última estancia, llegan a ser vinculantes. Similarmente, algunos factores determinantes de crecimiento en un país pueden ser más favorables que otros, conllevando, por ejemplo, a fuertes inversiones de capital físico, sin una adecuada mejora en la tecnología, que reduce los beneficios de las inversiones. Cuando distinguimos entre competitividad actual y crecimiento de la competitividad, podemos pintar un cuadro mucho más rico de las circunstancias y perspectivas económicas de una nación.

Es muy posible tener un país rico que es probable que crezca lentamente en el futuro (*stock* de capital y de tecnología alto, pero baja predisposición al ahorro y a la innovación). Es también posible tener un país pobre que es probable que crezca rápidamente en el futuro (bajo *stock* de capital y de tecnología actual, pero alta predisposición a ahorrar y a adoptar nuevas tecnologías desde el extranjero).

1.3.2.- EL MODELO DEL *INTERNATIONAL INSTITUTE FOR MANAGEMENT DEVELOPMENT* (IMD)

Mientras que el estudio de la competitividad internacional del WEF basa su teoría en el modelo del diamante de Porter, el estudio del IMD aplica su propia teoría que puede denominarse como *las cuatro fuerzas fundamentales*. Este modelo afirma que los países gestionan sus entornos de acuerdo a las cuatro fuerzas fundamentales que conforman el entorno competitivo de un país. Estas dimensiones son a menudo el

resultado de tradición, historia o sistemas de valores y están profundamente enraizadas en el “*modus operandi*” de un país que, en la mayoría de los casos, no son claramente definidas o explicadas.



Figura 1.12: El cubo competitivo del IMD

(Fuente: IMD (2001: 45).

La figura 1.12 ilustra el cubo competitivo del IMD, donde se puede comprobar las cuatro fuerzas fundamentales que conforman el entorno competitivo de un país y que integran una teoría general que, a su vez describe las relaciones entre los cuatro ejes. Las dimensiones ilustradas en la figura y explicadas a continuación son las siguientes:

- a) Atractividad *versus* Agresividad
- b) Proximidad *versus* Globalidad
- c) Activo *versus* Procesos
- d) Toma de Riesgos Individual *versus* Cohesión Social

a) Atractividad *versus* Agresividad

La primera dimensión, atractividad *versus* agresividad, se basa en que las naciones varían en la forma en que gestionan sus relaciones con la comunidad empresarial mundial. Tradicionalmente, la competitividad estaba relacionada con la agresividad internacional de países, es decir, con las exportaciones y con la inversión directa en el extranjero. Más recientemente, algunas naciones gestionan su competitividad siendo atractivas para localizar empresas. La agresividad genera renta en el país local, pero no trabajo. Por otro lado, la atractividad crea trabajos en los países de inversión directa pero pueden escasear en renta debido a la inversión en incentivos. Esto significa que, incluso las naciones ricas, no deben ignorar la importancia de la atractividad, especialmente por su impacto en el empleo. Por lo tanto, si las naciones quieren competir, deben considerar tanto la agresividad como la atractividad.

b) Proximidad *versus* Globalidad

En segundo lugar, se debe considerar que el sistema económico no es generalmente homogéneo. En la mayoría de los casos, las naciones deben tratar con dos tipos de economías coexistentes: la economía de proximidad y la economía de globalidad. La primera conlleva actividades tradicionales: artesanía y servicios personales, tales como doctores y profesores; actividades administrativas como el gobierno y la justicia; y actividades de apoyo al consumidor como servicio postventa y personalización de la clientela. Esta economía ofrece valor añadido por estar localizada próxima al usuario final. Se trata de una economía generalmente proteccionista y cara. La economía de globalidad, por otro lado, es una economía competitiva donde se incluyen empresas con operaciones internacionales, y se asume que la producción no necesita necesariamente estar cerca del usuario final. Este tipo de economía se beneficia de las ventajas comparativas de los mercados mundiales, especialmente con respecto a los costes operativos. La proporción entre estas dos economías para la prosperidad nacional varía con respecto al tamaño y el desarrollo económico de un país. Los países pequeños son más dependientes de su economía de proximidad y los países grandes de su economía de globalidad.

c) Activos versus Procesos

La tercera dimensión, activos *versus* procesos, indica que aquellas naciones que son ricas en activos –tierra, personas, y recursos humanos– no son necesariamente competitivas. Sin embargo, las naciones pobres en activos pero que se han apoyado esencialmente en procesos de transformación son mucho más competitivas. Adicionalmente, el estudio explica que los activos heredados no son necesariamente los recursos naturales. La infraestructura, la educación y las capacidades pueden ser consideradas también naturales ya que han sido acumulados en el pasado por otras generaciones. Esta dimensión examina por lo tanto, hasta donde un país depende de sus recursos naturales para obtener sus ingresos en vez de desarrollar procesos de transformación que apoyen la creación de valor. Como explica Garelli (1998), la distinción entre activos y procesos también realiza la diferencia entre conceptos de riqueza (heredados) y de competitividad (creados).

d) Toma de Riesgo Individual versus Cohesión Social

La última dimensión que conforma el entorno competitivo de un país es la distinción entre un sistema que promueve el riesgo individual y otro que preserva la cohesión social. El modelo anglo-sajón se caracteriza por la toma de riesgo, la desregulación, privatización y responsabilidad del individuo a través de un acercamiento mínimo del sistema del bienestar. Por el contrario, el sistema europeo continental depende en gran parte del consenso social, y tiene un enfoque más igualitario hacia las responsabilidades y un amplio sistema de bienestar. Sin embargo, como comenta Garelli (1999 y 2000), aunque ambos modelos hayan competido por muchos años, es el modelo anglo-sajón el que prevalece. La legislación de la Unión Europea ha tendido más hacia la desregulación y la privatización.

Resulta posible combinar algunas de las dimensiones analizadas con el fin de establecer modelos de competitividad. Así, nos podemos encontrar con la combinación de las diferentes dimensiones, por ejemplo, *Globalidad y Toma de Riesgo*. Según el IMD, la economía global está influida principalmente por un enfoque de toma de riesgo individual, o también llamado modelo anglo-sajón. Tanto Estados Unidos como el

Reino Unido han sido pioneros en tener negocios internacionales, por lo que no sorprende sus influencias en el modelo global sobre la idea de competitividad basada en la desregulación, la privatización y el espíritu empresarial. *Proximidad y Cohesión Social* representan la segunda combinación de las dimensiones del modelo. La cohesión social siempre ha estado asociada con una economía de proximidad donde las consecuencias sociales de la competitividad importan y es en el modelo europeo donde destaca este tipo de enfoques. Adicionalmente, el IMD propone que un factor importante y determinante de la competitividad de las naciones es el sistema de valores de un país. Este sistema de valores atraviesa cuatro fases bien diferenciadas:

1. Trabajo duro: las personas se dedican por completo a los objetivos corporativos del país y trabajan un gran número de horas (por ejemplo, en Corea).
2. Riqueza: aunque está arraigado el valor del trabajo duro, las personas prestan una mayor atención a incrementar sus propios ingresos (por ejemplo, en Singapur).
2. Participación social: las personas están menos interesadas en trabajar duro y más implicadas en la formación de su sociedad (por ejemplo, en Estados Unidos y en Europa a finales de los 60).
3. Auto-realización: las personas están más interesadas en desarrollar sus propias vidas, en vez de perseguir un cambio social (por ejemplo, en Estados Unidos y en Europa en la actualidad).

La figura 1.13 muestra una evolución natural de los valores, desde una perspectiva colectiva a una perspectiva individual. Este proceso no es reversible pero si es posible gestionarlo. Japón, por ejemplo, está en la transición de valores colectivos a valores individuales, lo que implica una reforma con profundidad de sus sistemas políticos, sociales y económicos. El argumento se basa en que los países van cambiando sus sistemas de valores a medida que van avanzando en el desarrollo económico. Así, los países asiáticos se encuentran en la actualidad en el sistema de valores que dominó a Europa y Estados Unidos en el siglo XIX, consistente en trabajo duro, disciplina, lealtad y ahorro. En cambio, los europeos y los americanos se encuentran en la actualidad, inmersos en un sistema de valores individuales que escenifica el logro personal.

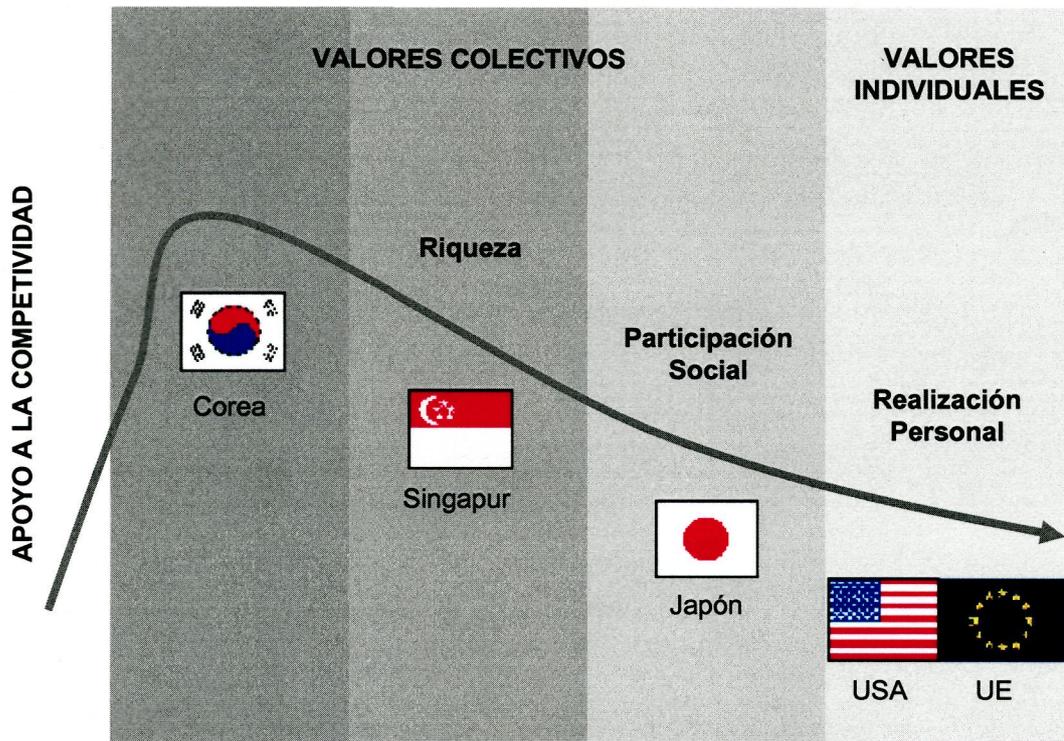


Figura 1.13: Evolución del sistema de valores

Fuente: IMD (2001:47).

1.3.3.- EL MODELO DE LA UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA

El estudio de la competitividad regional de la Universidad de Las Palmas “La competitividad internacional de Canarias” se basó en una adaptación de los dos modelos anteriores - WEF e IMD -, para permitir realizar un análisis comparativo de la competitividad internacional de Canarias. Este modelo, comparte la definición de competitividad de IMD y WEF y define, por lo tanto, a la misma basándose en el medio y largo plazo. En tal sentido, la competitividad de una región se basa en la capacidad que tenga el entorno regional para crear un crecimiento económico sostenible que impulse la competitividad de sus empresas.

Asimismo, en este estudio el término región se puede considerar tanto en el nivel micro como en el nivel macro, enlazando la idea de las economías actuales basadas en el binomio local-global. La idea es que la necesidad de hacer una comparativa con otras regiones rompe las barreras de dimensión geográfica, demográfica o económica, más de acorde con la economía tradicional, donde los recursos naturales por sí solos constituían la clave del éxito nacional. En vez de ello, y con el fin de evaluar, en la economía vigente, la ventaja competitiva regional, este modelo realiza una comparativa de las capacidades de la región para alcanzar un crecimiento económico sostenible. Esta comparativa se realiza con independencia del tamaño de las regiones, los sectores económicos clave o la riqueza medida por el PIB.

De cualquier forma, la región evaluada en este modelo presenta rasgos distintivos que le permiten perfectamente ser comparada con otras regiones de nivel micro como ella, o de nivel macro como una nación o estado. En este sentido, y con relación a su naturaleza, Canarias es una economía perteneciente al Estado español como una de sus Comunidades Autónomas y a la Unión Europea, como una de sus regiones ultraperiféricas.⁸ En el ámbito demográfico, Canarias tiene una densidad de población de 245 habitantes por kilómetro cuadrado, propia de países o regiones altamente

⁸ Canarias se convirtió oficialmente en una Comunidad Autónoma española en agosto de 1982. Y desde que entró en vigor el artículo 299 (2), del Tratado de Ámsterdam, en mayo de 1999, Canarias es reconocida oficialmente como una de las siete regiones ultraperiféricas de la Unión Europea.

industrializadas, o de servicios internacionales especializados, como Suiza, Luxemburgo y Holanda. Y en el ámbito económico se trata de una región cuyo Producto Interior Bruto por habitante en el año 2000 (13.566 \$) la situó en el posicionamiento 25 del ranking de 50 territorios del mundo, clasificándola así como una región perteneciente al mundo de la “sociedad opulenta” y frente a países aún en vías de desarrollo (e.g., India, cuyo PIB por habitante asciende a no más de 444\$).

Por otro lado, la utilización de los estudios de WEF y de IMD como base para realizar el análisis y cuantificación de la competitividad internacional de Canarias, permitió que el modelo de Canarias presentase un carácter amplio, incluyendo más de 300 variables de medición. Este grupo de variables estaba integrado por una serie de datos descriptivos del entorno regional y otras variables de percepción de una selección de empresarios locales. Además, la integración de las variables de los modelos WEF e IMD resultó en un estudio mucho más completo y detallado de competitividad que los dos anteriores. Igualmente, el modelo de la Universidad de Las Palmas permitió hacer una comparativa internacional y posicionar a la región de Canarias en un ranking de competitividad mundial, traspasando así las fronteras europeas a las que por su naturaleza pertenece.

Este estudio de la Universidad de Las Palmas se caracteriza por su carácter integrador, donde se incluyen variables de naturaleza secundaria pertenecientes a la economía local, aunque también integra todas las otras variables de la misma naturaleza pertenecientes a las economías nacionales con las que se compara a nivel internacional. Asimismo, el modelo incluye variables de naturaleza primaria de las economías nacionales con las que igualmente se compara y variables de la misma naturaleza recogidas a nivel local. Como particularidad de este modelo, que no ocurrió con los estudios del WEF y de IMD, las variables de naturaleza primaria fueron recogidas directamente por la investigadora del estudio mediante una encuesta personal, totalmente estructurada. En este sentido, la calidad de la información obtenida fue más precisa y en su respuesta los encuestados se posicionaron en un entorno internacional, lo que enriquece aún más la calidad del estudio.

Así, y en relación a las variables de naturaleza primaria procedentes de las percepciones de expertos cualificados se utilizó un cuestionario formado por 133

ítemes, donde se analizan distintas áreas que en las economías actuales resultan necesarias para identificar los factores determinantes del nivel de competitividad de cualquier territorio. A continuación se explican las diferentes áreas analizadas para identificar los factores determinantes de la competitividad internacional de Canarias.

- **Infraestructura:** incluye cuestiones relativas a las infraestructuras de distribución de bienes, la red de carreteras, el transporte aéreo y marítimo, y las telecomunicaciones.
- **Recursos humanos:** analiza aspectos tales como la cualificación del capital humano, el sistema educativo, el sistema sanitario, la problemática social (discriminación social por raza o sexo, etc.), la pro actividad de la población activa y el poder de los sindicatos.

Ciencia y Tecnología: integra aspectos relativos a la existencia y reputación de los centros de investigación científica, la disponibilidad de ingenieros cualificados y recursos económicos, el nivel de transferencia tecnológica y la importancia de la tecnología en la sociedad canaria.

- **Finanzas:** analiza áreas relacionadas con la dificultad en la obtención de créditos, el grado de utilización de recursos propios por parte de las empresas canarias para financiar sus inversiones, el grado de utilización de las empresas de los mercados de capitales extranjeros y la existencia de profesionales especializados en el área de finanzas.
- **Competencia:** centrada básicamente en los aspectos relacionados con el entorno competitivo de la región, los factores que dificultan la creación de nuevos negocios, la sofisticación de los consumidores y las características de los proveedores con relación a eficiencia y competitividad.
- **Gestión Empresarial:** determina el nivel de calidad existente en la gestión empresarial ejercida por nuestros empresarios, analizando aspectos relacionados con el activo principal sobre el que se sustenta la ventaja competitiva de las empresas, al desarrollo tecnológico, a la innovación de productos, a los acuerdos

comerciales con empresas extranjeras, al nivel de internacionalización. Asimismo, se analizan las características de la política de recursos humanos implantada, los órganos de poder, la reputación del empresariado canario en nuestra sociedad, a la disponibilidad, formación y experiencia de los directivos, el grado de identificación de los empleados con los objetivos de las empresas y el nivel de responsabilidad social de las mismas en sus actuaciones.

- **Internacionalización:** Integradas en esta área se encontraban básicamente la actitud de las empresas hacia la globalización, el atractivo de los incentivos a la inversión, el apoyo de las instituciones públicas a la exportación, la percepción sobre la existencia de barreras de entrada y la imagen de Canarias en el extranjero.
- **Gobierno Regional e Instituciones:** la evaluación de aspectos tales como el nivel de aceptación de los presupuestos públicos, la idoneidad de las subvenciones públicas, los factores y agentes tenidos en cuenta por el gobierno para formular sus políticas, el nivel de profesionalidad del estamento funcional, el fraude fiscal, la independencia política del órgano judicial, las características del sistema judicial, la eficacia del sistema de seguridad y la eficacia del gobierno atendiendo a diferentes criterios.
- **Economía Regional:** se incluye aspectos tales como la distribución de la renta, el nivel de desarrollo de la economía regional y la economía sumergida.

Por otro lado, la utilización de los modelos WEF e IMD permitió realizar un estudio detallado de sus metodologías, identificándose errores, confirmados posteriormente por las fuentes, así como permitir una validación más completa de los datos utilizados a nivel local, a través de la verificación de las fuentes utilizadas y un equipo de investigación. Asimismo, este estudio tiene carácter innovador y único al no existir estudio previo de esta índole, a nivel local o europeo, lo que ha provocado el interés de investigadores a nivel internacional, ya que permite la conexión de la dimensión micro y macro. Por otro lado, a nivel local tampoco se tiene constancia de un estudio anterior similar, con una comparativa internacional, y utilizando más de 300 variables y

50 países para la medición de la competitividad de Canarias.

La figura 1.14 representa los objetivos y análisis que integran el modelo de la competitividad internacional de Canarias. En la misma, se presentan las fuentes utilizadas tanto de información secundaria como primaria y se enumeran los objetivos propuestos en el estudio.

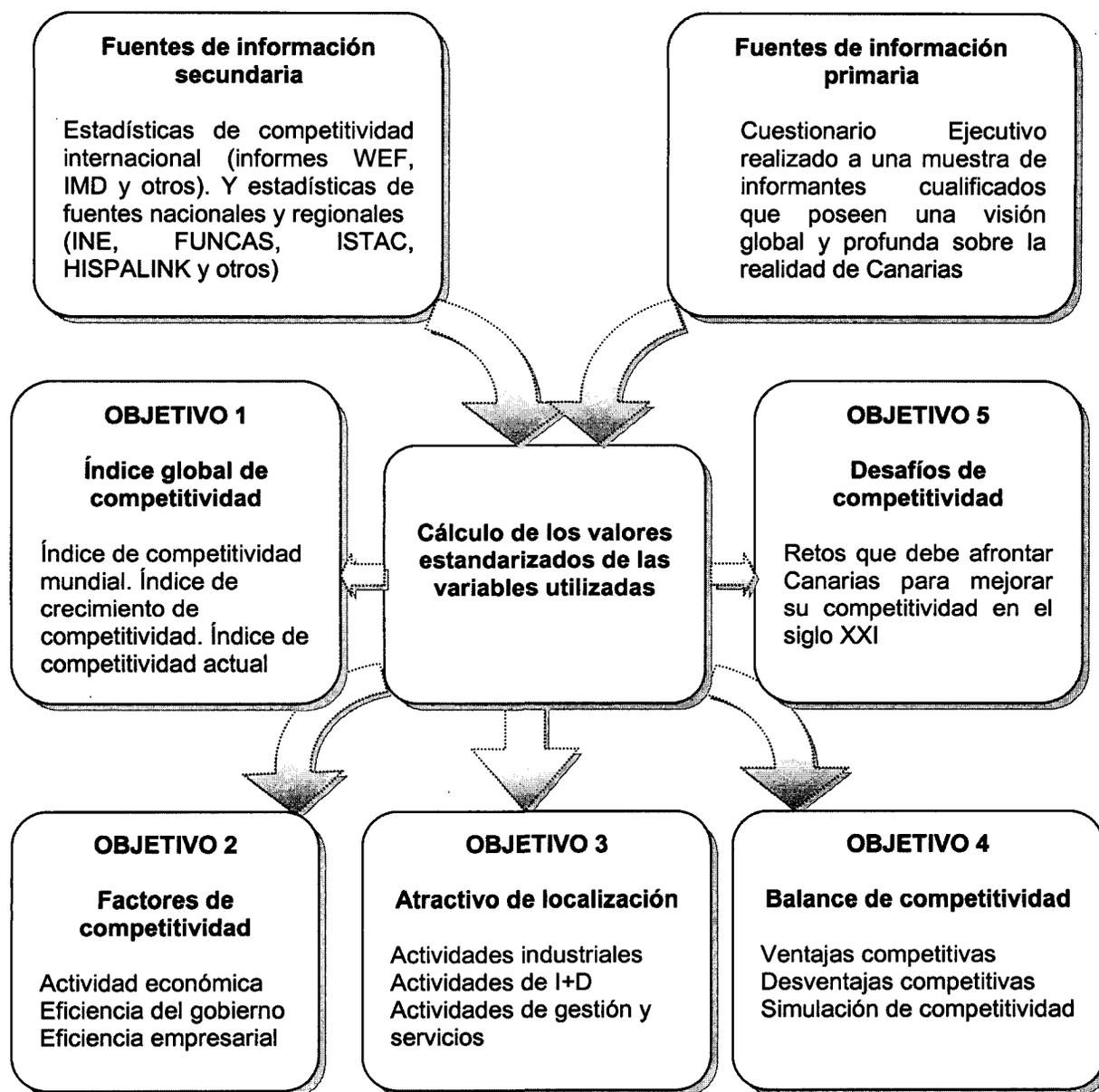


Figura 1.14: Fuentes de información, objetivos y análisis de resultados

(Fuente: Ramos 2002).

Como se desprende de la figura, cinco han sido los objetivos planteados para realizar el estudio. El *objetivo 1* se basó principalmente en determinar el índice de competitividad global de Canarias en el concierto internacional y analizar su posicionamiento competitivo con respecto a otros países, así como la relación existente entre el PIB *per capita* y los factores determinantes de la competitividad. Por otro lado, el *objetivo 2* se centró en identificar los factores y subfactores determinantes de la competitividad de Canarias relacionados con la actividad económica, la eficiencia del gobierno, la eficiencia empresarial y la calidad de las infraestructuras, así como su estructura de competitividad en comparación con la nacional y la del líder de competitividad mundial.

Asimismo, con el *objetivo 3* se analizó el balance de competitividad de España en general y de Canarias en particular identificando para ello las ventajas y desventajas competitivas que presentan con respecto a otros territorios competidores y, adicionalmente, simular el índice de competitividad admitiendo “qué ocurriría si” el valor de sus factores más débiles fuese sustituido por el valor medio respectivo de todos los países. Igualmente, con el *objetivo 4* se pretendió determinar el atractivo que presenta Canarias para las empresas foráneas para la localización de las actividades industriales, las actividades de investigación y desarrollo y las actividades de servicios.

Finalmente, en el *objetivo 5* del estudio se identificaron los principales desafíos competitivos mundiales y los que tiene que afrontar Canarias en los próximos años con el fin de mejorar su competitividad en el concierto internacional.

El estudio de La Competitividad Internacional de Canarias ha sido la fuente de inspiración y será la base de análisis del sistema avanzado de conocimiento propuesto en el capítulo 3 de la tesis. Este nuevo sistema avanzado de conocimiento suprimiría las limitaciones que los modelos de competitividad presentan en la actualidad y que están relacionadas con ser documentos físicos, unipersonales, compactos e inflexibles, tanto a la hora de intentar conectar las fuentes utilizadas, como al realizar análisis de inferencia de conocimiento automatizado. Ambos tópicos serán estudiados en el capítulo 3 de la tesis.

Asimismo, la integración de las tecnologías de la información en la gestión del conocimiento sobre competitividad cubre una laguna importante en los trabajos de investigación que hasta ahora se han realizado y que están relacionados con la gestión del conocimiento. Es decir, se ha estado infravalorando la importancia de la tecnología cuando la misma, representa en realidad el principal ingrediente en la implantación con éxito de esa gestión del conocimiento.

Iniciamos pues, el capítulo 2 de la tesis, realizando una revisión de las tecnologías de la información aplicables a la gestión del conocimiento.



Capítulo 2

Tecnologías aplicadas a la Gestión del Conocimiento

“La pregunta, sobre si un ordenador puede pensar no es más interesante que la pregunta sobre si un submarino puede nadar”

(Adaptado de E.W. Dijkstra).

2.1. TECNOLOGÍAS DE LA INFORMACIÓN Y LA GESTIÓN DEL CONOCIMIENTO

La historia del ser humano ha estado muy ligada a la historia de sus máquinas (Weizenbaum, 1976). Las máquinas han permitido al hombre transformar su entorno físico, construir ciudades, abrir canales, cultivar tierras. Y, como argumenta Weizenbaum, las máquinas del hombre han determinado de forma considerable su capacidad de entendimiento del mundo que le rodea y de sí mismo. En este sentido, la representación subjetiva de su entorno, mediante modelos simbólicos (basados en símbolos), y la inferencia que realiza sobre estos modelos han sido aspectos cruciales para la historia del ser humano. Esto ha permitido, como consecuencia, no sólo cambiar su realidad sino también sus campos, sus ciudades, sus mercados, sus industrias, sus empresas etc... Además, le han aproximado al conocimiento, no sólo de sí mismo sino también de su contexto y de sus naciones. Por consiguiente, podemos afirmar que las máquinas se han convertido en agentes que facilitan la propia evolución humana.

Asimismo, estas máquinas han evolucionado de ser sólo físicas a convertirse en abstractas y virtuales (software), avanzado de la gestión exclusiva de objetos físicos a manipulación de símbolos abstractos, datos, información y conocimiento. Y, donde antes eran modeladas sólo por el ser humano ahora pueden ser autoconfigurables. Las nuevas máquinas se han convertido así en una pieza clave de las economías actuales, basadas en las Tecnologías de la Información y de las Comunicaciones (TICs), y cubren un amplio abanico que entrelazan el quehacer diario de los ciudadanos, y una laguna importante existente en la gestión del conocimiento.

La posibilidad de explicitar la percepción y el conocimiento acerca del entorno ha sido posible gracias a la representación simbólica a través de números, letras, palabras, sonidos, imágenes etc., en definitiva, de datos, y ha facilitado a su vez, la transferencia de dicha percepción y conocimiento a otros miembros de la sociedad, incluso de forma atemporal. Para la Real Academia Española (RAE), *un dato se define*

como una representación de una información de manera adecuada para su tratamiento por un ordenador. En este sentido, en ningún caso, es necesario que se facilite la interpretación de ese dato, es decir, implica simplemente la representación simbólica de un hecho que está desprovisto de significado para el receptor. Sólo cuando el dato tiene significado para el receptor y facilita la toma de decisiones de este último, entonces hablamos de *información*.

Existe otro término muy importante que se utiliza en las investigaciones contemporáneas sobre las TICs, para identificar los datos que hacen referencia a otros datos, llamados *metadatos*, es decir, datos acerca de datos. Estos *metadatos* describen el contenido, la calidad, la condición y otras características de esos otros datos de nivel inferior, y representan, a su vez, un elemento clave en los avances actuales y futuros de la *World Wide Web* y de los *Sistemas de Información*.

El conocimiento es otro concepto clave y que ha sido ampliamente y durante mucho tiempo debatido por muchos autores. Fue durante los siglos diecisiete y dieciocho cuando la discusión entre el racionalismo y el empirismo tuvieron sus puntos más álgidos en relación a una parte de la epistemología, la naturaleza del conocimiento humano. Para los empiristas, como David Hume (1711-1776) o Isaac Newton (1642-1727), el conocimiento humano se deriva de la interacción, sensorial o empírica con el mundo. En cambio, los racionalistas niegan esto, afirmando que no todo el conocimiento surge de la experiencia. Representantes racionalistas clásicos más destacados han sido autores como Rene Descartes (1596-1650), Gottfried W. von Leibnitz (1646-1716), Benedict Spinoza (1632-1677), y más contemporáneo Immanuel Kant (1724-1804), que incluyeron el conocimiento de Dios, la sustancia y las ideas (Wilson R. y Keil F., 1999).

El estudio de la cognición tiene sus bases filosóficas en estos dos movimientos. Mientras el empirismo estudia la percepción, el aprendizaje y la memoria; el racionalismo estudia el pensamiento, el razonamiento y la resolución de problemas. Pero aún existen diferencias en las perspectivas sobre el concepto de conocimiento, que todavía hoy están presentes entre las disciplinas sociales y técnicas. Si bien el campo de la Informática y la Ciencia del Conocimiento llevan años estudiando las formas de representación del conocimiento, la inferencia, la resolución de problemas y

el aprendizaje; ha sido con posterioridad, y a raíz de la importancia que ha tomado la información y el conocimiento como un recurso intangible más, dentro de las organizaciones, cuando se ha empezado a estudiar la Gestión del Conocimiento y la Innovación como valores intangibles de las organizaciones. Como consecuencia, existe una mayor integración entre las Tecnologías de la Información y los Recursos Humanos, y mayores niveles de comunicación y transferencia de información y conocimiento hacia medios informáticos (Jiawei y Fu, 1994). Aunque, sin embargo, la investigaciones sobre gestión del conocimiento han estado infravalorando el papel de las tecnologías en el área.

Por ello y como punto de inicio, haciendo referencia de nuevo, a la RAE, podemos definir *conocimiento como la acción y efecto de averiguar por el ejercicio de las facultades intelectuales la naturaleza, cualidades y relaciones de las cosas*. Durante muchos años, diversos autores han presentado distintas tipologías de conocimientos, si bien, un ejemplo de ello, lo tenemos en el año 1949 cuando el filósofo inglés Gilbert Ryle publica su libro *The Concept of Mind*, y dedica acertadamente, un capítulo a distinguir entre el conocimiento sobre las acciones y procedimientos, conocido como *know-how*, y el conocimiento acerca de los hechos (*facts*), también llamado *know-that*. Además, dentro de esta distinción presenta la diferencia entre *estupidez*, no saber cómo, e *ignorancia*, no saber qué (Ryle, G. 1949). Posteriormente, otros autores han definido otras categorías de conocimiento, ampliamente conocidas, como es el caso del conocimiento explícito, es decir, el conocimiento del que es consciente el ser humano y puede describirlo de manera simbólica; y el conocimiento tácito, aquel que aún residiendo en el ser humano, éste no puede ser identificado o descrito de forma sistemática.

Asimismo, los Sistemas de Información, a través de sus bases de datos, han permitido gestionar las afirmaciones (*facts*) de su entorno a través del almacenamiento de datos; y en la secuencia de acciones (la programación), la forma de realizar las operaciones y desarrollar los procesos que se llevan a cabo en las organizaciones. Así, se ha podido conseguido hacer explícito una parte importante del *know-how* y del *know-that*. Por consiguiente, resulta conveniente en este punto, considerar que:

- Por lo general, el uso de los Sistemas de Información provoca un aumento del *know-that*, a través del incremento de registros transaccionales en la base de datos, y que en muchas ocasiones no son conocidos por los usuarios del sistema, pudiéndose identificar aspectos tales como tendencias de mercado, perfiles de usuarios o puntos de inflexión, por lo que ha surgido la necesidad de aplicar técnicas de *Minería de Datos* para descubrir este nuevo conocimiento.
- La modificación del *know-how*, a raíz, por ejemplo, del entorno cambiante de las organizaciones, ha sido una de las causas más importantes de la necesidad del costoso mantenimiento de los Sistemas de Información. Este hecho ha provocado la investigación de *Sistemas Adaptativos* y metodologías de desarrollo rápido de sistemas como es el caso de *Agile Software Development* (Cockburn, 2002).
- Y, además, los Sistemas de Información no poseen una meta descripción operativa (semántica), no sólo de su *know-that* sino tampoco de su *know-how*. Este hecho ha quedado patente en los últimos años con el desarrollo de la *World Wide Web*, que ha provocado la necesidad de gestionar contenido más que datos o simples símbolos debido al desbordamiento de la información y la desinformación publicada diariamente en la World Wide Web (WWW). Como resultado, se ha fomentado la investigación en el campo de la Semántica y la Ontología como es el caso de la *Web Semántica*. En este sentido, estas investigaciones facilitarán la obtención de un mejor nivel de comunicación *hombre-máquina* y *máquina-máquina*, y un nivel más eficiente de *adaptación-mantenimiento* de los Sistemas de Información.

No obstante, el problema central del la Gestión del Conocimiento Explícito no es su representación, sino conocer qué lo importante en el momento de captar o filtrar sólo aquellos elementos necesarios para ese conocimiento. Representa así, un fenómeno integrador en el que cada pieza que conforma ese conocimiento se relacionada con cada una de las piezas restantes (Chunk 1995).

En este segundo capítulo de la tesis, se presentan algunos de los aspectos más significativos de las investigaciones sobre las TICs aplicadas a la Gestión del

Conocimiento. Para ello, su contenido se estructura en siete apartados: en el primer apartado *2.2.- Los Sistemas de Información e Internet*, se incluye una introducción al binomio de los Sistemas de Información e Internet, intentado plasmar algunas de las tendencias actuales de esta área de investigación, que provocan y sin duda provocarán, en el futuro, algunos cambios muy significativos de nuestras organizaciones. Posteriormente, En el apartado *2.3.- Representación del Conocimiento basado en Símbolos*, se presentan las técnicas más importantes para la representación de los diferentes tipos de conocimiento explícito. Por su parte, el apartado *2.4.- Representación No-Simbólica del Conocimiento*, se analizan las principales técnicas de representación no-simbólica del conocimiento. En el apartado *2.5.- Inferencia del conocimiento*, se describen las formas de inferencia más relevantes.

Asimismo, el apartado *2.6.- Descubrir Conocimiento*, integra las metodologías de aplicación y las fuentes de datos existentes. En el apartado *2.7.- Funciones Básicas de la Minería de Datos*, se incluyen las funciones de mayor interés en la minería de datos. Y, por último, en el apartado *2.8.- Tendencias Tecnológicas*, se analizan las principales tendencias relacionadas con las TICs y la WWW que dibujarán el nuevo escenario tecnológico de los próximos años.

2.2. LOS SISTEMAS DE INFORMACIÓN E INTERNET

Los programas informáticos no se crean sólo por el hecho de que una persona se sitúe delante de un ordenador y lo programe. De hecho, muchos de los programas actuales pueden contener más de un millón de instrucciones, ejecutarse de manera distribuida y coordinada con otros programas, incluso de forma remota, lo cual supera la capacidad de comprensión de cualquier ser humano, independiente de su grado de pericia en el área. A esto hay que añadir la complejidad del proceso de desarrollo (análisis, diseño, programación, auditoría y mantenimiento), que hace necesario equipos completos de profesionales y expertos multidisciplinares (Norman A. Donald, 2000).

No obstante, los Sistemas de Información juegan en la actualidad un papel fundamental en todas nuestras transacciones diarias, pero requieren a su vez, una adaptación al entorno cambiante y dinámico en el que vivimos. La aparición de *Internet* y la *World Wide Web* está provocando la necesidad de integrar los datos, la información, el conocimiento y los servicios que ofrecen las organizaciones a través de Internet y la *Web*. Como consecuencia, existe la necesidad de desarrollar tecnologías que cubran esa laguna y faciliten no sólo el interface de estos sistemas, sino también la seguridad de las comunicaciones.

Aunque en los últimos años, el desarrollo de la primera fase de la telefonía móvil sólo haya requerido el uso de terminales de telefonía para la comunicación interpersonal mediante la voz, sonidos y pequeños mensajes de textos (*Short Message Systems*, SMS). A medida que se amplíe el ancho de banda de estas redes y se provoque un desarrollo y una integración, no sólo de las tecnologías hardware utilizadas en estas terminales, sino también de las tecnologías software y de los servicios que se puedan ofrecer, se observará una mayor necesidad de integración de la telefonía móvil, Internet y los Sistemas de Información. Un ejemplo básico y elemental de los inicios de esta integración la podemos observar en el uso de los SMS en la comunicación desde *Sistemas de Información* y la *Web* a los usuarios de las terminales.

La empresa Mobipay España S.A., por ejemplo, presenta una de las aplicaciones arriba comentadas, al facilitar el pago seguro a través de los teléfonos móviles, mediante la activación de tarjetas físicas o virtuales, de crédito, débito o prepago, en comercios físicos o virtuales, estáticos o móviles (ver figura 2.1).

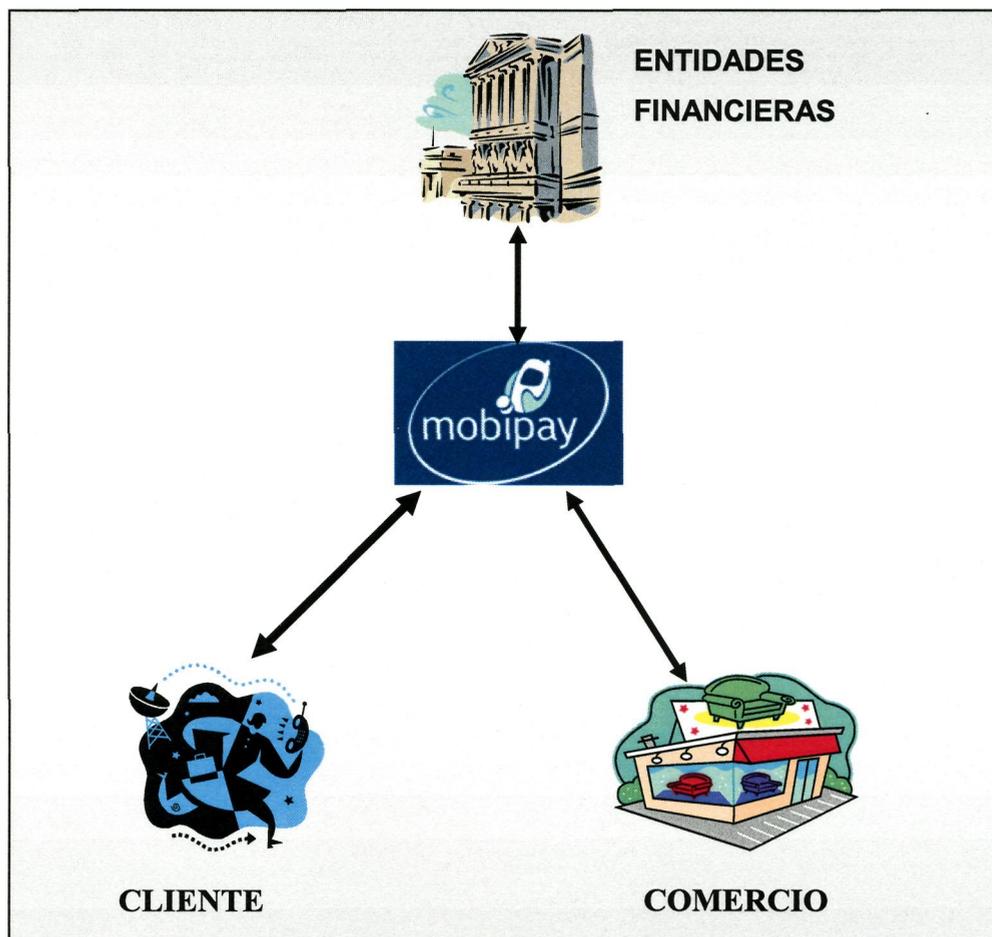


Figura 2.1: Sistema de Pago por teléfono móvil Mobipay

Como se puede apreciar en la figura 2.1, cuando se realiza un pago, por ejemplo en un comercio, el comprador o cliente facilita al comerciante el número de teléfono o un número de identificación del sistema mobipay. Dicho comerciante introduce inmediatamente esta identificación en el terminal punto de venta, y en pocos segundos, el cliente recibe un mensaje en su móvil solicitando la autorización de la transacción. Asimismo, esta autorización la realiza el cliente mediante la introducción del número secreto asociado al sistema mobipay. Por su parte, el sistema mobipay realiza las comprobaciones necesarias y confirma el pago de la operación mediante un mensaje al comerciante y al cliente. Este tipo de operativa es similar en el caso de comercios en movilidad (ejemplos taxis o servicios a domicilio).

Por otro lado, si el pago se realizara a través de Internet, el comprador que desea ejecutar una operación de compra a través de Internet, selecciona los servicios o

productos que desean y a través del website solicita posteriormente pagar el importe de su compra, utilizando el servicio mobipay. En este caso, el comercio virtual envía un número de referencia que aparecerá en la pantalla de ordenador del cliente. Posteriormente, el cliente introduce en su móvil el número de referencia y su número secreto. De esta forma, la transacción queda abonada. Algo parecido sucede en máquinas expendedoras y en otros contextos de pago.

Entre las operaciones de pago que se pueden realizar se encuentran: pagos en comercios, pagos en comercios móviles (por ejemplo en un taxi), pagos por Internet, pagos en máquinas expendedoras, envío de dinero a otra persona, recarga del teléfono móvil, reserva y pago de servicios o pago de facturas, entre otras. En relación a los aspectos de seguridad, el sistema utiliza los siguientes elementos: número secreto asociado, sistemas de alarmas y bloqueos, comunicaciones encriptadas, asociación de la tarjeta SIM del cliente al sistema, PIN del teléfono móvil. Al mismo tiempo, el sistema ofrece una serie de ventajas a los comercios, tales como: rapidez en operación de cobro, seguridad ya que se elimina la manipulación de dinero físico, fiabilidad y control en las transacciones y, por último, facilidad de uso.

2.2.1.- Semántica y Ontología

El incremento impensable de los volúmenes de datos e información, su sobrecarga, y la necesidad de optimizar la gestión de los recursos intangibles hace necesaria la aplicación de mecanismos que faciliten la gestión inteligente de contenidos más allá de la gestión de datos. En este sentido, en los últimos años se han producido avances muy significativos en esta área de conocimiento.

Uno de los campos donde se ha observado un mayor avance ha sido en la *Web*. Como consecuencia del desarrollo de tecnologías y lenguajes como el XML (Bray *et al.*, 2000; Patel-Schneider y Siméon, 2002); el XHTML (W3C HTML Working Group, 2003); el RDF (Broekstra *et al.*, 2001); y el OWL (Smith *et al.*, 2003), entre otras. Surge así, la *Web Semántica* por uno de sus precursores más importantes y padre de la *Web*, Tim Berners-Lee. Este conocido investigador en una de sus publicaciones más importantes, expone:

“Tuve un sueño de la Web....y tenía dos partes:

La Web, en su primera parte se convirtió en un poderoso mecanismo de colaboración global entre las personas.... En su segunda parte, esta colaboración se extendió a los ordenadores. Las máquinas serían capaces de analizar todos los datos que existieran en la Web. .. Una vez que la segunda parte se completaba, la Web sería un lugar donde el capricho de un ser humano y el razonamiento de una máquina coexistían en una mezcla ideal y poderosa”.

(Adaptado de Berners-Lee, 2000)

Para Berners-Lee y otros autores, la *Web Semántica* nos va a facilitar el procesamiento de contenidos. Eso no quiere decir que la *Web* piense, pero igual que el submarino no nada y nos permite adentrarnos en profundidades marinas que jamás podríamos llegar nadando, la *Web Semántica*, junto a los agentes inteligentes, la minería de datos, los algoritmos evolutivos y el resto de tecnologías, nos pueden ofrecer capacidades jamás imaginadas por los seres humanos. Sin embargo, el camino que queda por andar es largo y falta mucho por hacer (Patel-Schneider y Fensel, 2002).

Dichos avances tecnológicos, no sólo afectan a la *Web*, también tienen su efecto en los Sistemas de Información y en las Organizaciones. En este sentido, los Sistemas de Información se convierten en Sistemas de Gestión de Conocimiento o Sistemas Basados en el Conocimiento, interconectados sobre la plataforma Internet y, en un futuro, con la informática móvil (Sure *et al.*, 2002). Como consecuencia de ello, estos avances tendrán importantes efectos sobre las organizaciones y en especial sobre los trabajos basados en el conocimiento (Gordon, 2002). Aunque, todavía esta sigue siendo una laguna importante que los investigadores sobre gestión del conocimiento han infravalorado, y que esta investigación pretende resaltar.

Así, nuevos protocolos de comunicación y lenguajes de programación están siendo desarrollados rápidamente para intentar estandarizar las formas en las que los sistemas describen lo que hacen (Hendler *et al.*, 2002). El punto de inicio que ha sido utilizado para conseguir este nuevo escenario es el uso de la Ontología, entendida *como la especificación explícita y formal de una conceptualización compartida por humanos y programas informáticos, de un dominio o área de conocimiento específica*. Esta tarea se hace posible gracias al empleo de metadatos, que permiten una meta descripción simbólica de los datos, eventos, procesos u objetos a procesar.

Las Ontologías están presentes cada vez más en aplicaciones informáticas complejas, lo que permite tareas tales como el intercambio y almacenamiento de datos, el razonamiento basado en Ontologías, la navegación basada en la Ontologías y la evolución de Ontologías. Para desarrollar este tipo de sistemas se requiere, en la mayoría de los casos, la combinación de más de una simple herramienta de software (Regan y Slator, 2003).

2.2.2.- Agentes Inteligentes

Los agentes Inteligentes representan una de las tecnologías claves para la evolución no sólo de Internet y la Web, sino incluso de las telecomunicaciones. Estos programas informáticos (o sistemas) pueden tener acceso a fuentes de información múltiples, heterogéneas y geográficamente distribuidas. Facilitan la gestión de la información y el conocimiento, además de recuperar, analizar, gestionar e integrar multitud de datos e información necesaria, en un tiempo record.

Existen multitud de métodos y aplicaciones desarrolladas y basadas en Agentes Inteligentes disponibles para *Bases de Datos Avanzadas* (Perkowitz y Etzioni, 1994; McKay *et al.*, 1990; Pastor *et al.*, 1992; Pitoura y Bhargava, 1995), *Sistemas Basados en el Conocimiento* (Haugeneder *et al.*, 1994), y *Sistemas de Información Distribuidos* (Berghoff *et al.*, 1996; Satyanarayanan *et al.*, 1994), etc... Pero, ha sido durante los últimos años, con la necesidad de desarrollar la *Web Semántica* (McDermott *et al.*, 2001; Stuckenschmidt, 2002) y la *Minería de Datos* (Bala *et al.*, 2002; Reiterer *et al.*,

2000), cuando los Agentes Inteligentes se han convertido en un área clave de investigación.

Son muchos los tipos de Agentes Inteligentes existentes, aunque se puede establecer algunas categorías básicas como son:

- *Agentes Sociales o Multi-agentes*, poseen capacidades similares a los grupos sociales. Estos agentes envuelven una abstracción de la teoría Sociológica y de la Organización, que permiten modelos sociales de agentes. Uno de los sistemas con mayor presencia en la investigación de modelos multi-agentes lo encontramos en el sistema Swarm (Daniels, 1999). Un software que permite la simulación multi-agente de sistemas complejos. Sus aplicaciones son muy variadas y cubren desde las aplicaciones en la localización distribuida de fallos en redes (White, Bieszczad y Pagurek, 1998) hasta la simulación empresarial (Terna, 2003).
- *Agentes Racionales*, desde un sentido económico. En este caso los agentes cooperan con el fin de maximizar el beneficio (Wooldridge y Rao, 1999). Por lo tanto, la acción racional es aquella que maximiza las preferencias. Una de las aplicaciones más importantes de los Agentes Racionales se encuentra en los problemas de negociación (Sandholm y Lesser, 1995)
- *Agentes Adaptables o Evolutivos*, capaces de adaptarse a los cambios del entorno. Un claro ejemplo de este tipo de Agentes son los Asistentes personales (Maes P., 1994), aprendiendo y personalizando el interface, modelando al usuario del sistema. En esta categoría se incluyen sistemas de agentes que incluyen otras tecnologías como puede ser: los Algoritmos de Aprendizaje (automático o supervisado), Algoritmos Evolutivos y la Minería de Datos.
- *Agentes Móviles*, capaces de viajar de forma autónoma a través de redes como Internet (Dömel, 1996). Las aplicaciones de estos Agentes son muy variadas, desde seguridad en redes (Tardo y Valente, 1996) hasta telefonía móvil (Adler et al., 2001).

- *Agentes Interactivos*, tienen la capacidad de interactuar con otros agentes que existen y actúan en una cierta proximidad. Esta interacción puede ser intencionada o no. Y en la mayoría de los casos esta comunicación implica un proceso mucho más complejo que una simple descripción programada ya que puede presentar aspectos tales como: el discurso, la negociación, la conformidad, las dimensiones del significado, significado convencional frente al significado personal o individual, la perspectiva (el agente que emite el mensaje, el receptor del mensaje y el resto de los agentes de la sociedad u otros), la semántica y el contexto. Aplicaciones de Agentes Interactivos como Agentes Inteligentes en simulación de entornos dinámicos (Tambe *et al.*, 1995).
- *Agentes híbridos*, son aquellos que combinan varias de las características representativas. En este caso las posibles aplicaciones de estos agentes puede ser cualquiera como por ejemplo en el desarrollo de una nave espacial autónoma (Pell *et al.*, 1998).

En la figura 2.2, se puede observar cual es el escenario que se propone para la *Web Semántica*, propuesta por Berners Lee, en la que se combinan Agentes Inteligentes (Agents) que interactúan con los Usuarios, Portal Comunitario Web (Community Portal) utilizado por los Agentes y los Usuarios, Bancos de Metadatos (Metadata Repository), Motor de inferencia (Inference Engine), Páginas Web, Ontologías, Software de Gestión Ontológico, Herramientas para la Construcción de Ontologías y Programas de Edición de Páginas Web. Todo ello integrado de forma harmónica, permite la ejecución de tareas, basadas en contenidos, de forma automática o semiautomática y facilita al usuario y a las organizaciones unos niveles de aplicación, que distan de forma significativa y cualitativa de las posibilidades actuales de la WEB. Produciéndose, de esta forma, una evolución en el uso y el significado que tiene la WWW para los mercados, los usuarios y las organizaciones, pero que abre, a su vez, una ventana hacia un nuevo entorno de requerimientos tecnológicos, que podría provocar mayores niveles de diferenciación tecnológica, afectando de manera significativa a las capacidades de las organizaciones y los usuarios.

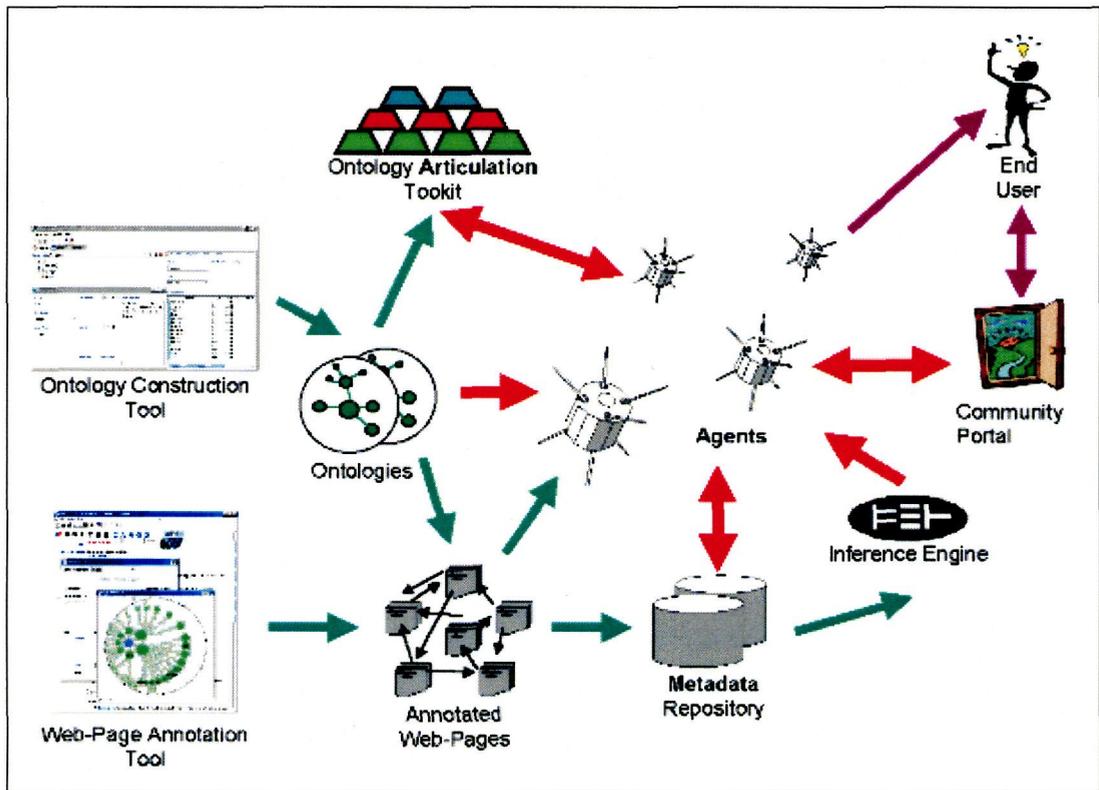


Figura 2.2: La figura presenta el portal comunitario en relación a otras partes de la Web Semántica. (Fuente: <http://www.semanticweb.org/about.html#bigpicture>)

A continuación se muestra una relación de las herramientas software (ver cuadro 2.1), que facilitan o permiten el desarrollo de Sistemas de Agentes. Esta relación se recoge en la Publicación Europea *Review of Software Products for Multi-Agents Systems* (Mangina 2002). En el cuadro se puede observar el nombre del producto, el tipo de producto (lenguaje de programación, entorno de desarrollo u otro), el tipo de licencia de la que dispone (comercial, libre, académica u otra) y el proveedor que suministra el producto.

Como se puede observar en la lista de proveedores, cabe destacar organizaciones de reconocido prestigio internacional como el DARPA (Defense Advance Research Project Agency), el Media Lab del MIT (Massachussets Institute of Technology), la Universidad de Carnegie-Mellon, la Universidad de Birmighan, la Universidad de Bonn, la Universidad de Chicago. Asimismo, en el grupo de empresas se incluyen: Toshiba, Motorola, Siemens, BT o Logica UK.

Es importante reconocer que estos productos se encuentran en continua evolución y también que esta relación de productos ha sido generada en un momento determinado, con las limitaciones que este tipo de estudios tiene que afrontar. Además, el cuadro recoge varias de las herramientas de las que se hace uso en el desarrollo del sistema de gestión avanzado para la gestión del conocimiento aplicado a la competitividad internacional de los territorios propuesto en esta tesis. Estas herramientas están integradas específicamente por Jess y Jade, que serán analizadas y aplicadas en el capítulo 3 con una simulación del sistema propuesto.

ID	PRODUCTO	TIPO	LICENCIA	PROVEEDOR
1	ADK	Lenguaje o entorno	Comercial	Tryllian
2	AgentSheets	Lenguaje o entorno	Comercial	AgentSheets Inc.
3	AgentTool	Enfoque de ingeniería de Sistema Multiagente	Académica (Libre)	Kansas State
4	Bee-gent	Lenguaje o entorno	Comercial (Libre)	Toshiba Corporation
5	CABLE	Lenguaje o entorno	Comercial	Logica UK Ltd
6	Comet Way	JAK para servicios automatizados	Comercial (Libre)	Comet Way Ltd.
7	CORMAS	Lenguaje o entorno	Comercial (Libre)	CIRAD
8	Cougaar	Lenguaje o entorno	(Libre)	(DARPA) / BBN
9	DECAF	Lenguaje o entorno	(Libre)	University of Delaware
10	EXCALIBUR	EXCALIBUR Agent	Architecture	AI center
11	FIPA-OS	Lenguaje o entorno	Comercial (Libre)	Emorphia Ltd.
12	Grasshopper	Lenguaje o entorno	Comercial	IKV++ Technologies
13	IDOL	Lenguaje o entorno	Comercial	
14	IMPACT	Metodología para diseño/desarrollo	Académica	University of
15	JACK	Lenguaje o entorno	Comercial	Agent Oriented
16	JADE	Software de soporte	Comercial (L general)	TILAB
17	JADE / LEAP	Distributed Agent	Comercial (L general)	Motorola, ADAC, Uni Parma, Siemens
18	JAFMAS /	Lenguaje o entorno	Académica	University of Cincinnati
19	JATLiteBean	Software de soporte	Académica (Libre)	University of Otago
20	JESS	Lenguaje o entorno	Académica (Libre)	Sandia Nacional Laboratory
21	Kaariboga	Lenguaje o entorno	Other (Libre - BSD-	Other
22	LEE	MAS modelados/simulación	Académica (Libre)	University of
23	Living	Lenguaje o entorno	Comercial	Living Systems AG
24	MAML	Basado en MAS modelados/simulación	Comercial (Libre)	Agent-Lab Ltd.
25	MAP / CSM	Simulación toolkit	Académica (Libre)	University of Bonn
26	MASSVVE KIT	Lenguaje o entorno	Académica (Libre)	Federal University
27	NARVAL	Entorno de desarrollo	Comercial (GN general)	Logilab
28	RePast	Basado en MAS modelados/simulación	Académica (Libre- BSD)	University of Chicago
29	RETSINA	Lenguaje o entorno	Académica	Carnegie Mellon
30	SEMOA	Software de soporte	Other (Libre)	Fraunhofer Institut fuer Graphische Datenverarbeitung
31	SIM_AGENT	Lenguaje o entorno	Académica (Libre)	University of Birmingham
32	StarLogo	Lenguaje o entorno	Académica (Libre)	MIT MEDIA lab
33	TuCSon	Software de soporte	Académica	DEIS Universita' di Bologna
34	Voyager	Software de soporte	Comercial	Recursion Software,
35	Xraptor	Código fuente	Académica (Libre)	Johannes Gutenberg
36	ZEUS Agent Toolkit	Lenguaje o entorno	Comercial (Libre)	BTextact Technologies

Cuadro 2.1: Lista de Productos para el desarrollo de Sistemas Multi-agentes

Algunas de las áreas de aplicación donde los Agentes Inteligentes juegan un papel fundamental se corresponden con:

- *Inteligencia Ambiental*, la distribución de informática ubicua, la comunicación continua y los interfaces de usuario inteligentes para dispositivos personales o industriales, etc....
- *Grid Computing* (Informática-procesamiento distribuido), en los que el uso de sistemas multi-agentes permiten el acceso a recursos con grandes capacidad de procesamiento aplicables a la ciencia, la medicina, la ingeniería, las aplicaciones comerciales e industriales etc...
- *Negocio Electrónico*, donde los sistemas de agentes están ya permitiendo la automatización o semi-automatización de las tareas de obtención de datos, y transacciones de compra a través de Internet.
- *Web Semántica*, donde los agentes cumple con tareas como facilitar el acceso a servicios y optimizar el uso de recursos en ocasiones en coordinación con otros agentes.
- *Bioinformática y la Biología Informática*, donde los Agentes Inteligentes pueden realizar la explotación coherente de la revolución de datos que ocurre en biología, y otros; incluyendo monitorización, control, gestión de recursos, y espacios, y otras aplicaciones como militares o industriales (Luck, McBurney y Prist, 2003).

No obstante, existen algunos desafíos que aun no han sido cubiertos y que requieren trabajo de investigación para los próximos años. Entre los más destacados se encuentran:

- Incremento de la calidad de las metodologías y las herramientas de desarrollo de sistemas multi-agentes a estándares industriales.
- Acuerdos estándar satisfactorios para el desarrollo de sistemas abiertos.

- Infraestructuras semánticas para comunidades abiertas de agentes.
- Desarrollo de capacidades de razonamiento para agentes en entornos abiertos.
- Desarrollo de habilidades de agentes para entender los requerimientos de usuarios.
- Desarrollo de habilidades de Adaptabilidad de Agentes a entornos dinámicos.
- Desarrollo de la confidencialidad y confianza de Usuarios en los Agentes.

2.2.3.- Introducción al Conocimiento Explícito

Como ya explicaba Drucker en su libro *Managing in a Time of Great Change* (Ruggles, 1998), el conocimiento se ha convertido en un recurso económico clave y – quizás la única – fuente dominante de la ventaja competitiva de las empresas.

El conocimiento en general en las organizaciones, se encuentra en diferentes formas, *implícito* en los agentes que componen la organización y su entorno, es decir, sus miembros, sus agentes externos (*ej.* proveedores), etc.; y *explícito* en sus sistemas informáticos y documentos que contienen información sobre dichas organizaciones. Estas formas integran las dos clasificaciones de conocimiento más aceptadas en la literatura y que se centran en el grado de codificabilidad del conocimiento, es decir la distinción entre conocimiento tácito y explícito (eg. Polanyi, 1958; Nelson y Winter, 1982; Van Geenhuizen, 1993; Hedlund, 1994; Nonaka y Takeuchi, 1995; Grant, 1996a, 1996b y 1997; Lyles y Steemsma, 1997; El Sawy, et al., 1997; Leonard y Sensiper, 1998).

El conocimiento tácito es el conocimiento personal, específico del contexto y difícil de formalizar y comunicar (Garud y Nayyar, 1994; Nonaka y Takeuchi, 1995; y otros). Se encuentra inmerso en los procedimientos y rutinas, las acciones, las emociones, valores e ideas y en el compromiso (Cohen y Bacddayan, 1994; Schön, 1983; Winter, 1987). Y, como consecuencia, integra un conjunto de creencias, valores y percepciones que influyen sobre el conocimiento explícito (Fahey y Prusak, 1998). El conocimiento explícito, por otro lado, se define como aquel conocimiento que se puede transmitir en lenguaje formal y sistemático, al ser codificado en datos, números palabras, fórmulas matemáticas, manuales, etc. (Nonaka, et al., 2001). En resumen, en símbolos.

Algunos autores tales como Teece (1998) argumentan que el conocimiento explícito se trata a través del uso de las tecnologías de la información como herramienta central para su transferencia, dado que no se requiere contactos personales. Por consiguiente, en el transcurso de esta investigación, relacionada con las tecnologías de la información aplicadas a la gestión del conocimiento, trataremos principalmente la tipología de conocimiento explícito.

A medida que la organización formaliza sus operaciones diarias e interacciona con su entorno, evoluciona internamente, se adapta y aprende. De esta forma, se genera una cantidad considerable de datos que pueden convertirse en información y que podría facilitar el aprendizaje y adaptación de los miembros de la organización. Aunque, en la mayoría de los casos, los datos pasan inadvertidos y sólo una pequeña parte de los mismos, se considera para la toma de decisiones. Pero, no sólo el *know-that* o *facts* de las transacciones no se optimizan, de igual forma, una importante cantidad de conocimiento tácito de los miembros de la organización, tanto en *know-that* como en conocimiento know-how, tampoco se optimiza.

Se han realizado múltiples trabajos de investigación dirigidos a transformar el conocimiento tácito en conocimiento explícito (Rodhain, 1999; Busch *et al.*, 2001; Conrado y Dybå, 2001; Busch y Richards, 2001) y se ha conseguido simular capacidades humanas difíciles de explicitar, como la visión (Hertz, 1988; Bowden *et al.*, 2002; Freeman *et al.*, 2000), el tacto (Basdogan *et al.*, 2000), caminar, saltar, nadar (Sims, 1994), planificar (Husbans *et al.*, 1994; Aler *et al.*, 2001), diseñar (Bandini y Manzini, 2001) etc... En la mayoría de los casos, se ha conseguido esta simulación, haciendo uso de sistemas *no simbólicos* tales como *Redes Neuronales Artificiales* o *Sistemas Evolutivos*. Una de las dificultades que se encuentra en los sistemas *no simbólicos* es la falta de meta descripción del aprendizaje o razonamiento. Es decir, cuando se consigue el objetivo o el nivel de aprendizaje o búsqueda deseado, es bastante difícil obtener una descripción simbólica y entendible por el ser humano. En este sentido, este aspecto nos podría indicar que puede existir una relación entre el conocimiento difícil de explicitar, el conocimiento tácito, y los sistemas *no simbólicos*. Una interesante y futura materia de estudio.

En la mayoría de los casos, el uso de los sistemas no simbólicos aparecen en dominios muy pequeños y específicos, adquiriendo un nivel de éxito bastante considerable. No obstante, es posible conseguir la sinergia combinando el carácter no simbólico de estas tecnologías con el carácter simbólico de otras como son los Sistemas Ontológicos. Esta es, a su vez, un área de gran interés para la comunidad científica en la actualidad.

En el momento de explicitar conocimiento, existen dos aspectos básicos a estudiar. Por un lado, la formalización o técnica de representación del conocimiento (*know-that*), y por otro, el mecanismo de inferencia o razonamiento que se utiliza (*know-how*). En la mayoría de los casos, el dominio o contexto del conocimiento define la técnica de representación a utilizar, y esta a su vez, define el mecanismo de inferencia. En el próximo apartado, se estudia las técnicas más comunes en la representación de conocimiento, que hasta hoy han sido aplicadas en muchas aplicaciones reales y, en algunos casos, desde hace ya varios años. En segundo lugar, se hace referencia a las técnicas de inferencia que son utilizadas con carácter general.

2.2.4.- Representación del Conocimiento

La Representación del Conocimiento es un área de estudio que incluye la tarea de explicitar información y conocimiento, que puede ser codificada y utilizada en modelos cognitivos formales o informatizados. Las sub-áreas de la representación del conocimiento pueden abarcar desde los aspectos puramente filosóficos de la epistemología, hasta los problemas más prácticos de gestión de grandes cantidades de datos. Esta diversidad se unifica en todas sus formas en el problema central de codificación del conocimiento humano, para que dicho conocimiento pueda ser posteriormente utilizado. Asimismo, esta meta está representada de forma resumida en la hipótesis de representación de conocimiento de Smith (1982)

Según este autor, cualquier proceso inteligente y mecánicamente compuesto, estará integrado por ingredientes estructurales que, por un lado, nosotros como observadores externos, y de forma natural, tomamos para representar una estructura proposicional

del conocimiento que presenta el proceso en su totalidad.; y por otro lado, que independientemente de dicha atribución semántica externa, el proceso juega un papel formal, aunque causal y esencial en la creación del comportamiento que manifiesta ese conocimiento.

Una representación con éxito de algún tipo de conocimiento debe, por consiguiente, estar en una forma entendible por los humanos y debe también hacer que el sistema utilice el conocimiento y se comporte como si lo conociese. Los ingredientes estructurales que alcanzan estas metas se encuentran normalmente en los lenguajes para la representación del conocimiento. Dichos lenguajes han sido desarrollados durante años.

Sin embargo, no debe obviarse, que la programación informática es un proceso de representación del conocimiento, esencialmente, *know-how*. El medio utilizado para esta representación, los lenguajes de programación, no satisfacen por completo el criterio propuesto en la hipótesis anterior. Asimismo, estos lenguajes afectan, principalmente, al comportamiento que es consistente con el conocimiento representado. Aunque, es conocida su dificultad para ser entendida por el ser humano. Por lo general, los sistemas que en la actualidad presentan un cierto nivel de desarrollo en lo que se refiere a la representación del conocimiento, poseen un modelo algo parecido al que se recoge en la figura 2.3.

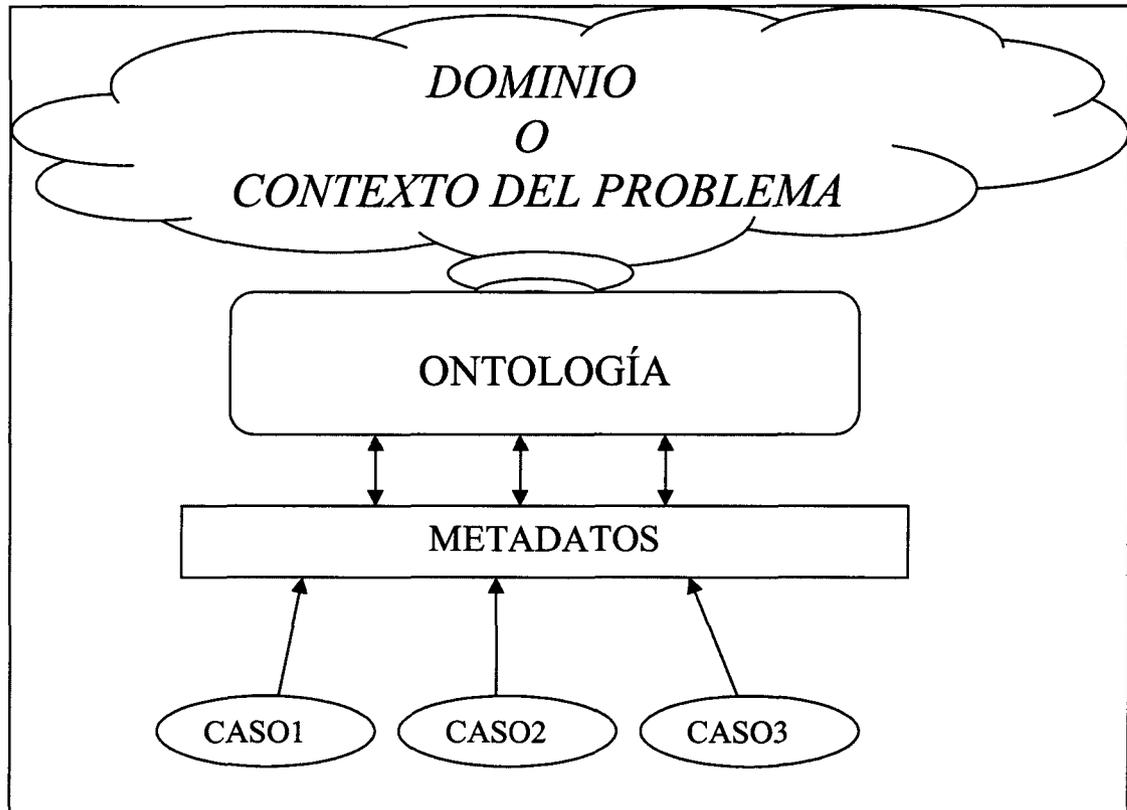


Figura 2.3: Estructura jerárquica de la semántica basada en ontología

Como se puede apreciar en la figura, varios pueden ser los niveles que integran la estructura jerárquica de la semántica basada en modelos ontológicos. El dominio o contexto del problema, queda representado en el modelo ontológico (ontología) a través de la descripción de sus contextos y la relación entre los mismos. Por ejemplo, la competitividad internacional, como dominio o contexto del problema a tratar, puede ser representada en un modelo ontológico en el que se describan los conceptos como país, organización internacional, factores de competitividad, etc..., así como las relaciones entre ellos. Estas descripciones de conceptos y sus relaciones, representan a su vez una meta descripción (metadatos) de cualquiera de los casos o *facts* y reglas sobre la competitividad. Como puede ser el caso del factor *actividad económica* o del país *España*.

Asimismo, las técnicas de representación del conocimiento se pueden clasificar en dos grandes grupos: aquellas técnicas que se basan en símbolos para presentar el

conocimiento, y las que no requieren de símbolos o entidades simbólicas, y que representan conceptos, eventos, objetos etc.... Estas últimas, tienen sus fundamentos, esencialmente, en modelos estocásticos o matemáticos.

2.3. - REPRESENTACIÓN DEL CONOCIMIENTO BASADO EN SÍMBOLOS

En los últimos años, han sido muchos los esquemas de representación que se han propuesto e implementado, cada uno de ellos posee sus fortalezas, debilidades y áreas de aplicación. Mylopoulus y Levesque (1984), propusieron una clasificación basada en cuatro categorías básicas: Representación Lógica, Representación Procedural, Representación basadas en Redes y Representación Estructurada.

Como se puede observar en el cuadro 2.2, los tipos de representación simbólica de conocimiento explícito, es decir, basada en símbolos, pueden ser clasificados en cinco grupos, estos son: las técnicas de representación simbólica basada en la lógica, las técnicas de representación simbólica Procedural, las técnicas de representación simbólica basada en Redes, las técnicas de representación simbólica basadas en Estructuras y, finalmente, las técnicas de representación simbólica Híbridas. Para cada uno de estos tipos de representación existen las técnicas de representación más significativas. El cuadro incluye: la lógica clásica, la lógica difusa, la lógica temporal, la lógica espacial, la lógica causal, la lógica modal y la lógica no monótona, incluidas todas en el tipo de representación basada en lógica. Asimismo, encontramos los grafos conceptuales para la representación basada en redes. Y los marcos, scripts y casos, para la representación en estructuras. Cabe añadir que para la representación procedural y la representación híbrida se utilizan diferentes tipos de técnicas.

A continuación, se describen cada uno de los tipos de representación simbólica, arriba enumeradas, y sus respectivas técnicas de representación.

TIPO DE REPRESENTACIÓN SIMBÓLICA	TÉCNICA
Representación basada en Lógica	<ul style="list-style-type: none"> • Lógica Clásica • Lógica Difusa • Lógica Temporal • Lógica Espacial • Lógica Causal • Lógica Modal • Lógica No Monótona
Representación Procedural	<ul style="list-style-type: none"> • Varios tipos
Representación basada en Redes	<ul style="list-style-type: none"> • Grafos conceptuales
Representación en estructuras	<ul style="list-style-type: none"> • Marcos • Scripts • Casos
Representación híbrida	<ul style="list-style-type: none"> • Varios tipos

Cuadro 2.2: Técnicas de representación simbólica del conocimiento.

2.3.1.- Representación Lógica

En este caso, el conocimiento se representa a través de expresiones lógicas formales. La Lógica de predicados de primer orden es la más utilizada. Tiene sus orígenes en las Matemáticas y esta a su vez en la Filosofía. Un ejemplo de representación lógica se presenta a continuación (ver cuadro 2.3):

Ejemplo:

1: A ="La Lógica es un formalismo de representación"

2: $A \rightarrow B$ ="Si la Lógica es un formalismo de representación **entonces** la Lógica de Proposiciones de Primer Orden también lo es"

3: B ="La Lógica de Proposiciones de Primer Orden es un formalismo de representación"

1: A (es verdad)

2: $A \rightarrow B$ (es verdad)

3: B (es verdad)

A	B	$A \rightarrow B$
Verdad	Verdad	Verdad
Verdad	Falso	Falso
Falso	Verdad	Verdad
Falso	Falso	Verdad

Cuadro 2.3: Tabla de Verdad

Como se puede comprobar en el cuadro, si a) la lógica es un formalismo de representación; y b) la lógica de proposiciones de primer orden es un formalismo de representación, entonces A implica B. Es decir, si A es cierto entonces B también lo es. Si A es cierto y B no, entonces la implicación es falsa. Sin embargo, si A es falso y B es verdadero la implicación sigue siendo cierta. Y si A es falso y B es falso la

implicación sigue siendo cierta.

La Lógica nos proporciona un conjunto de símbolos y reglas de inferencia para expresar razonamientos válidos, independientemente de la interpretación semántica que tengan dichos símbolos en el mundo real. Existen extensiones de la Lógica Clásica que intentan cubrir una necesidad de aplicación específica y que amplían las capacidades de representación e inferencia de la misma. Entre estas tenemos:

- *Lógica Difusa o Borrosa*: Surge como necesidad de aplicar reglas lógicas a situaciones en las que los umbrales de decisión no son discretos, con valores de verdad-falso, sino continuos, difusos o borrosos. Es decir, cuando es preciso formalizar conocimiento incierto, aproximado o difuso.

Asimismo, resulta difícil localizar el momento exacto en el que se planteó la utilidad de este tipo de Lógica, aunque en 1965, el Catedrático Lotfi A. Zadeh en su artículo "*Fuzzy sets*", establece las bases de lo que hoy se conoce como *Lógica Difusa o Borrosa*. Es por ello, por lo que es conocido como el padre de *Lógica Difusa*.

En la realidad existen atributos que requieren valores continuos o grados de pertenencia μ a un grupo, como pueden ser conceptos tales como lleno, bello, grande, caro etc..

Los operados básicos de este tipo de lógica se establecen de la siguiente forma (ver figuras 2.4, 2.5 y 2.6):

1. Intersección: $\mu(A \text{ y } B) = \min\{\mu(A), \mu(B)\}$

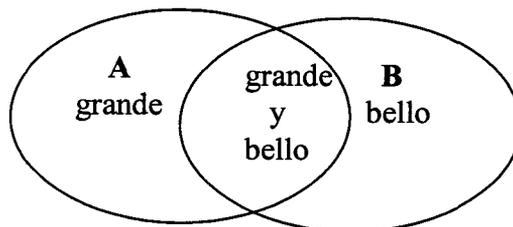


Figura 2.4: Representación gráfica del operador lógico Intersección

2. Unión: $\mu (A \text{ o } B) = \max\{ \mu (A), \mu (B)\}$

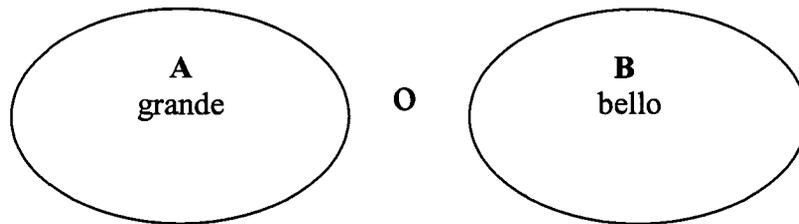


Figura 2.5: Representación gráfica del operador lógico Unión

3. Complemento o negación: $\mu (\text{no } A) = 1 - \mu (A)$

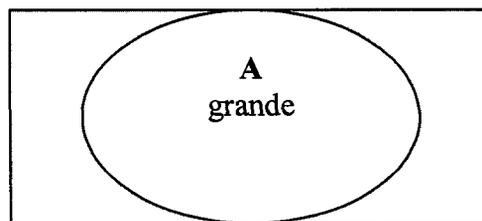


Figura 2.6: Representación gráfica del operador lógico Complemento

Este tipo de lógica ha abierto un mundo de posibles aplicaciones en entornos de toma de decisión basados en incertidumbre, como ocurre en la economía actual. Y permite el *razonamiento difuso* en el que una conclusión difusa o inexacta es inferida desde premisas difusas o inexactas.

Los *controladores difusos* representan una de las aplicaciones con más éxito de la Lógica Difusa. Básicamente, se ha aplicado esta técnica en sistemas de

control que están presentes en nuestra vida diaria, desde sistemas de control industrial hasta aplicaciones en *Business Intelligence*. Con los controladores difusos, las técnicas de la *Lógica Difusa* se utilizan para convertir heurísticas, e incluso reglas de control expresadas verbalmente, en estrategias de control automático.

Otra característica de la *Lógica Difusa* es la facilidad con la que se integra con otras tecnologías como las *Redes Neuronales Artificiales* o los *Algoritmos Genéticos*. Estas tendencias han dado como resultado un área de investigación de gran importancia y versatilidad llamada *Soft Computing*.

- *Lógica Temporal*: Este es un término utilizado ampliamente para hacer referencia a la representación de información temporal en un contexto lógico. Considerada una extensión de la *Lógica Proposicional Clásica*, fue propuesta inicialmente por Arthur Norman Prior (1914-69), reconocido como el fundador de la *Lógica Temporal*. Su uso ha servido para clarificar cuestiones filosóficas acerca del tiempo. Lo que ha permitido, como consecuencia, representar el cambio, las leyes causales, las relaciones entre espacios temporales, eventos, y conceptos tales como el presente, el pasado y el futuro.
- *Lógica Espacial*: Esta extensión de la *Lógica Clásica o Proposicional* nos permite representar e inferir conocimiento haciendo uso de reglas lógicas sobre las posiciones de los objetos en los espacios y las relaciones espaciales de los mismos. Se diferencia de la *Lógica Temporal* en dos características básicas:
 1. El tiempo es unidimensional mientras que el espacio es tridimensional.
 2. El tiempo tiene cierto grado de dirección mientras que el espacio no presenta esta distinción.

Algunas de las aplicaciones de la *Lógica Espacial* las podemos encontrar en consultas topológicas en bases de datos espaciales (Papadimitriou *et al.*, 1996), razonamiento espacial cualitativo (Cohn, 1995), formalización de conocimiento espacial (Kuipers, 1977), estrategias de búsquedas sobre ontologías espaciales

(Pais y Pinto-Ferreira, 1998), etc...

- *Lógica Causal*: El término Causal puede provocar controversia, ya que muchas personas piensan que este término es una invención del ser humano, y además, en muchas ocasiones, resulta bastante difícil establecer una relación entre la causa y el efecto. No obstante, es importante representar y razonar acerca de la causalidad de los hechos o eventos.

Otra forma de representar la causalidad está presente en las llamadas *redes causales, bayesianas o probabilísticas* que son representaciones gráficas de dependencias aplicadas al razonamiento probabilístico. Las *redes bayesianas* son una alternativa a la minería de datos, la cual tiene ventajas tales como:

- Permitir aprender sobre relaciones de dependencia y causalidad.
 - Permitir combinar conocimiento con datos.
 - Evitar el sobre-ajuste de los datos.
 - Permitir gestionar bases de datos incompletas.
- *Lógica Modal*: Esta extensión de la Lógica Clásica es también conocida como lógica de la posibilidad y la necesidad. Incorpora operadores modales que permiten inferir conocimiento como son \Box o \mathcal{L} (ver cuadro 2.4). Estos operadores operan en sentencias tales como:

$$\Box\Phi$$

$$\mathcal{L}(\Phi)$$

Por ejemplo, en una frase estos operadores pueden ser interpretados como “es sabido que Φ es verdad” o “ Φ es necesariamente cierto”. Así, la frase “Se sabe que la subasta tuvo éxito”, se expresaría como:

$$\mathcal{L}(\text{éxito}(\text{subasta}))$$

donde $\Phi = \text{éxito}(\text{subasta}) \rightarrow$ la subasta tuvo éxito

Existe un operador dual a \mathcal{L} que es \mathcal{M} o \diamond . Por lo que si se interpreta $\Box\Phi$ como “ Φ es necesariamente cierto”, entonces $\diamond\Phi$ puede interpretarse como “ Φ es posiblemente cierto”.

\mathcal{L}	\mathcal{M}
necesariamente	posiblemente
sabe/conoce	Cree
cree	no duda
cierto siempre	Cierto en ocasiones
obligatoriamente	Lícito

Cuadro 2.4: Posibles interpretaciones de los operadores modales.

A continuación, se presentan algunos ejemplos de formulas bien formadas.

- Es necesariamente cierto que todos los triángulos tienen tres lados.

$\mathcal{L}(\text{todos los triángulos tienen tres lados})$

- Es necesariamente cierto que si N es divisible por 8 entonces N es divisible por 4.

$\mathcal{L}(\text{si N es divisible por 8} \rightarrow \text{N es divisible por 4})$

La *Lógica Modal* tiene una de sus aplicaciones en los modelos de creencias. Y resulta de gran utilidad en sistemas de explicación o respuesta o planificación cuando se requiere la representación o inferencia de conocimiento basado en creencias, o en ambientes de movilidad (Gardelli y Gordon, 2000). También ha sido utilizada en sistemas en los que se aplica inferencia sobre conocimiento temporal o replicas de razonamiento legal.

- *Lógica No Monótona*: Esta extensión de la *Lógica Clásica* surge como necesidad, por ejemplo, de representar e inferir conocimiento basado en el **sentido común**, donde la consistencia de la veracidad o falsedad en un modelo depende de la temporalidad o la racionalidad limitada de los agentes en dicho dominio. Este aspecto, no es posible describirlo a través de la *Lógica Clásica*, es por ello que se propone una formulación menos rígida del problema como consecuencia, de las necesidades iniciales del área conocimiento. Para solventar esto, se crean reglas o leyes que permitan un tipo de razonamiento llamado razonamiento por defecto (default reasoning).

Cabe destacar los trabajos de McCarthy (1980) sobre Circunscripción como un método formal para representar razonamiento por defecto, haciendo uso de Cálculo de Predicados.

2.3.2.- Representación Procedural

Esta perspectiva contrasta con los métodos de representación declarativas propuestos en la Representación Lógica o en la Representación basadas en redes.

En la Representación Declarativa la representación del conocimiento es estática, es decir, se describen los objetos, los estados etc., pero se requiere de un motor de inferencia o programa que infiera ese conocimiento estático. En el caso de la Representación Procedural, la información necesaria para gestionar el conocimiento está implícita en sí mismo aunque requiere de un interprete para seguir las

instrucciones especificadas en la base de conocimiento.

Por consiguiente, en la Representación Procedural, el conocimiento se representa como un conjunto de instrucciones para solventar un problema. Este tipo de representación la asocian algunos investigadores con el método de resolución inteligente de problemas que aplicamos los seres humanos, basados en regularidades y en la ejecución de reglas (Newell y Simon, 1972; 1961). Como consecuencia de lo anterior, este tipo de representación es aplicable en áreas de conocimiento específicas por lo que se pueden utilizar Heurísticas (o conocimiento específico del dominio), y extensiones de la Lógica Clásica. Pero, este hecho conlleva el no poder cubrir todo el conocimiento posible del dominio en cuestión, solamente el explícito, y por lo tanto, los razonamientos o inferencias no cubren todas las posibles situaciones. Además, cualquier cambio en la base de conocimiento puede afectar a la totalidad de la base.

Los Sistemas basados en Reglas, tienen su fundamento en la idea de que el experto realiza la resolución de problemas en forma “situación–actual”. Existen dos tipos de Sistemas basados en Reglas, los sistemas de inferencia hacia adelante (*forward chaining*), y los sistemas de inferencia hacia atrás (*backward chaining*).

Los sistemas de inferencia hacia adelante encuentran la solución o la conclusión desde las premisas iniciales. Este tipo de sistemas son óptimos para aplicaciones de diagnóstico, donde la conclusión depende de consideraciones iniciales. De esta forma sus aplicaciones se circunscriben a contextos donde el número de posibles soluciones es limitado. En caso contrario, el sistema tiende a ser ineficiente y lento en su uso.

Los sistemas de inferencia hacia atrás son convenientes en situaciones en las que la población de soluciones es alta, y desde la solución, se busca las premisas que cumplen los criterios descriptivos.

Una de las tareas más importantes que se encuentran en estos sistemas son los procesos de búsqueda. Es por ello, por lo que las heurísticas que dirigen las búsquedas hacia ramas específicas del árbol de conocimiento resultan fundamentales.

El conocimiento incierto o difuso también tiene cabida en este tipo de sistemas y en especial en las reglas a través de valores continuos de certidumbre (ej., con certidumbre de 0.6).

2.3.3.- Representación en Redes

El conocimiento se presenta como un gráfico en el que los nodos representan objetos o conceptos en el dominio del problema, y los arcos representan las relaciones o asociaciones entre ellos.

- *Grafos Conceptuales:* son estructuras de conceptos y relaciones conceptuales, donde cada arco enlaza un nodo con otro nodo conceptualmente ligado. Las sintaxis abstractas de las estructuras y limitaciones de los Grafos Conceptuales, sin comprometer ninguna implementación o notación concreta. Para ilustrar lo anterior, a continuación se presenta, de forma gráfica, la siguiente frase "Esteban va a Londres en avión"

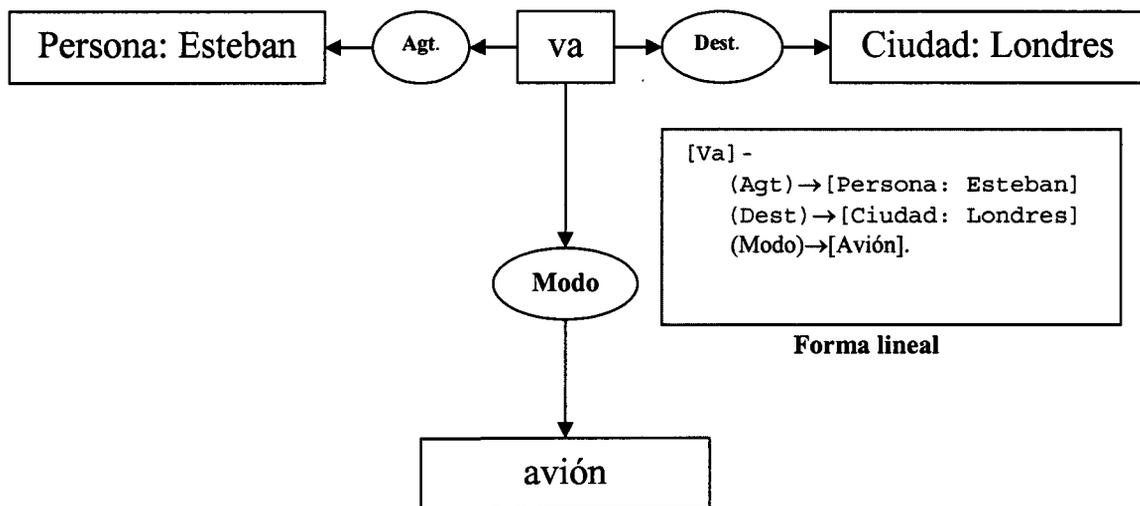


Figura 2.7: Forma gráfica de un ejemplo de Grafo Conceptual

En este ejemplo de la figura 2.7, los conceptos son representados por rectángulos [Persona: Esteban], [Ciudad: Londres], [va] y [avión]. Las relaciones

conceptuales están representadas por los óvalos Agente (Agt.), Destino (Dest.) y Caso (Caso). La Forma Gráfica y la Forma Lineal son tipos de representación válidas para la comunicación humano-humano y humano-máquina. En el caso del intercambio de Grafos Conceptuales entre ordenadores existe un formato llamado *conceptual graph interchange form* (CGIF), que se transcribe, para el ejemplo anterior, de la siguiente forma:

[Va *x] (Agt ?x [Persona: Esteban]) (Dest ?x [Ciudad: Londres]) (Modo ?x [Avión])

Este formato, a su vez, puede ser transcrito a otro formato como puede ser Knowledge Interchange Format (KIF) (Genesereth y Fikes, 1992). (ver cuadro 2.5)

```
(exists ((?x Va) (?y Persona) (?z Ciudad) (?w Avión))
  (and (Name ?y Esteban) (Name ?z Londres)
    (Agt ?x ?y) (Dest ?x ?z) (Modo ?x ?w)))
```

Cuadro 2.5: Representación del grafo ejemplo en KIF

El KIF es un lenguaje orientado a los ordenadores para el intercambio de conocimiento entre programas informáticos. Posee algunas características de interés, tales como semánticas declarativas, posibilidad de comprensión lógica, representación de meta-conocimiento, representación de reglas de representación no monótona y definiciones de objetos, funciones y relaciones.

A raíz de los trabajos de investigación relacionados con la representación de conocimiento basado en redes, ha surgido un nuevo paradigma denominado *Topic Maps*, descrito por primera vez y de forma completa, en el estándar ISO/ES 13250:2000 "*Topic Maps*". Este estándar está basado en el lenguaje SGML y HyTime. Esta especificación suministra un modelo y una gramática para representar la estructura de recursos de información utilizados para definir tópicos y asociaciones entre tópicos. Los nombres, los recursos y relaciones representan las características de objetos abstractos llamados tópicos. Este tipo de estándar, está siendo

ampliamente aplicado a la Web, a través del estándar XTM (XML *Topic Maps*) (Topic Maps.org, 2001).

2.3.4.- Representaciones estructuradas

Los lenguajes de representación estructurada extienden las redes permitiendo a cada nodo convertirse en estructura con propiedades con valores.

- *Marcos o Frames*. Una de estas formas de representación son los Marcos o *Frames*, representación propuesta por Minsky (1974; 1975; 1980), como un mecanismo para representar el conocimiento de un dominio. Un *Frame* se define como una estructura de datos que representa un concepto o situación y que como puede ser, por ejemplo, “país”, “región” etc...;e incorpora diferentes tipos de información, que pueden ser descriptivas, de definición o de instrucción sobre el uso del *Frame*. A partir de la propuesta inicial de Minsky, este esquema de representación ha evolucionado hacia los Grafos Conceptuales y otros formalismos como la *Lógica Modal*.

Las aplicaciones de los sistemas basados en *Frames* son óptimas para reconocimiento o inferencias sobre estructuras. Un ejemplo de ello lo tenemos en el análisis de documentos científicos (Cattoni *et al.*, 1998)

- *Scripts* o Guiones: Otro tipo de representación basada en estructuras se encuentra en los llamados *Scripts*, caracterizados por la similitud que tiene con los *Frames*, ya que permite la representación de una situación prototípica, aunque con una diferencia esencial, en que tiene como finalidad la descripción de una secuencia estereotipadas de eventos o acciones. El *Script* posee un conjunto de *slots* con información acerca de personas, objetos y acciones involucrados en los eventos. Algunos de sus elementos típicos incluyen:
 - *Condiciones (entry conditions)*: describen lo que se debe de satisfacer para poder aplicar el *Script*.

- *Objetos (props)*: incluyen los objetos que se usan en la secuencia de eventos.
- *Roles*: el/los agentes involucrados.
- *Variantes (track)*: variantes que pueden ocurrir en un determinado *Script*.
- *Escenas*: descripción de la secuencia de eventos.
- *Resultados*: las condiciones que existen después de aplicar el *Script*

Un ejemplo de *Script* se puede observar en el cuadro 2.6: *Script de un restaurante japonés*, es decir, un cliente en un restaurante japonés. El *Script* está compuesto de los elementos descritos anteriormente, como pueden ser nombre del *Script*, la variante, los roles (cliente, camarero y cajero), los objetos (caja, menú, comida, dinero, servilletas, sal/salsa), las condiciones (el cliente tiene hambre y dinero), las posibles escenas (entrada, solicita, comer etc...) y los posibles resultados.

La descripción de eventos tipificados por apartados como es el caso del ejemplo, permite realizar un razonamiento basado en casos y la inferencia por reconocimiento de patrones. Una de las aplicaciones más típicas de los Scripts es la "comprensión" del lenguaje natural. La identificación parcial a través de un caso incompleto de ejemplo con un Script, puede provocar el reconocimiento del resto de la información no suministrada o la consulta estructurada del resto de la información necesaria. Esta ventaja puede ser muy importante en algunas situaciones prácticas en las que la información suministrada sea insuficiente y se requiera un ejercicio de reconocimiento. Otra área muy significativa de aplicación de los Scripts es el razonamiento por ejemplos (Frank y Foley, 1994)

Script Restaurante Japonés

Variante (track): comida japonesa

Roles: cliente, camarero, cajero

Objetos (props): caja, menú, comida, dinero, servilletas, sal/salsa

Condiciones: el cliente tiene hambre y dinero

- Escena1: entrada
 - el cliente estaciona su vehículo
 - el cliente entra al restaurante
 - el cliente espera a que lo atiendan
 - el camarero lleva al cliente a una mesa desocupada
- Escena2: solicita
 - el camarero trae la carta al cliente
 - el cliente escoge que comer de la carta
 - el cliente indica la selección al camarero
- Escena3: comer
 - el camarero trae la comida
 - el cliente come su comida
- Escena4: salida
 - el cliente pide la cuenta
 - el camarero le trae la cuenta
 - el cliente va al cajero y paga su cuenta
 - el cliente sale del restaurante
 - el cliente se va en su vehículo
- Escena4a: comida-corrída (opcional)
 - el cliente sale corriendo del restaurante
 - el cliente se va en su vehículo rechinando llanta

Resultados:

- el cliente ya no tiene hambre
- el cliente tiene menos dinero
- el cliente esta satisfecho (?)
- el cliente no esta satisfecho (?)
- el cliente esta demasiado lleno (?)
- el cliente tiene dolor de estómago (?)

Cuadro 2.6: Ejemplo de Script de un restaurante japonés.

- **Casos:** Es el elemento básico de los Sistemas Basados en Casos (Case Based Systems). Para estos sistemas, un Caso es un problema o situación que ha sido experimentado o resuelto por el sistema con anterioridad, o un nuevo problema a solucionar. Históricamente y filosóficamente fue una reacción a los *Sistemas Basados en Reglas* y tienen parte de su fundamento en trabajos de investigación sobre el comportamiento psicológico de los seres humanos, en la resolución de problemas (Ross, 1989; Anderson, 1983; Kolodner, 1983).

Estos casos se almacenan con información sobre sí mismo, que permita una futura identificación.

Algunas aplicaciones de los Sistemas Basados en Casos se pueden encontrar en las organizaciones (Purvis *et al.*, 2000), en la Gestión del Conocimiento (Watson, 2002), en la toma de decisiones cualitativas (Slade *et al.*, 1994) o en entornos de educación virtual (Regan y Slator, 2002).

2.3.5.- Representación Simbólica Híbrida

La combinación de varios métodos de representación dirigida a la resolución de un problema es característica de los sistemas híbridos. Estos sistemas pueden clasificarse en dos tipos:

- **Sistema híbridos externos:** cuando el sistema está organizado por módulos independientes con métodos de representación desiguales, y la comunicación entre estos, se realiza a través de canales o variables compatibles entre ambos módulos.
- **Sistemas híbridos internos:** cuando el sistema incorpora, dentro del mismo módulo, diferentes métodos de representación como pueden ser reglas y marcos.

Los tipos de representación simbólica híbrida normalmente suelen incluir las técnicas de representación en redes y las técnicas de representación en estructuras. Por

ejemplo, combinaciones tales como: grafos con casos o grafos con marcos. E incluso, representaciones híbridas dentro de un mismo grupo de técnicas, como es el caso de los casos con los *scripts* o con las técnicas de marcos.

Una vez introducida la representación del conocimiento basada en símbolos, a continuación se introduce la representación no-simbólica del conocimiento

2.4. - REPRESENTACIÓN NO-SIMBÓLICA DEL CONOCIMIENTO

Algunos ejemplos de conocimiento tácito pueden ser simulados con sistemas no simbólicos como pueden ser las Redes Neuronales y los Sistemas Evolutivos, entre otros. El cuadro 2.7 ilustra los tipos de representación no simbólica y las técnicas de representación respectivas. Entre los tipos de representación no simbólica se incluyen: las Redes Neuronales, los Sistemas Evolutivos y la Representación Híbrida. En relación a las técnicas de representación, se presentan las Redes Asociativas, las Redes de Hopfield, las Redes Estocásticas y las Redes de Kohonen, utilizadas para la representación de redes neuronales. Además, se integran en el cuadro los algoritmos genéticos, la programación genética, las estrategias evolutivas, la programación evolutiva y los sistemas clasificadores, que componen el grupo de técnicas utilizadas para la representación de sistemas evolutivos. Y, por último, las redes neuronales difusas, redes evolutivas y sistemas difusos evolutivos, representan el grupo de técnicas utilizadas para la representación híbrida.

A continuación, se describen cada uno de los tipos de representación no simbólica, arriba enumeradas, y sus respectivas técnicas de representación. Iniciamos entonces este apartado introduciendo el concepto de redes neuronales.

TIPO DE REPRESENTACIÓN NO SIMBÓLICA	TÉCNICA
Redes Neuronales	<ul style="list-style-type: none"> • Redes Asociativas • Redes de Hopfield • Redes Estocásticas (Máquinas de Boltzman) • Redes de Kohonen
Sistemas evolutivos	<ul style="list-style-type: none"> • Los Algoritmos Genéticos. • La Programación Genética. • Las Estrategias Evolutivas. • La Programación Evolutiva. • Los Sistemas Clasificadores.
Representación híbrida	<ul style="list-style-type: none"> • Redes Neuronales Difusas • Redes Evolutivas • Sistemas Difusos Evolutivos

Cuadro 2.7: Técnicas de representación no simbólica del conocimiento.

2.4.1.- Redes Neuronales

No existe una definición universalmente aceptada de Red Neuronal, aunque se podría describir como un modelo matemático compuesto por una red de procesadores

numéricos elementales (llamados nodos), que suelen contener una pequeña cantidad de memoria y que se conectan entre sí mediante canales llamados *conectores* que suelen transmitir, en la mayoría de los casos, datos codificados.

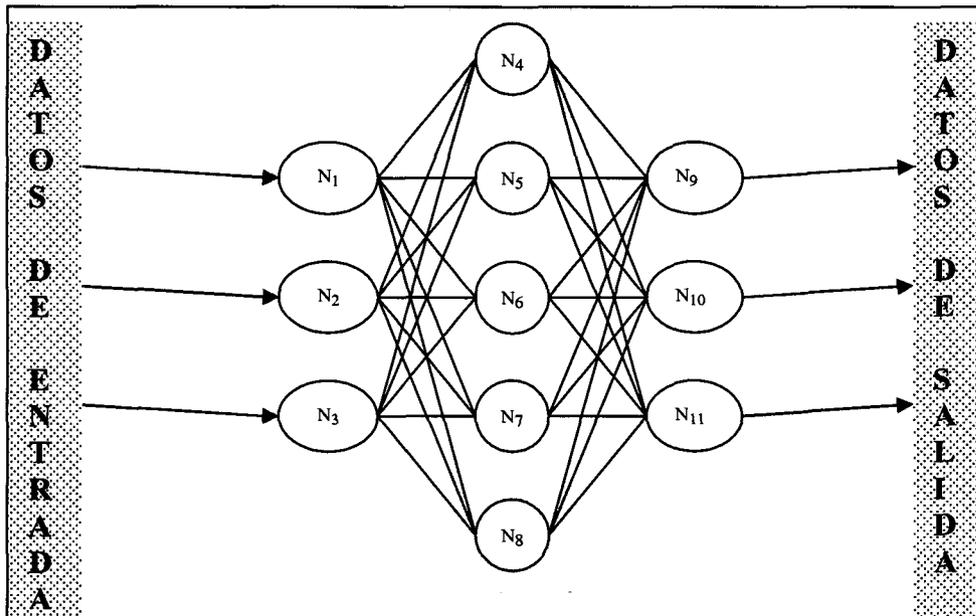
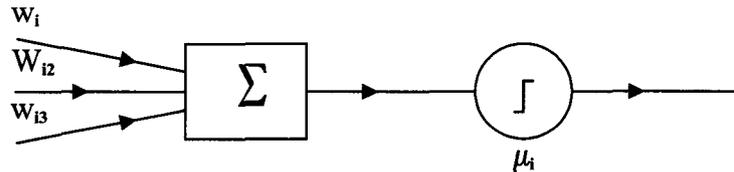


Figura 2.8: Representación gráfica de un ejemplo de Red Neuronal Artificial

La figura 2.8 contiene una representación gráfica de un ejemplo de Red Neuronal Artificial donde tenemos una entrada de datos y una salida de datos. Está compuesta por once nodos interconectados entre sí a través de conectores. Del nodo N1 al N3 encontramos la capa de entrada de datos, que son los encargados de recibir los datos cualificados del exterior de la red. Del nodo N4 al N8, nos encontramos con los nodos de la capa intermedia, también llamada oculta, cuyos nodos se encuentran interconectados con los nodos de la capa de entrada y salida de datos. Y, por último, del nodo N9 a N11 tenemos la capa de salida, que son los encargados de transmitir al exterior los datos codificados de entrada que han sido procesados por la red.

Un modelo inicial de Red Neuronal fue el propuesto por McCulloch y Pitts (1943) (ver figura 2.9). Muchos de estos modelos se han inspirado en el sistema neurofisiológico

humano, compuesto por 10^{11} células nerviosas de muchos tipos, llamadas *neuronas*. La capacidad de procesamiento en paralelo de tantos procesadores existentes en el cerebro humano ha sido uno de los mayores atractivos para los investigadores.



$$n_i(t+1) = \theta\left(\sum_j w_{ij}n_j(t) - \mu_i\right) \quad \Theta(x) = \begin{cases} 1 \rightarrow \text{si } \dots x \geq 0; \\ 0 \rightarrow \text{el } \dots \text{resto} \end{cases}$$

Figura 2.9: Diagrama esquemático y ecuaciones de la neurona de McCulloch-Pitts (1943)

Las Redes Neuronales son también conocidas como *Redes Neuronales Artificiales*, para diferenciar a éstas, de las *Redes Neuronales Biológicas*, que son mucho más complejas que las Artificiales.

La aplicación de estos modelos es muy variable. En principio, una *Red Neuronal Artificial* puede procesar cualquier función que pueda ser programada por ordenador, haciendo uso de cualquier lenguaje de programación (Valiant, 1988; Siegelmann y Sontag, 1999; Orponen, 2000; Sima y Orponen, 2001), e incluso aplicaciones en que la su informatización sea dudosa (Siegelmann, 1998 y Hadley, 1999). No obstante, se ha encontrado importantes aplicaciones de las mismas, especialmente en problemas de clasificación y funciones de aproximación o *mapping*, donde se requiere niveles de tolerancia a la imprecisión, existiendo datos muestrales en abundancia y donde el diseño de reglas rápidas o complejas no pueda llevarse a cabo con facilidad.

Una de las tareas más importantes a las que se enfrenta un investigador en el momento de hacer uso de una de estas técnicas es el poder interpretar el problema desde la perspectiva más rentable de la técnica. Es decir, saber elegir y aplicar el

modelo correcto al tipo de problema.

Existen muchos tipos de redes neuronales artificiales. Podríamos decir que nadie conoce exactamente cuántos tipos existen. Sin embargo, los dos grupos más importantes de algoritmos de aprendizaje están clasificados en supervisados y no supervisados. En el aprendizaje supervisado, los resultados correctos se dividen a su vez en dos grupos: los valores objetivo y las salidas deseadas. Éstos son conocidos y aportados a la red neuronal durante un proceso de entrenamiento. De tal forma que dicha red pueda ajustar sus pesos, para intentar equiparar las salidas a los valores deseados. Después del entrenamiento, la red neuronal es comprobada, dándole sólo valores de entrada y no valores de salida. Y, de esta forma, observar qué nivel de corrección tienen los valores de salida en función de los que ya se conocen. Con respecto al aprendizaje no supervisado, a la red neuronal no se le suministra los resultados correctos durante el entrenamiento. Este tipo de redes neuronales no supervisadas, por lo general realizan algún tipo de comprensión de datos tal como: reducción de la dimensionalidad o clustering.

En función a la topología de la red, existen dos tipos principales de redes neuronales: *feedforward* y *feedback*. En las primeras, alimentación hacia delante, las conexiones entre los nodos no forman ciclos. Este tipo de redes normalmente producen una respuesta rápida. La mayoría de estas redes pueden ser entrenadas usando una amplia variedad de métodos numéricos, convencionales y muy eficientes. Además de los algoritmos inventados por los investigadores sobre las redes neuronales.

Por otro lado, en las redes neuronales de retroalimentación o recurrentes, las conexiones entre los nodos forman ciclos. En algunas de estas redes, cada vez que se produce una entrada de datos, la red neuronal realiza varios ciclos tomando la salida como entrada, en cada vuelta. Como consecuencia de esto, suelen tener un tiempo de respuesta más largo y suelen ser más difíciles de entrenar que las redes *feedforward*.

Las redes neuronales a su vez pueden ser clasificadas en función al tipo de datos que aceptan, existiendo dos tipos de datos básicos: los categóricos y los cuantitativos. Dentro del primer grupo, es decir las redes neuronales que aceptan variables categóricas, sólo pueden tomar un número finito de posibles valores. Y, normalmente,

suelen existir más de un caso en cada categoría. Las variables categóricas pueden tener valores simbólicos tales como: grande-pequeño, pobre-rico,...etc, que tienen que ser codificados en números antes de ser pasados a la red. El aprendizaje supervisado con valores categóricos, y los no supervisados con salidas categóricas, se denomina clasificación.

Las variables cuantitativas representan medidas numéricas de algunos atributos, tales como la distancia en kilómetros. Las medidas deben ser realizadas de tal forma que al menos algunas relaciones aritméticas de las medidas reflejan relaciones análogas entre los atributos de los objetos que se miden. El aprendizaje supervisado con los valores objetivo cuantitativos se denominan regresión.

- **Redes Asociativas:** La red asociativa es una red que toma un vector de entrada y devuelve otro vector de salida. Existen dos tipos de redes asociativas: las hetero-asociativas y las auto asociativas. Las hetero-asociativas toman un vector de entrada y devuelven un vector completamente diferente. Por ejemplo, una red hetero-asociativa podría ser utilizada para tomar un fichero de sonidos y devolver un texto que lo represente. Las auto asociativas, por su parte, toman un vector de entrada y devuelve el mismo vector de salida. Este tipo de redes suelen aprender varios vectores de tal forma que cuando un vector corrupto le es asignado, dicha red puede corregirlo y devolver el vector correcto. Asimismo, las redes auto asociativas son utilizadas para el reconocimiento parcial o corrupto de patrones (Matthews, 1999).
- **Redes de Hopfield:** la red de Hopfield es una red auto asociativa, utilizada con frecuencia para el reconocimiento mono cromático de imágenes. Fue introducida en 1982 por J.J. Hopfield, como un nuevo concepto en el que los datos de salida eran retroalimentados a las entradas. Este tipo de redes están compuestas de una capa cuyas unidades de salida retroalimentan todas las otras entradas por sí mismas.
- **Redes Estocásticas:** estas redes son una extensión de los modelos neuronales básicos, las cuales tienen función de activación no determinista, y su aprendizaje es también estocástico, haciendo una selección de pesos de forma

aleatoria y comprobando su efecto en la salida de datos. En estas redes también se busca obtener un nivel mínimo de error y para ello se recurre a un símil termo-dinámico, tomando la red como un material formado por partículas, las cuales tienen un nivel de energía probabilística y la energía global del sistema depende de la energía individual de cada partícula y de su temperatura, buscando siempre el estado de mínima energía. En este tipo de redes neuronales, cada neurona funcionará como una partícula y la red completa el material. El mecanismo que permite la obtención del mínimo error se conoce con el nombre de temple simulado, en el cual se eleva la energía del sistema (*simulated annealing*) y se baja progresivamente, haciendo que la red caiga en el mínimo global y evitando los mínimos locales, suministrando algunas veces energía para salir de ello.

- *Redes de Kohonen*: Las redes de Kohonen, son también conocidas como mapas auto organizados de Kohonen. Representan un tipo de red auto organizadas, no supervisadas. Como ya se mencionó anteriormente, el aprendizaje no supervisado no requiere de un proceso de entrenamiento o maestro externo que indique a la red que está operando de forma correcta o incorrecta. Por lo que la red se auto organiza descubriendo por sí misma los rasgos comunes, regularidades, correlaciones o categorías en los datos de entrada. Se argumenta en este caso, que las neuronas deben auto organizarse en función de los estímulos procedentes del exterior.

2.4.2.- Sistemas Evolutivos

La teoría de la evolución natural propuesta por Darwin (1859) y los trabajos del monje Gregor Mendel reconocido como el padre de la genética actual, promueven una nueva directriz de investigación en la informática y en la resolución de problemas. La base de esta propuesta, se basa en la búsqueda aleatoria de soluciones desde poblaciones de soluciones representadas por estructuras de datos, basadas en el ADN, que evoluciona de generación en generación (herencia). Esto ha permitido el desarrollo de técnicas de optimización que equilibran la exploración y la explotación de los espacios

de búsqueda. Los algoritmos evolutivos comprenden un conjunto de tecnologías entre las que se incluyen: los *Algoritmos Genéticos*, la *Programación Genética*, las *Estrategias Evolutivas*, la *Programación Evolutiva* y los *Sistemas Clasificadores*, entre otros.

En la figura 2.10 se representa cómo podría ser un espacio de soluciones y los puntos azules representan los máximos u objetivos de la búsqueda.

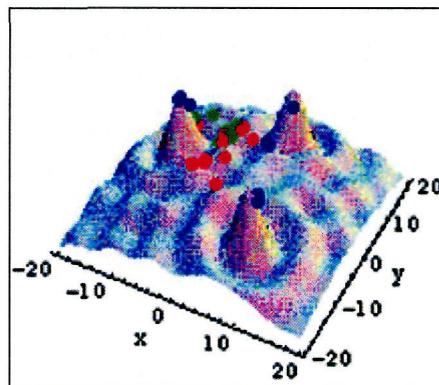


Figura 2.10: Representación gráfica del espacio de soluciones donde se realizan las búsquedas (Fuente: <http://pages.cpsc.ucalgary.ca/~jacob/>)

Todas las tecnologías anteriores, comparten una base conceptual de simulación de la evolución natural de Estructuras Individuales, pertenecientes a Poblaciones de Estructuras a través reglas de selección, y operadores de Mutación y Reproducción. Por lo general, la selección viene definida por una función de rendimiento o supervivencia, de cada Estructura Individual en la Población.

Secuencia de búsqueda*Establecer el contador de tiempo a cero**Iniciar una población de soluciones generalmente de forma aleatoria**Evaluar el rendimiento de cada uno de los individuos iniciales de la población**Comprobar el fin de la búsqueda**(por obtener un rendimiento deseado o por el n° de generaciones)**Incrementar el contador de tiempo**Seleccionar una subpoblación de soluciones (al azar)**Realizar una reproducción entre las soluciones seleccionadas**Realizar una mutación sobre las soluciones obtenidas del paso anterior**Evaluar el rendimiento de las nuevas soluciones**Selección natural sobre las soluciones de la población incluyendo las nuevas soluciones**Ir a Comprobar**Fin de la búsqueda***Cuadro 2.8:** Método evolutivo generalizado

En el cuadro 2.8., se presenta el método evolutivo generalizado como secuencia de búsqueda. Como se puede observar, se establece inicialmente el contador de tiempo a cero. Seguidamente, se genera una población de soluciones de forma estocástica, que permita la distribución de las soluciones iniciales en todo el espacio de búsqueda. A continuación, se evalúa el nivel de eficiencia de cada una de las soluciones iniciales y entra en un ciclo representativo que terminará, o por llegar al número de generaciones límites, o por obtener la/s soluciones deseadas. Dentro del bucle de búsqueda, los pasos que se realizan son los siguientes: se incrementa el contador de tiempo; se realiza una selección al azar de una subpoblación de soluciones. Se combinan dichas Soluciones entre sí, creando un nuevo conjunto de soluciones combinadas o modificadas. Estas soluciones combinadas son mutadas y, posteriormente evaluadas, con el fin de realizar una selección natural sobre el conjunto total de la población, incluyendo las nuevas soluciones.

Por lo tanto, el objetivo de los algoritmos evolutivos no es más que la búsqueda de una o varias *soluciones deseadas*. Lo importante pues, es definir qué se entiende por *solución* (modelo de respuesta) y qué se entiende por *solución o soluciones deseadas* (modelo del objetivo). Por ejemplo, una solución puede indicar el camino más corto de un vendedor (TSP - Travel Salesman Problem) (Grefenstette *et al.*, 1985; Beyer, 1992; Boseniuk, 1990), que tiene que cubrir una ruta de visitas, el diseño de una campaña de marketing (Eiben, Euverman *et al.*, 1996; Bhattacharyya, 2000), con el máximo impacto a un menor coste, o el plan de control de optimización de inventario (Eriksson R. y Olsson B., 1997).

Muchas son las áreas de aplicación de los Sistemas Evolutivos, si bien cabe destacar las relacionadas con problemas de optimización, adaptación y aprendizaje.

2.4.3.- Representación No-Simbólica Híbrida

Cada técnica de representación ofrece ciertas ventajas, como por ejemplo las redes neuronales, el reconocimiento de patrones o clustering, o la lógica difusa, o la inferencia de conocimiento de impreciso. Pero, al mismo tiempo, cada una de estas técnicas posee debilidades o desventajas. Es por ello, que en muchas ocasiones, y en función del objetivo a cubrir, se busca la combinación de varias técnicas con el fin de conseguir el objetivo marcado o cumplir los requerimientos del sistema, como es el caso de los sistemas híbridos. Entre las combinaciones de las representaciones no simbólicas más típicas se encuentran:

- **Redes Neuronales Difusas:** En este caso, las redes neuronales reemplazan los valores discretos de entrada en la capa de entrada de datos, por valores difusos y la aritmética que se realiza en los nodos es una aritmética difusa y los valores de los pesos también pueden ser difusos. Las aplicaciones de este tipo de modelos son muy variadas, clasificadores difusos (Tschichold-Gürman, 1994), clustering difuso (Kulkarni y Muniganti, 1996; Wu y Lu, 2000), predicción de series temporales (Maguire *et al.*, 1998), e incluso, correo electrónico activo

(Kaurnoskos y Vasilakos, 2002).

- *Redes Evolutivas*: Los algoritmos evolutivos pueden ser utilizados para evitar dos tareas difíciles de realizar como son el entrenamiento y el diseño de las redes neuronales. Existe otra aplicación muy significativa de los algoritmos evolutivos y las redes neuronales y es crear un proceso evolutivo artificial de redes neuronales. En este caso, también, las aplicaciones de este tipo de técnicas es muy variada, como puede ser la navegación en espacios desconocidos a través de sistemas multi-agentes basados en redes neuronales evolutivas (Wang y Mckenzie, 1999) o la evolución de criaturas virtuales (Sims, 1994).
- *Sistemas Difusos Evolutivos*: Los sistemas difusos requieren de una selección de parámetros y decisiones de diseño que son bastante importantes y tediosas, y que en muchas ocasiones se parece a las decisiones que se deben tomar en el caso de las redes neuronales. Es esta la principal razón, por la que los sistemas evolutivos pueden ser utilizados para automatizar y optimizar la selección de estos parámetros y decisiones en el diseño de los sistemas difusos. Un caso típico de este tipo de tecnologías aparece en el diseño de controladores difusos (Chiu y Hunter, 1999).

2.5.- INFERENCIA DEL CONOCIMIENTO

El método de inferencia de conocimiento viene, en la mayoría de los casos, establecido por la técnica de representación de conocimiento aplicada en el sistema. Si bien, cabe resaltar algunos tipos de inferencias o similares que son característicos y diferenciales. Entre estos se encuentran los sistemas de búsqueda, y la inferencia distribuida, entre otros.

2.5.1.- Sistemas de Búsqueda

La resolución de problemas o la realización de procesos basados en tareas pueden, en muchas ocasiones, ser interpretados como una búsqueda de solución. Pero el modelado de estas soluciones requiere, en la mayoría, de los casos la consideración de aspectos tan importantes como pueden ser:

- El espacio de búsqueda y su tamaño
- El modelo del problema
- El posible cambio del espacio o la solución a través del tiempo.
- Las restricciones del problema y la solución o soluciones.
- La tipología del problema (encontrar soluciones o probar conclusiones)

Por todo ello, se requiere definir conceptos tales como:

1. El tipo de representación
2. El objetivo a alcanzar
3. La función de evaluación
4. La definición del problema de búsqueda
5. Puntos máximos o mínimos y vecindades (neighborhoods) del espacio de búsqueda.
6. Métodos de búsqueda local (tipo Hill-climbing)

Uno de los grandes problemas con los que se encuentran los investigadores en el momento de encontrar soluciones en espacios de búsqueda, es el tamaño del espacio de búsqueda, que en muchas ocasiones, hace casi imposible la tarea de búsqueda. Para solucionar este problema se aplica conocimiento (heurísticas) que permite dirigir de forma inteligente el proceso de búsqueda. Entre los métodos de búsqueda más utilizados, en función al tipo de aplicación o requerimiento, se pueden encontrar:

1. *Búsqueda exhaustiva* (no informada): Este tipo de método realiza una comprobación de cada una de las soluciones en el espacio de búsqueda hasta encontrar la mejor o la más óptima. Este tipo de búsqueda es aconsejable cuando los espacios de búsqueda no son demasiado grandes. Se podría decir que ha servido de base para el desarrollo de otros métodos.
2. *Búsqueda con soluciones incompletas*: En este caso, la solución deseada surge de la manipulación de soluciones parciales o incompletas, a través de pasos de construcción. El método más conocido de búsqueda con soluciones incompletas se denomina algoritmos *greedy*.
3. *Búsqueda local*: En contraposición al caso anterior, en muchas ocasiones resulta preciso concentrar la atención del proceso de búsqueda en zonas bien localizadas o vecindades en torno a una solución determinada. Esto se hace posible realizando pasos o recorridos muy cercanos a dicha solución o punto del espacio. Uno de los problemas que aparecen con este tipo de método es el hecho de que la búsqueda quede bloqueada en una solución óptimo local, sin que por ello dicha solución sea la mejor del espacio de búsqueda.
4. *Escapatoria de soluciones óptimas locales*: Para solucionar la debilidad del grupo anterior, se encuentran los métodos que permiten escapar de las soluciones óptimas locales mediante mecanismo como el propuesto en el algoritmo Simulated Annealing (enfriamiento simulado; Kirkpatrick 1983).

2.5.2.- Inferencia Distribuida

Son muchas las ocasiones en las que sería ideal el poder realizar inferencia distribuida. Las razones de esta decisión pueden ser:

- Los seres humanos utilizan un modelo psicológico distribuido para inferir conocimiento.
- Se puede realizar procesamiento distribuido, bien a través de las redes

informáticas como puede ser Internet, o bien a través de los sistemas de clustering de ordenadores para obtener procesamiento en paralelo.

- Muchas aplicaciones y procesos podrían mejorar considerablemente su rendimiento, por ejemplo los procesos de búsqueda.
- Se podría hacer un desarrollo de sistemas más modular.
- Se podrían utilizar diferentes técnicas de inferencia e incluso representación de conocimiento.
- Se podrían desarrollar sistemas de información más fiables.

2.5.3.- Otros tipos de Inferencias

Son muchos las formas de inferir conocimiento almacenado en soportes magnéticos si bien cabe destacar, el razonamiento causal y sus aplicaciones (Hudlicka y Corker, 1988; Wilkins 1988), el razonamiento cualitativo (Yuen, Ho y Zeleznikow, 1992; Kuipers, 1986, 1994; Lee *et al.*, 1987), y el razonamiento basado en casos (CBR-Case Based Reasoning; Stepankova, 1992; Forbus, 1985; Aarts, 1998; Watson, 2000; Purvis *et al.*, 2000; Regan y Slator, 2002).

2.6.- Descubrir conocimiento.

En muchas ocasiones se puede entender la minería de datos, y el Knowledge Discovery (*descubrir conocimiento*) en bases de datos, como conceptos similares. La Minería de Datos puede entenderse como un paso esencial hacia el descubrimiento de conocimiento en bases de datos.

La Minería de Datos surge como resultado de la evolución de las Tecnologías de la Información, el incremento de volúmenes ingentes de datos y el interés que se tiene por convertir estos datos en información útil y conocimiento. Ya desde los años sesenta (ver figura 2.11), las tecnologías de la información y las bases de datos han ido evolucionando desde el procesamiento de ficheros primitivos, hasta sistemas de bases de datos sofisticadas y muy poderosas. La investigación y el desarrollo de los sistemas

de bases de datos desde los años setenta, ha progresado desde las bases de datos en red y jerárquicas, al desarrollo de bases de datos relacionales, herramientas para crear modelos de datos, técnicas de organización de datos e indexación. Además, los usuarios han mejorado los modos de acceso a través de lenguajes estructurados de consulta, los interfaces y la gestión de transacciones. Los métodos eficientes de procesamiento de transacciones en línea (OLTP – On-line transaction processing), donde las consultas son vistas como transacciones de sólo lectura, han contribuido de forma sustancial a la evolución y a la aceptación de tecnología relacional, como la mayor herramienta para el almacenamiento inteligente, la recuperación y la gestión de grandes volúmenes de datos.

Desde mediados de los años ochenta, las tecnologías de bases de datos se han caracterizado por la elección de tecnología relacional y un resurgimiento de actividades de investigación y desarrollo de nuevos y poderosos sistemas de bases de datos (ver figura 2.11). Estos sistemas, hacen uso de modelos avanzados, tales como relacional-extendido, orientado a objeto, relacional-objeto y modelos deductivos (Bostan y Yazici, 1998; Yiu-Kai y Seetharaman 1997), entre otros. Asimismo, en los últimos años, han aparecido sistemas de bases de datos orientados a aplicaciones específicas como pueden ser, bases de datos espaciales (El-Geresy y Abdelmoty, 2001; Xiaowei *et al.*, 1998), temporales (Kurutach, 1985; Koubarakis, 1993; Xiaowei *et al.*, 1998), multimedia (Grosky, 1997; Marcus y Subrahmanian, 1996), activas, científicas (Spars, 1982), del conocimiento (Houstis *et al.*, 2000) o información de oficina. También se han estudiado temas relacionados con la distribución, diversificación y compartir datos. Los sistemas de bases de datos heterogéneos y la WWW e Internet, surgen para jugar un papel muy importante en la industria de la información, es por ello que ha surgido el interés por la minería de datos en la Web (Berendt, Hotho y Stumme, 2002).

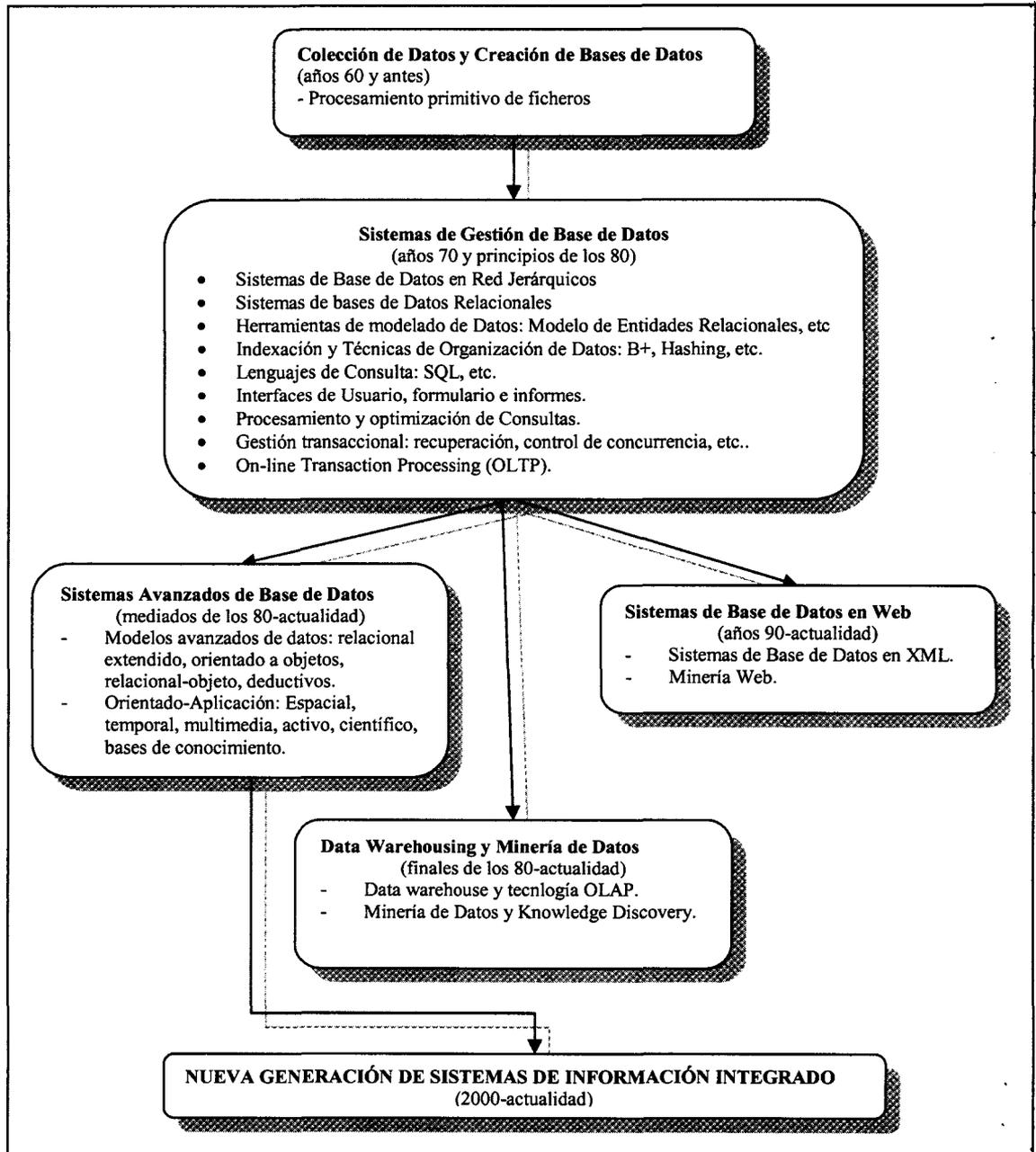


Figura 2.11: La evolución de la tecnología de la base de Datos

2.6.1.- Metodologías de aplicación

El descubrimiento de conocimiento es un proceso que puede ser dividido en los siguientes pasos (ver figura 2.13):

1. Depuración en el que se retiran datos inconsistentes o con ruido.
2. Integración de múltiples fuentes de datos para resultar en un *data warehouse* o base de datos.
3. Selección de los datos relevantes para la investigación y recuperación de la base de datos o data warehouse.
4. Transformación de los datos seleccionados en el paso anterior en formas apropiadas para la minería a través de operaciones tales como agregación.
5. Minería es la principal fase en la que los métodos avanzados se aplicados para conseguir la extracción de patrones de datos.
6. Evaluación de los patrones encontrados para identificar los que representan realmente conocimiento.
7. Presentación del conocimiento utilizando métodos de presentación del conocimiento descubierto con el fin de que sea entendido por los usuarios.

La figura 2.12 representa de forma gráfica un cubo de 4 dimensiones que puede representar datos en cualquier contexto.

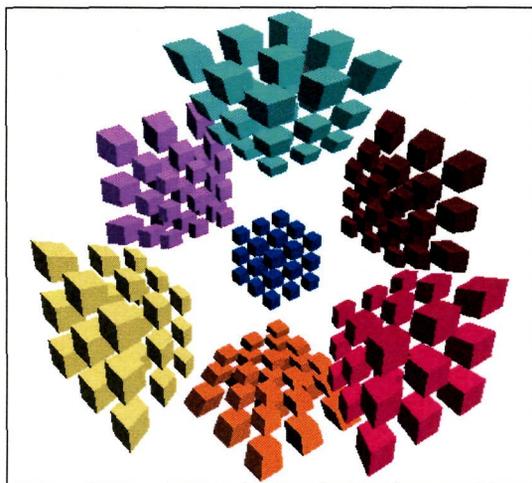


Figura 2.12: Representación gráfica de la simulación de un cubo de 4 dimensiones

La figura 2.13 ilustra el proceso de Knowledge Discovery desde bases de datos y ficheros planos.

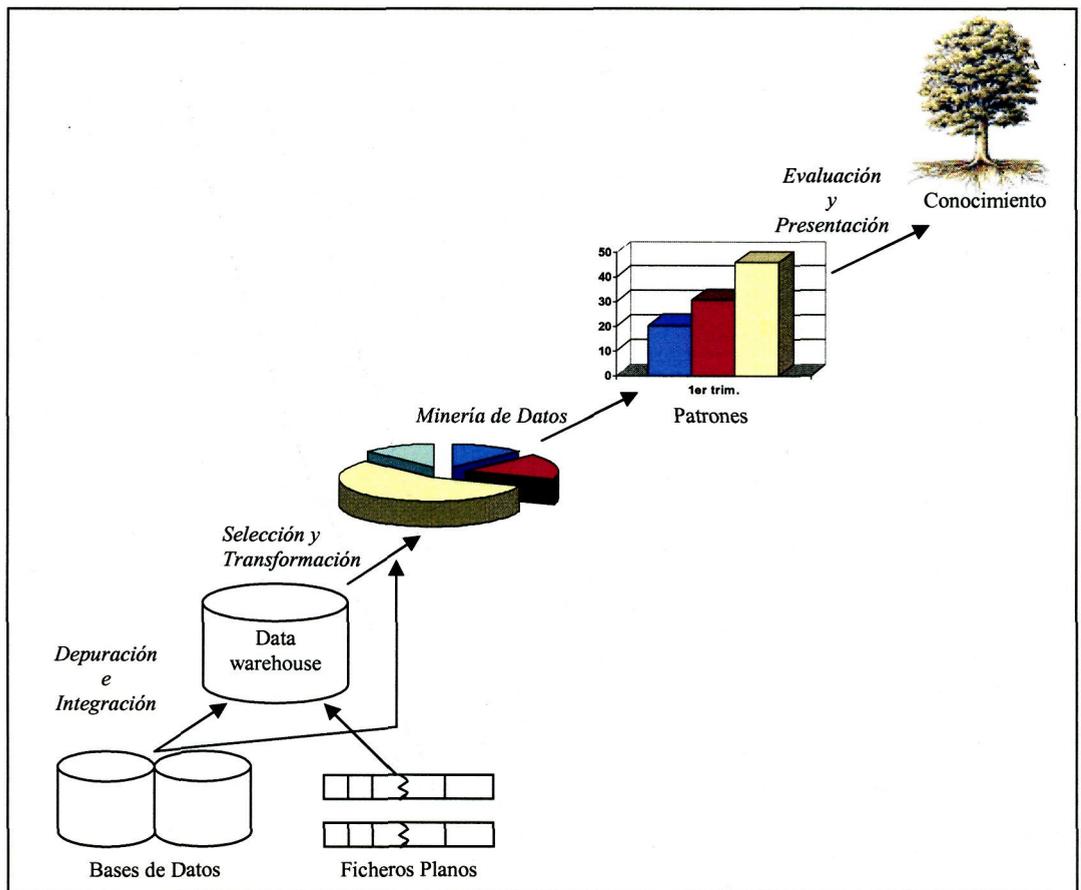


Figura 2.13: La minería de Datos como paso hacia el descubrimiento de Conocimiento (adaptado de Han y Kamber, 2001)

Es importante resaltar que el objetivo final de la Minería de Datos está orientado a aplicaciones más que a la naturaleza básica de los fenómenos. Es decir, la Minería de Datos no está dirigida a identificar relaciones específicas entre variables. Su atención se centra en producir una solución que permita generar predicciones o conocimientos útiles. Por lo tanto, para muchos autores esta tecnología es como una "caja negra" para explorar o descubrir

conocimiento, por lo que se combinan con técnicas como las Redes Neuronales Artificiales, los Algoritmos Evolutivos y las Máquinas Vectoriales (*Support Vector Machines*)

2.6.2.- Fuentes de datos

La minería de datos puede tener sus fuentes de datos en diferentes orígenes, es decir de diferentes procedencias (Web, Bases de Datos etc...). En teoría, la minería de datos debería ser aplicable a cualquier almacén de datos y en cualquier circunstancia, pero en la práctica su aplicación es mucho más viable y eficiente en fuentes tales como Base de Datos Relacionales, Data Warehouses, Ficheros Planos, la World Wide Web, Base de Datos Transaccionales y Bases de Datos Avanzadas.

Las Bases de Datos Avanzadas incluyen, entre otras, Bases de Datos orientadas a Objetos, Bases de Datos Espaciales, Bases de Datos Multimedia etc..

- *Bases de Datos Relacionales* (RDBMS – Relational Data Base Management Systems), compuestas por una colección de *Cuadros*, cada una de las cuales tiene un nombre único. Cada Cuadro consiste en un conjunto de atributos (nombre de columnas) y almacena, por lo general, grandes grupos de registros o filas. Cada registro en una Cuadro relacional representa un objeto identificado con un identificador único y compuesto por los valores de los atributos de ese objeto. (ver figura 2.14)

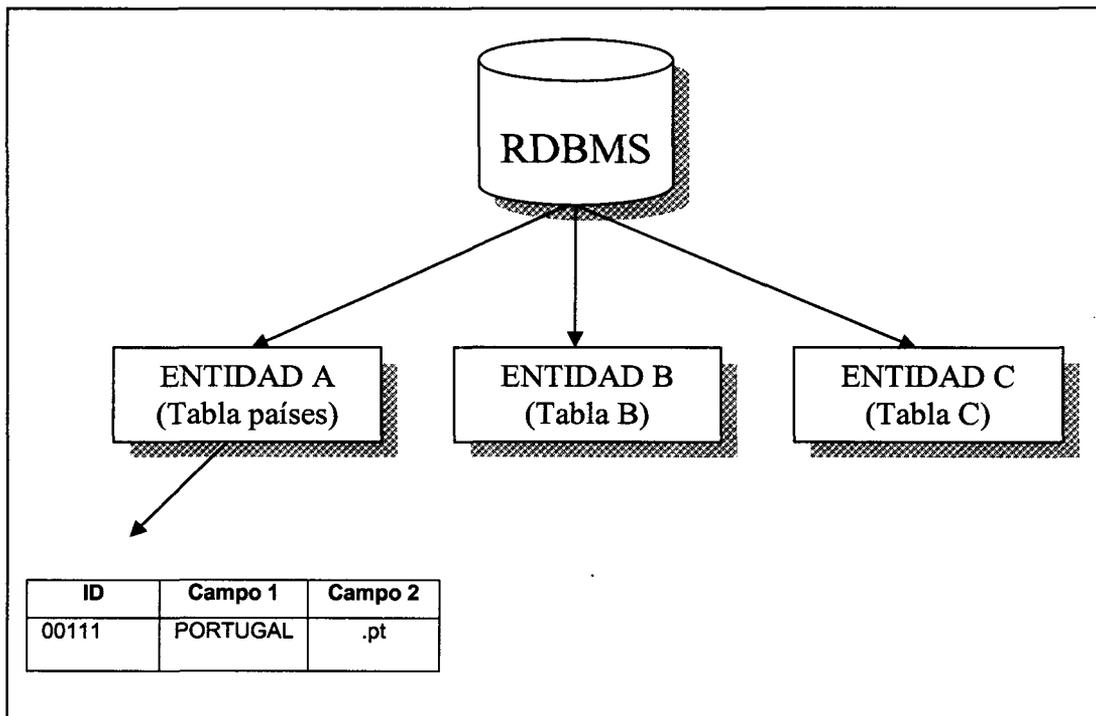


Figura 2.14: Representación gráfica del modelo de Bases de Datos Relacional

Para desarrollar una *Base de Datos Relacional*, se define inicialmente un modelo semántico de datos, tal como el modelo ER (entity-relationship, entidad-relación), donde se describe las entidades y sus relaciones.

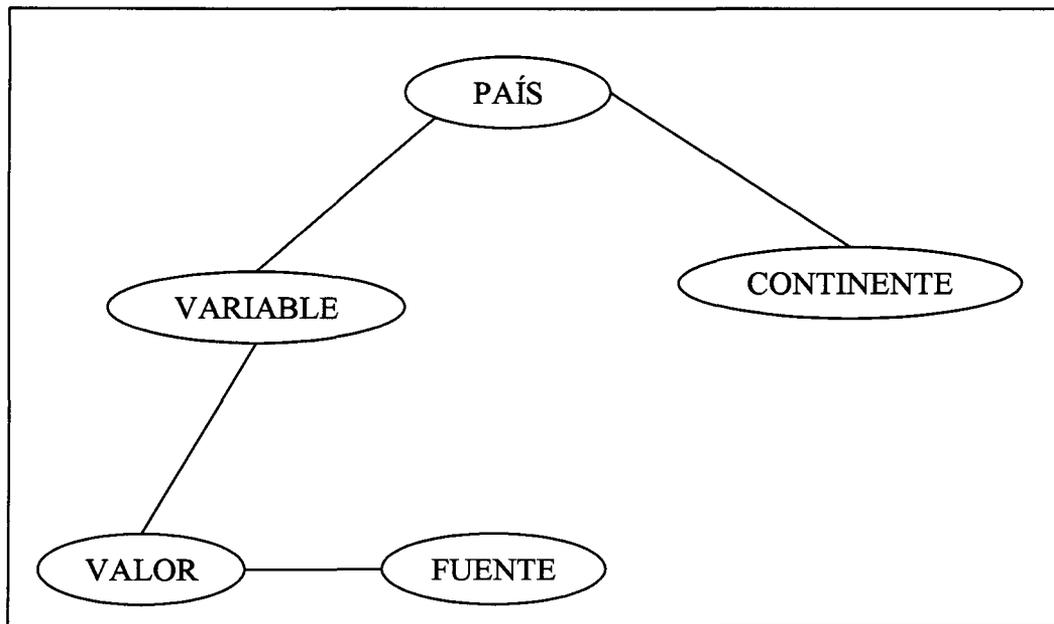


Figura 2.15: Diagrama Simplificado de Relaciones de Entidades

La figura 2.15 muestra un ejemplo de diagrama simplificado de relaciones entre entidades, que en este caso son: país, continente, variable, valor y fuente. Basados en los estudios de competitividad, las relaciones que pueden existir se encuentran entre país continente, país variable, variable valor y valor fuente. Este tipo de diagrama facilita el entendimiento y la comprensión del área de conocimiento así como las relaciones que existen entre los diferentes conceptos. El acceso a estas bases de datos se realiza mediante lenguajes tales como SQL (Structured Query Language; American National Standards Institute, 1992) o mediante asistentes gráficos. La Minería de Datos aplicada sobre estas bases de datos permite realizar estudios longitudinales y de tendencia, así como desviaciones y comparaciones.

Las Bases de Datos Relacionales son una de las fuentes de datos más populares y ricas.

- *Data Warehouse*, por lo general, se modela sobre una estructura de base de datos multidimensional, en la que cada dimensión se corresponde con un atributo o conjunto de atributos del esquema. Cada celda almacena el valor de una medida

agregada tal como la suma del importe de exportaciones. La estructura física de un *Data Warehouse* puede ser una *Base de Datos Relacional* o Cubo multidimensional de Datos. Esto permite el rápido acceso y precómputo de vistas multidimensionales de datos y resúmenes de datos.

Con el fin de facilitar la toma de decisiones, estos almacenes de datos se organizan en relación a aspectos de gran importancia dentro del área de conocimiento. Esto hace posible el uso de On-Line Analytical Processing, o OLAP, una tecnología software que permite a los analistas o directivos analizar y navegar en tiempo real a través de los datos de la organización en base a dimensiones lógicas de los datos y haciendo uso de sistemas avanzados de visualización. Entre las operaciones básicas que se pueden realizar con sistemas OLAP cabe destacar:

- Calcular y modelar a través de dimensiones y jerárquicamente.
 - Análisis de tendencias a través de series temporales.
 - Realizar cortes laterales entre dimensiones, para su visualización en pantalla.
 - Perforar a niveles más bajos de consolidación.
 - Rotación hacia nuevas comparaciones dimensionales
- *Bases de Datos Transaccionales*, en este caso, cada registro de un fichero corresponde con una transacción. A su vez, cada registro contiene un campo clave (*trans_ID*), y un conjunto de atributos que pertenecen a la transacción (ver cuadro 2.9). Este fichero puede adjuntar otros ficheros auxiliares donde se almacenarían más atributos de la transacción.

venta_id	Lista de artículos
S1230455	BC321, DP122, AR101
S1230456	H291 ,B321, AP122

Cuadro 2.9: Representación de los registros de varias transacciones.

- *Bases de Datos Avanzadas*, como se indicó anteriormente en el texto, estas bases de datos surgen como necesidad de expresar y gestionar datos para aplicaciones específicas, como son: bases de datos espaciales, temporales, multimedia, activas, científicas, del conocimiento o información de oficina.

Aunque este tipo de base de datos requiere operadores de gestión específicos para la manipulación de datos espaciales o de multimedia, también se debe reconocer que los frutos de esa gestión son mayores que los de las Bases de Datos Relacionales, mencionadas anteriormente.

2.7.- Funciones básicas de la Minería de Datos

Las funcionalidades presentes en la Minería de Datos son utilizadas para especificar los patrones que deben buscarse en las operaciones de minería. Estas operaciones pueden clasificarse, a grandes rasgos, en dos categorías: Descriptivas y Predictivas.

Las operaciones o tareas Descriptivas permiten caracterizar los datos de la base de datos. Las operaciones Predictivas permiten inferir los datos actuales con el fin de hacer predicciones.

En muchas ocasiones, los usuarios de estos sistemas no conocen inicialmente qué tipos de patrones son de su interés, que pueden precisar hacer búsquedas en paralelo de diferentes tipos de patrones. Por consiguiente, es de interés para ese usuario, que el sistema de Minería de Datos facilite la minería en diferentes fuentes de datos y con diferentes niveles de abstracción, además de facilitar o guiarle en la búsqueda de dichos patrones.

Una de los procesos de más interés para los usuarios es la identificación de los datos asociados a clases o conceptos. Esto se puede conseguir a través de:

- La Caracterización de Datos, representa la recapitulación o síntesis de las características generales o atributos de la clase objetivo de los datos. Existen varios métodos para realizar la caracterización efectiva, como puede ser la operación de OLAP roll-up, sobre un cubo de datos usada y controlada por el usuario, para conseguir la síntesis en una de las dimensiones del cubo. Otro mecanismo puede ser la técnica *inducción orientada al atributo*, utilizada para la generalización y caracterización sin la interacción del usuario. Los resultados de esta caracterización pueden presentarse en varios formatos gráficos o mediante relaciones generalizadas o reglas características.
- La Discriminación de Datos, se realiza a través de la comparación de las características generales de la clase objetivo frente a las características generales de los objetos, de una o varias clases a contrastar. Las clases objeto y de contraste pueden ser especificadas por el usuario, igual que los objetos de clase recuperadas a través de la consultas a la base de datos. Los resultados de esta operación no distan de la forma utilizada en la Caracterización de Datos, aunque el resultado debe contener una medida de comparación entre las clases objetivo y las clases a contrastar. Las descripciones de discriminación expresadas en forma de regla, se conocen con el nombre de reglas discriminantes.

Algunas de las funcionalidades a las que se puede acceder a través de la Minería de Datos son las siguientes:

2.7.1.- Análisis de Asociación

Con el Análisis de Asociación se pretende descubrir las reglas de asociación presentando las condiciones atributo-valor que tienen lugar a un mismo tiempo y de forma frecuentemente en el conjunto de datos. El Análisis de Asociación se utiliza principalmente, para el *análisis de datos de transacción* y el *análisis de cesta de mercado*. En los últimos años, son muchos los algoritmos que han sido propuestos para realizar una minería eficiente de reglas de asociación (Mandan *et al.*, 2003; Lu at

al, 2000, Inokuchi at al., 2003).

La minería de Reglas de Asociación facilita la búsqueda de relaciones interesantes entre ítemes, en una muestra de datos. La pregunta sobre ¿qué tipos de reglas de asociación pueden encontrarse en el proceso de minería?, un ejemplo de respuesta se presenta a continuación:

Reglas de Asociación

La minería de Reglas de Asociación permite la búsqueda de relaciones entre artículos en conjunto de datos. *Market Basket Analysis* es solo una forma de minería de regla de asociación. Estas pueden clasificarse de varias formas:

En función a los tipos de valores que son utilizados en la regla: Regla de Asociación Bolean, cuando la regla está relacionada con asociaciones entre presencia o ausencia de ítemes; o Regla de Asociación Cuantitativa cuando la regla describe las asociaciones entre ítemes o atributos cuantitativos.

En función a las dimensiones de los datos implicados en la regla: Regla de Asociación Unidimensional, cuando los ítemes que están presentes en la regla hacen referencia a una sola dimensión; o Regla de Asociación Multidimensional, cuando los ítemes presentes en la regla hacen referencia a más de una dimensión de dato.

En función a los niveles de abstracción de los datos envueltos en el conjunto de reglas. Por ejemplo:

ingresos(X, "25 mil35 mil") → compra(X, "ordenador portátil")

ingresos(X, "25 mil ... 35 mil") → compras(X, "ordenador")

En este ejemplo, se puede observar que existen dos niveles de abstracción con Ordenador portátil y ordenador. En este caso se habla de Reglas de Asociación Multi-nivel. En el caso en que el nivel de abstracción sea uno, entonces se habla de Reglas de Asociación de un solo nivel.

2.7.2.- Clasificación o predicción

La *clasificación* es el proceso de encontrar un conjunto de modelos o funciones que describan y distingan las clases de datos o conceptos, con el fin de usar el modelo para predecir la clase de objetos cuya etiqueta se desconoce. Es decir, la *clasificación* y la *predicción* son dos formas de análisis de datos que pueden ser utilizadas para extraer modelos, que describen importantes clases de datos o para predecir tendencias en datos futuros. Mientras la *clasificación* predice etiquetas de categorías, la *predicción* modela funciones de valor continuo.

Muchos han sido los métodos de *clasificación* y *predicción* propuestos por los investigadores: *machine learning*, *sistemas expertos*, *estadísticas* o *neurobiología*. Aunque en la mayoría de estos casos, los sistemas permiten trabajar con un conjunto reducido de datos, porque son sistemas residentes en memoria y eso implica que su capacidad está en función de la capacidad de memoria. Pero, en los últimos años la investigación en esta área ha permitido el desarrollo de técnicas de *predicción* y *clasificación* escalables, capaces de gestionar grandes cantidades de datos residentes en memoria - de disco. En muchas ocasiones, estas técnicas hacen uso de procesamiento en paralelo o distribuido (Zak *et al.*, 1997; Shintani *et al.*, 1998; Li y Shasha, 1998).

Antes de realizar las operaciones de minería, sería aconsejable realizar los siguientes pasos, con el fin de mejorar la precisión y la escalabilidad del proceso de *clasificación* o *predicción*.

- Depuración de datos, es decir, preprocesar los datos con el fin de eliminar o reducir ruido y el tratamiento de valores inexistentes. Este paso reduce la confusión durante el proceso de aprendizaje.
- Análisis de relevancia, es decir, identificar y eliminar los atributos que son irrelevantes o redundantes.
- Transformación de datos, consiste en generalizar o normalizar los datos con el fin de optimizar y acelerar el proceso de minería.

Los métodos de clasificación y predicción pueden compararse en función a la precisión predecible, la velocidad de ejecución, la solidez de las predicciones, la escalabilidad de los modelos y a la capacidad de interpretación del modelo resultante.

El Cuadro 2.10 contiene una tipificación de los métodos de clasificación más utilizados en aplicaciones de minería de datos, donde se requiere realizar una operación de clasificación. A continuación presentamos cada tipo de clasificación y sus métodos.

TIPO DE CLASIFICACIÓN	MÉTODOS
Inducción de árboles de decisión	<ul style="list-style-type: none"> • <i>ID3</i>
Bayesiana	<ul style="list-style-type: none"> • <i>Bayesian Classifier</i> • <i>Naive Bayesian Classifier</i> • <i>Bayesian belief networks</i>
Retropropagación	<ul style="list-style-type: none"> • <i>Redes Neuronales Multicapa</i>
Reglas de asociación	<ul style="list-style-type: none"> • <i>Association Rule Clustering</i> • <i>Classification Asociativa</i> • <i>Classification by Aggregating Emerging Patterns</i>
Predicción	<ul style="list-style-type: none"> • <i>Regresión Logística</i> • <i>Regresión de Poisson</i>

Cuadro 2.10: Tipos de métodos de clasificación

Clasificación basada en la Inducción de Árboles de Decisión

Un árbol de decisión es como una combinación de estructura de árbol y diagrama de

flujo, donde cada nodo representa una toma de decisión sobre un atributo y cada rama de ese nodo representa el resultado de la decisión. Cada hoja o nodo final representa, a su vez, las clases o las distribuciones de clases. En el momento de clasificar un caso desconocido, los valores de los atributos del caso se enfrentan al árbol de decisión.

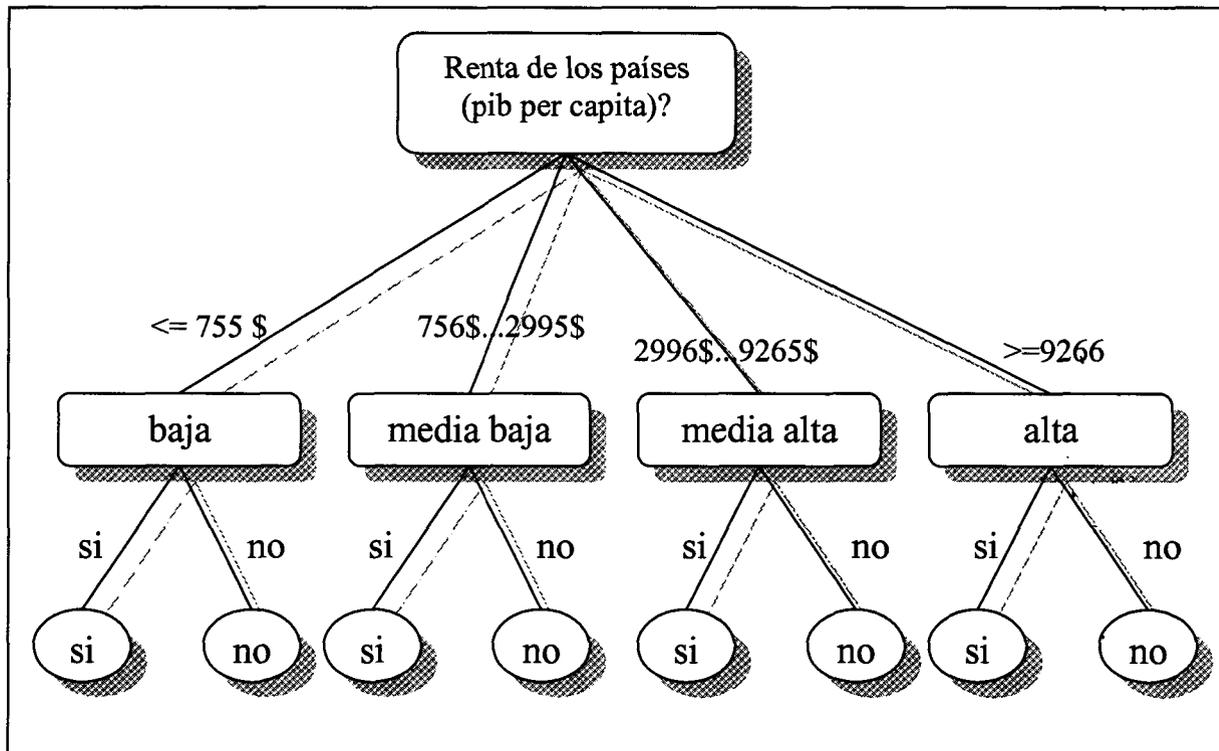


Figura 2.16: Ejemplo de árbol de decisión

En la figura 2.16 se representa un árbol de decisión que clasifica a los países por su renta, es decir, renta baja, media baja, media alta y alta, y en función de su PIB *per capita*. Esta solución puede ser generada de manera automática. Como se puede observar, el nivel inferior corresponde a dos nodos por cada tipo de renta (si – no). Para clasificar un país en función de su renta solo se sigue la secuencia *top-down*. Para entender cómo es el proceso que permite generar de manera automática el árbol sólo hay que seguir las instrucciones descritas en el cuadro 2.11, *Método general de ID3 para la generación de un árbol de decisión por inducción*.

Estos son los pasos básicos a seguir para la creación de un árbol de decisión por inducción, basado en una versión de ID3 (Flachsbart *et al.*, 1994).

Función Generar árbol de decisión (*casos, lista_de_atributos*)

1. crear un nodo N
2. si las *muestras* son todas de la misma clase C entonces
3. devolver N como el nodo final con la etiqueta C
4. si la *lista_de_atributos* esta vacía entonces
5. devolver N como el nodo final con la clase más común de la *muestra*
6. seleccionar el *atributo_de_test*, el atributo con mayor ganancia de información entre la *lista_de_atributos*
7. marcar la etiqueta del nodo N como *atributo_de_test*
8. para cada valor conocido a_i del *atributo_de_test*
9. crear una rama desde el nodo N para la condición *atributo_de_test*= a_i
10. hacer s_i el conjunto de casos del total de casos donde *atributo_de_test*= a_i
11. si s_i está vacía entonces
12. adjuntar la hoja o nodo etiquetado como el más común de la clase en la *muestra*
13. sino adjuntar el nodo devuelto por la función Generar árbol de decisión (s_i , *lista_de_atributos*)

Cuadro 2.11: Método general de ID3 para la generación de un árbol de decisión por Inducción

La estrategia básica de este proceso es la siguiente:

- El árbol empieza con un simple nodo representando la muestra de entrenamiento (paso 1).
- Si en la muestra se encuentran todos los de la misma clase, entonces el nodo se convierte en la hoja y es etiquetada con la clase (paso 2 y 3)
- En caso contrario, el algoritmo usa un medida basada en la entropía, conocida como ganancia de información, como si fuese una heurística para seleccionar el atributo que mejor se separa las muestras en cada clase (paso 6). Este atributo se convierte en el atributo de test o decisión en el nodo (paso 7). En este caso, los valores se consideran discretos, si fueran continuos éstos deben convertirse en discretos.
- Para cada valor desconocido del atributo de test se crea una rama, de igual forma se particionan las muestras (paso 8-10).
- El algoritmo usa el mismo proceso de forma recursiva para formar un árbol de decisión para cada partición de la muestra. Una vez que un atributo se encuentra en un nodo, no necesita ser considerado en ningún nodo descendente (paso 13).

- El particionado recursivo para solo cuando alguno de las siguientes condiciones es cierta:
 - Todas las muestras, para un determinado nodo, pertenecen a la misma clase (paso 2 y 3), o
 - No hay más atributos restantes en los que la muestra puede seguir dividiéndose (paso 4). En este caso, el nodo con la clase más común se usa (paso 5). En cualquier otro caso, la distribución de la clase de la muestra de nodos se puede almacenar.
 - No hay muestras para la rama $test_atributo=a_i$ (paso 11). En este caso, se crea la hoja con la clase mayoritaria en las muestras (paso 12).

Clasificación Bayesiana

Los clasificadores Bayesianos representan un tipo de clasificadores estadísticos. Pueden predecir la probabilidad de pertenencia a una clase, así como la probabilidad de que una muestra pertenezca a una clase en particular.

La Clasificación Bayesiana está basada en el teorema de la probabilidad condicional de Bayes (Chessman y Stutz, 1996).

$$P(H | X) = \frac{P(X | H)P(H)}{P(X)}$$

Donde $P(H | X)$ es la probabilidad a posteriori de la hipótesis H sobre la condición de X cuando es cierta, y $P(H)$ o $P(X)$ son probabilidades a priori, es decir, la probabilidad de que la hipótesis H es cierta a priori, igual que $P(X)$. Por ejemplo, en un mundo de datos donde se representan frutas descritas por el color y la forma (dos atributos), X puede representar el color naranja y la forma redonda, y H la hipótesis de que la fruta sea una naranja. Entonces $P(H)$ es la probabilidad a priori, de que la fruta sea una naranja, y $P(X)$ la probabilidad a priori de que la fruta sea de color naranja y forma redonda. $P(H|X)$ representa la probabilidad a posteriori de que la fruta sea una naranja conociendo la probabilidad a priori de que la fruta es de color naranja y forma redonda. Por lo tanto, el teorema de Bayes permite calcular la probabilidad a posteriori de $P(H|X)$ estimando a priori $P(X)$, $P(H)$ y $P(X|H)$, a partir de los datos de los que se

dispongan.

Algunos estudios han comparado algoritmos de clasificación y han encontrado un clasificador Bayesiano simple conocido como Naive Bayesian Classifier (George y Langley, 1995). Su simplificación estriba en considerar los atributos de la clase independientes. Esto ha permitido obtener resultados más rápido, ya que requieren menos cálculos y tan buenos o mejores que otros métodos de inducción.

En muchas situaciones reales, la independencia de los atributos de las clases u objetos no existe, sino que los atributos tienen cierto nivel de relación que puede ser estimada y utilizada en la red bayesiana. Estas redes bayesianas son conocidas como de creencia (bayesian belief networks). Permiten definir independencia condicional de clases entre subconjuntos de variables.

Clasificación por Retropropagación (Backpropagation)

La clasificación por retropropagación (o backpropagation) es aquella basada en un algoritmo de aprendizaje supervisado y corrector de errores por red neuronal (Rumelhart et al., 1987). Este algoritmo de red neuronal artificial es el más popular y el más utilizado en aplicaciones prácticas. Está representado por una red neuronal artificial de multicapa con alimentación hacia delante (multilayer feed-forward neural network), compuesta de una capa de entrada, una capa de salida y un número indefinido de capas ocultas. (ver figura 2.17)

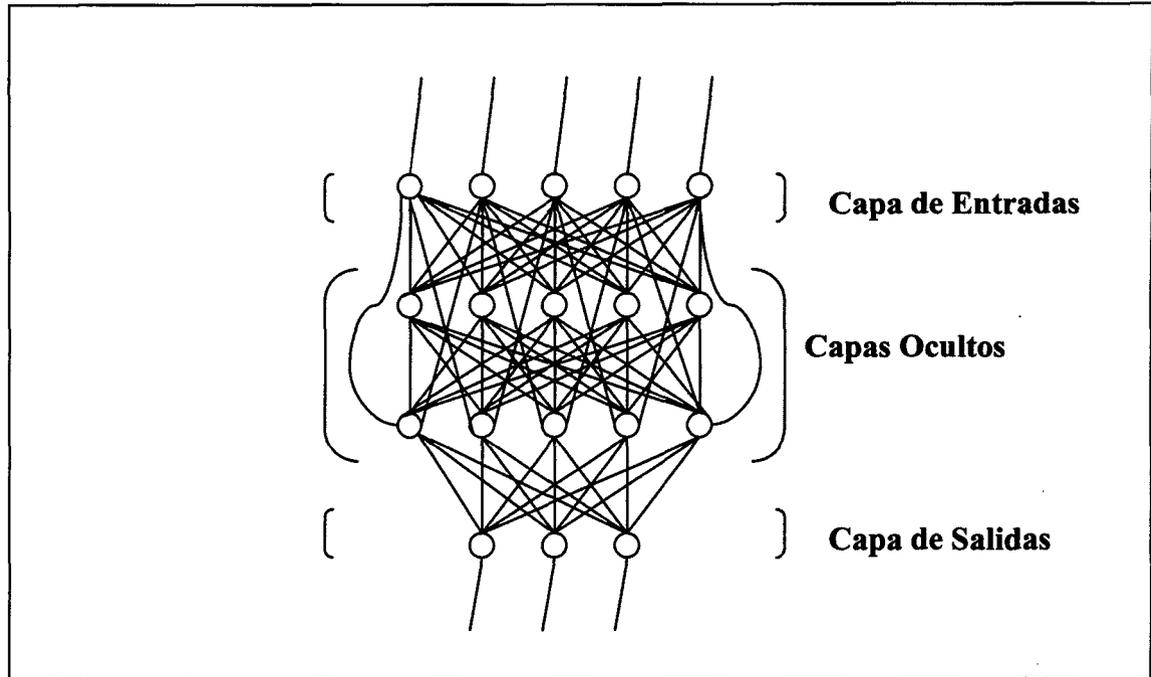


Figura 2.17: Esquema de red neuronal multicapa de alimentación hacia delante

Los pasos que se siguen en una red neuronal de aprendizaje para clasificación usando la retropropagación, son los que aparecen en el cuadro 2.12, que define el método general de aprendizaje de una red neuronal por retropropagación:

Inicializar todos los pesos y márgenes en la red
mientras la condición de finalización no se satisface **hacer**
para cada caso X de la muestra **hacer**
para cada unidad j de entrada u oculta **hacer**

$$I_j = \sum_i w_{ij} o_i + \theta_j$$
 //calcular la entrada de la unidad j en relación a las capas anteriores

$$o_j = \frac{1}{1 + e^{-I_j}}$$
 //calcular la salida para cada unidad j
//retropropagar los errores
para cada unidad j de salida **hacer**

$$Err_j = o_j(1 - o_j)(T_j - o_j)$$
 //calcular errores
para cada unidad j de las capas ocultas, desde la última a la primera **hacer**

$$Err_j = o_j(1 - o_j) \sum_k Err_k w_{jk}$$
 //calcular el error con respecto a la siguiente capa, k
para cada peso w_{ij} en la red **hacer**

$$\Delta w_{ij} = (l)Err_j o_i$$
 //incremento del peso

$$w_{ij} = w_{ij} + \Delta w_{ij}$$

para cada margen θ_j en la red **hacer**

$$\Delta \theta_j = (l)Err_j$$
 //incrementa el margen

$$\theta_j = \theta_j + \Delta \theta_j$$

Cuadro 2.12: Método general de aprendizaje de una red neuronal por retropropagación

Las entradas necesarias para realizar la clasificación con estos pasos son: los casos de entrenamiento, resto de los casos, el ratio de aprendizaje l , una red multicapa de alimentación hacia delante. El resultado de estas operaciones será una red neuronal entrenada para clasificar casos.

El problema característico de las redes neuronales artificiales es la falta de claridad a la hora entender en aprendizaje automático generado a través del entrenamiento supervisado. Aunque se han propuestos métodos de extracción de reglas de decisión de las redes neuronales entrenadas.

Clasificación basada en conceptos de minería de reglas de asociación

La minería de reglas de asociación es un área de investigación muy activa en minería de datos. Una de las aplicaciones más usuales se encuentra en las bases de datos (Jamil, 2002; Sarawagi *et al.*, 2000; Meo *et al.*, 1998). Esta técnica permite la obtención de reglas de asociación a través de la minería de datos. Existen varios métodos para la obtención de reglas de asociación aunque cabe destacar tres *Association Rule Clustering* (ARC) (Lent *et al.*, 1997), *clasificación asociativa* (Liu *et al.*, 1998) y *Classification by Aggregating Emerging Patterns* (CAEP) (Dong *et al.*, 1999).

Predicción

Una de las prácticas mas comunes es la minería de datos es la predicción de datos continuos. Muchos de los problemas basados en predicción de datos continuos, pueden ser resueltos a través de regresiones lineales, y en casos en los que sea necesario la aplicación de regresiones no lineales, éstas pueden ser convertidas a regresiones lineales.

Regresión lineal

$$Y = \alpha + \beta X$$

Donde Y es la variable a predecir con varianza constante, α y β son los coeficientes de regresión, que son estimados dando s casos o punto de datos $(x_1, y_1), (x_2, y_2), (x_3, y_3)$, con las siguientes expresiones matemáticas de mínimos cuadráticos

$$\beta = \frac{\sum_{i=1}^s (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^s (x_i - \bar{x})^2}$$

$$\alpha = \bar{y} - \beta \bar{x}$$

sabiendo que \bar{x} es el promedio de $x_1, x_2, x_3, x_4, \dots, x_n$, e \bar{y} es el promedio de $y_1, y_2, y_3, y_4, \dots, y_n$. Existen casos en los que las variables predictoras (Xs) son más de una, es decir, regresión múltiple, obteniendo una expresión como la siguiente:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2$$

En este caso, también se puede aplicar el método de mínimos cuadrados para calcular α, β_1, β_2 . Cuando la regresión es *no lineal*, por ejemplo, una *regresión polinomial*, entonces se realiza una transformación en las variables, convirtiendo el modelo no lineal a un modelo lineal que puede ser resuelto a través del método de mínimos cuadrados. Por ejemplo

Modelo *no lineal*

$$Y = \alpha + \beta_1 X + \beta_2 X^2 + \beta_3 X^3$$

Se definen nuevas variables que permitan la conversión

$$X_1 = X \quad X_2 = X^2 \quad X_3 = X^3$$

Por lo que el modelo *no lineal* pasa a convertirse en un modelo *lineal múltiple*.

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$$

Existen otros modelos de regresión como pueden ser los de *regresión logística*, de *Poisson* o *Log-lineales*. Estos modelos extienden las capacidades del modelo tradicional.

2.7.3.- Análisis de cluster

El término *clustering* hace referencia al proceso de agrupar un conjunto de objetos físicos o abstractos en clases de objetos similares. El análisis de cluster es un ejemplo de aprendizaje no supervisado. Un cluster se puede definir como una colección de objetos de datos que mantienen en común algún atributo o dimensión entre ellos y difieren con el resto de los objetos de datos de otros clusters. Las similitudes y diferencias son evaluadas a través de los atributos de los objetos de datos. Las

aplicaciones son muchas, desde la operaciones de clasificación en cualquier contexto (clientes, empresas, países etc...) hasta la caracterización o el cálculo de distancias. Algunos de los requerimientos para realizar la minería de datos por clustering son los siguientes:

- Habilidad para procesar diferentes tipos de datos, puesto que son muchas las aplicaciones en las que se requiere hacer clustering de varios tipos de datos, tales como binarios, categóricos, ordinales etc...
- Escalabilidad es una de las características necesarias en el análisis de cluster teniendo en cuenta que en ocasiones el número de muestras es grande y el resultado puede tener sesgo debido a la ausencia de esta característica. Por lo que se precisan métodos de clustering con un alto nivel de escalabilidad.
- Arbitrariedad de las formas de cluster, es decir, el identificar el cluster dependientemente de la forma de dicho cluster.
- Los parámetros de entrada del método de clustering deben ser lo más independientes posibles del contexto o dominio del conocimiento, ya que lo contrario requiere un estudio o análisis de dicho dominio.
- Posibilidad de gestionar datos con ruidos, algo que es muy fácil de encontrar en las bases de datos en cualquier situación real. La mayoría de los métodos de clustering suelen ser sensibles a este ruido (falta de datos, datos erróneos etc...)
- Insensibilidad en la entrada de registros de datos, ya que lo normal es encontrar métodos de clustering que requieran seguir un cierto orden en la entrada de registros de datos, de lo contrario se podría producir un resultado sesgado e inválido para el propósito que se busca.
- Alta dimensionalidad, puesto que en la mayoría de los casos las bases de datos reales suelen contener muchas dimensiones y en la mayoría de los casos, los métodos de clustering son sensibles al número de dimensiones.

- Clustering bajo restricciones, es decir, poder establecer las restricciones bajo los que es posible realizar clustering y obtener un resultado. Esta situación es muy típica en muchas aplicaciones donde existen restricciones que deben respetarse en la solución resultante.
- Capacidad para la Interpretación y la usabilidad, ya que en ocasiones el resultado que se obtiene es difícil de interpretar en el contexto o dominio de conocimiento.

En el análisis de cluster, existen algunas consideraciones que se deben recordar para tener éxito en la tarea. Entre estas consideraciones están:

- Los tipos de datos, estos suelen presentarse en forma de matriz de datos o matriz de distancias entre atributos de los objetos de datos.

$$\begin{pmatrix} d_{11} & \dots & d_{1i} & \dots & d_{1p} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ d_{i1} & \dots & d_{ii} & \dots & d_{ip} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ d_{n1} & \dots & d_{ni} & \dots & d_{np} \end{pmatrix} \quad \text{Matriz de datos } nxp$$

$$\text{Matriz de distancia } nxl \quad \begin{pmatrix} 0 & \dots & \cdot & \dots & \cdot \\ d(2,1) & 0 & \cdot & \cdot & \cdot \\ d(3,1) & d(3,2) & 0 & \dots & \cdot \\ \cdot & \cdot & & 0 & \cdot \\ \cdot & \cdot & & & \cdot \\ d(n,1) & d(n,2) & & \dots & 0 \end{pmatrix}$$

Donde $d(2,1)$ es una medida de la distancia del atributo 2 al atributo 1 de los objetos 2 y 1. A su vez, los datos pueden ser de varios tipos, tales como de intervalo o escala, binarios, ordinales, nominales, por ratio o mixtos.

- Los métodos de clustering, puesto que son muchos los métodos de clustering que en muchas ocasiones su aplicación depende del tipo de dato, del propósito y de la aplicación.

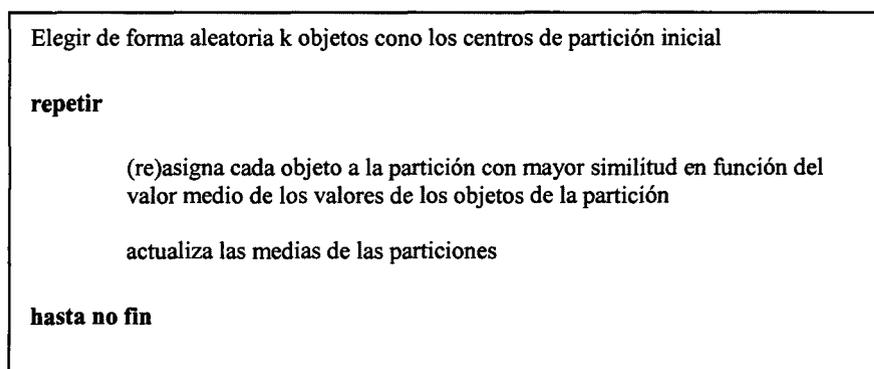
TIPO DE CLUSTERING	MÉTODOS
Particionado	<ul style="list-style-type: none"> • <i>k-mean</i> • <i>k-medoid</i> • CLARA
Jerárquicos	<ul style="list-style-type: none"> • Divisivos • Aglomerativos • Integrados
Densidad	<ul style="list-style-type: none"> • DBSCAN • GDBSCAN • OPTICS • DENCLUE
Rejas (Grid)	<ul style="list-style-type: none"> • STING • Wavelet • CLIQUE
Modelo matemático	<ul style="list-style-type: none"> • COBWEB • CLASSIT • AutoClass • P-AutoClass
Detección de <i>Outlier</i>	<ul style="list-style-type: none"> • Basados en estadística, • Basados en la distancia • Basados en la desviación

Cuadro 2.13: Tipos de métodos de clustering

Como se puede observar en el cuadro 2.13, los tipos de clustering pueden ser

clasificados en: particionado, jerárquicos, densidad, rejillas, modelo matemático y detección de outlier, con sus correspondientes métodos. A continuación, se presentan cada uno de ellos.

1. Los *métodos de particionado*, permiten la clasificación de un número n de objetos en k particiones tal que ($k \ll n$) y cada partición representa un cluster. (ver cuadro 2.14)



Cuadro 2.14: Proceso general particionado

Los datos de entrada que necesita este método son el número de clusters k , y el número de objetos n de la base de datos, el resultado que se obtiene es un conjunto de k clusters que cumplen los criterios de similitud. Los métodos a considerar de particionado incluyen *k-mean*, *k-medoid* y CLARA (clustering para grandes bases de datos) (Kaufman y Rousseeuw, 1990)

2. Los *métodos jerárquicos* de clustering permiten agrupar los objetos en un árbol de clusters. Pueden ser clasificados en divisivos (Ej. DIANA – Divisive Análisis), aglomerativos (Ej. AGNES - AGglomerative NESTing) o integrados (Ej. BIRCH - Balanced Iterative Reducing and Clustering Using Hierarchies, Chamaleon etc...) (Zhang *et al*, 1996, Karypies *et al*, 1999)
3. Los *métodos de clustering* basados en densidad son ideales cuando los

clusters tienen formas irregulares y se caracterizan por regiones de alta densidad de objetos, en el espacio de datos, frente a otras regiones con baja densidad. Entre los métodos basados en análisis de densidad se encuentran el DBSCAN (A Density-Based Clustering Method Base on Connected Regions with Sufficiently High Density – Un método de clustering basada en densidad sobre regiones conectadas con suficientemente alta densidad) (Ester *et al*, 1996), el GDBSCAN (Density-Based Clustering in Spatial Databases – Clustering basado en la densidad en bases de datos espaciales) (Sander *et al*, 1998), el OPTICS (Ordering Points To Identify the Clustering Structure – Ordenar Puntos para Identificar la Estructura de Clustering) (Ankerst *et al*, 1999), el DENCLUE (DENSisty-based CLUstEring – clustering basado densidad) (Hinneburg *et al.*, 1998), etc..

4. Los *métodos de clustering basados en rejillas (grid)* utilizan una estructura de datos de rejilla de multi-resolución. Es decir, estos métodos dividen el espacio en un número finito de celdas que forma las rejillas donde se realiza el proceso de clustering. Una de las ventajas de estos métodos es que el tiempo de procesamiento no depende del número de objetos de datos sino del número de celdas. Entre los *métodos de clustering basados en rejillas* se encuentran: STING (STastical INformation Grid – Reja de información estadística) (Wang *et al*, 1997) un método de clustering que divide el espacio en celdas rectangulares con varios niveles y jerarquías; WaveCluster, que es básicamente clustering haciendo uso de la transformación *wavelet*, entendiendo transformación *wavelet* como la técnica que descompone una señal en diferentes subbandas de frecuencia (Sheikholeslami *et al.*, 1998); CLIQUE (CLustering In QUEst,) (Agrawal *et al.*, 1998), donde se combina el *clustering basado en densidad* y el *basado en rejillas*, útil para espacios multidimensionales, que suelen aparecer a partir de grandes bases de datos.
5. Los *métodos basados en modelos* que pretenden acoplar los datos a algún modelo matemático. Estos métodos suelen tener dos tipos de enfoques: el estadístico o el de redes neuronales. Dentro de los métodos estadísticos cabe destacar COBWEB, un método conceptual e incremental de clustering que, dado un conjunto de pares atributo-valor, genera un clustering jerárquico

llamado árbol de clasificación (Fisher, 1987); CLASSIT, una extensión de COBWEB para datos continuos (Gennari *et al.*, 1989); AutoClass, método en el que se utiliza análisis estadístico Bayesiano para estimar el número de clusters (Cheesman *et al.*, 1989), y P-AutoClass una versión de procesamiento en paralelo del método AutoClass (Pizzuti *et al.*, 2003). En los métodos de redes neuronales cabe distinguir dos clases: las redes de aprendizaje competitivo (Rumelhart y Zipser, 1985), y los mapas de rasgos autoorganizados (o SOFM – Self-Organizing Feature Map) (Kohonen, 1995).

6. *Los métodos de análisis de outlier*, permiten la identificación de objetos de datos que no cumplen con el comportamiento, modelo o clasificación del resto de los objetos de datos de la muestra. En muchas ocasiones, estos objetos pueden ser de gran interés para la investigación ya que pueden contener información importante y oculta en relación a la generalidad de la muestra (Santos-Periera y Pires, 2002) Entre los más populares están los basados en estadística, los basados en la distancia y los basados en la desviación.

2.8.- TENDENCIAS TECNOLÓGICAS

Los trabajos de investigación en tecnologías de la información y comunicaciones que se pueden observar en la actualidad, están dirigiendo sus esfuerzos y sus mejores resultados hacia tres tendencias bien diferenciadas: la gestión de información y el conocimiento, las capacidades de los sistemas informáticos, y las comunicaciones (ver figura 2.18).

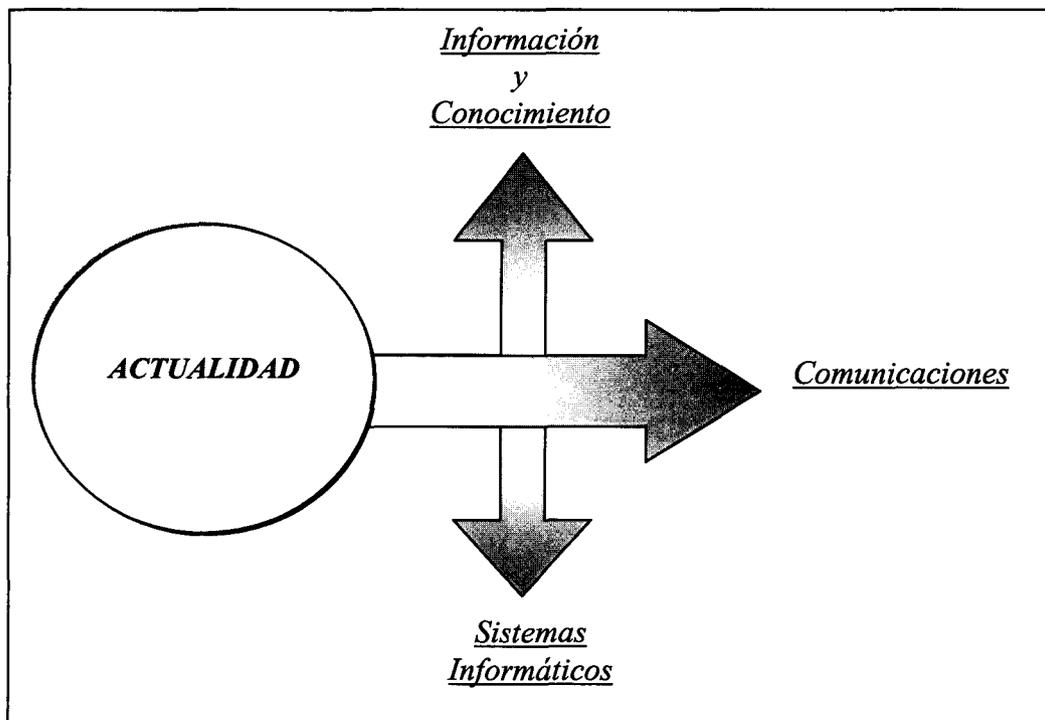


Figura 2.18: Diagrama de representación de tendencias

Cada una de estas tendencias afecta en gran medida no sólo a las organizaciones y a los mercados sino también al resto de las tendencias, debido al nivel de integración y relación que existen entre ellas. Iniciamos pues, este apartado introduciendo la Web Semántica.

2.8.1 La Web Semántica

Quizás una de las áreas con mayor nivel de evolución e influencia global sea la Web. Las tecnologías utilizadas para publicar información a través de la red de redes y disponible internacionalmente, se están integrando en los sistemas de información corporativos y llegan a ser requisito imprescindibles en estos sistemas. De esta forma, los sistemas de información corporativos se acercan a la Web y viceversa.

Los orígenes de la idea básica de la Web quedan reflejados en el artículo de Vannevar Bush, en 1945, *As we may think*. Pero su existencia y evolución ha sido posible gracias, principalmente, a los esfuerzos de Tim Berners-Lee, proclamado inventor de la Web y fundador de la World Wide Web Consortium (W3), ubicado en el Massachusetts Institute of Technology, Laboratory for Computer Science en colaboración con el CERN (Centro Europeo de Investigación Nuclear), donde se originó la Web, del DARPA (U.S. Defense Advanced Research Project Agency) y de la Comisión Europeo. Los trabajos que se realizan en el W3 suelen terminar convirtiéndose en estándares internacionales. Entre los documentos de trabajos publicados por la W3, se encuentran los relacionados con lenguajes como HTML, XML y OWL.

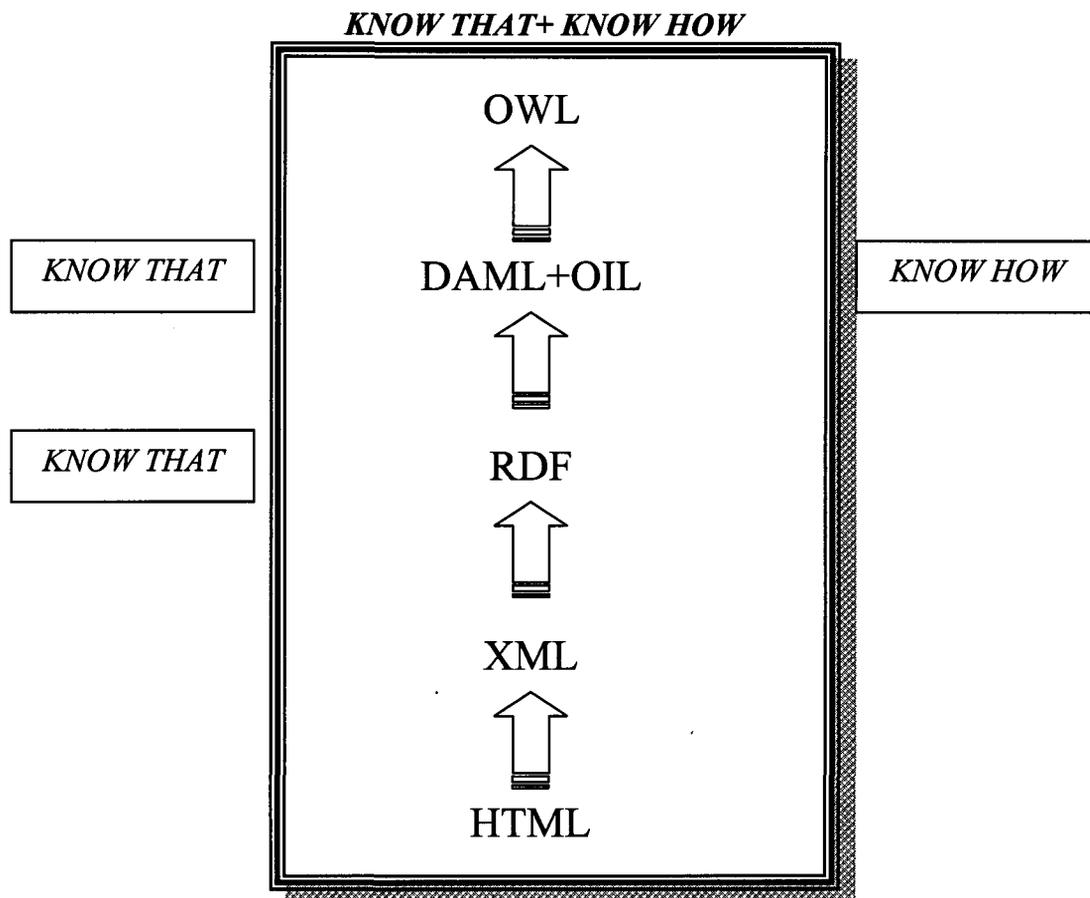


Figura 2.19: Diagrama de representación de tendencias en lenguajes aplicados a la WEB

La figura 2.19 presenta un diagrama de representación de tendencias en lenguajes aplicados a la WEB. El HTML (HyperText Markup Language), ha sido clave para el desarrollo de la World Wide Web (W3C, 1999). Utiliza un formato sin derechos de autor, basado en el lenguaje SGML (Standard Generalized Markup Language; SGML, ISO 8879:1986). Los documentos Web escritos en HTML pueden ser procesados por un amplio rango de herramientas, desde simples editores de texto hasta sofisticadas herramientas de autor WYSIWYG (What You See Is What You Get). Por otro lado, el XML (Extensible Markup Language) es un formato de texto bastante simple y flexible, desarrollado en base al SGML, que permite crear documentos estructurados con una sintaxis superficial, pero sin imponer limitaciones semánticas al significado de estos documentos. Sin embargo, el XML Schema, es un lenguaje que sí permite la creación restringida de estructuras en documentos XML.

El RDF, es un modelo de datos aplicados a objetos (recursos) y las relaciones entre estos, que aportan semánticas simples a este modelo de datos pudiendo ser representados en una sintaxis XML. Por otro lado, el RDF Schema, es un vocabulario para la descripción de propiedades y clases de estos recursos RDF, con semánticas para la generalización-jerarquías de dichas propiedades y clases. Sin embargo, el OWL, añade más vocabulario a la descripción de las propiedades y las clases, que incluye, las relaciones entre las clases (uniones, intersecciones etc.), cardinalidad (exactamente uno), igualdad, clases enumeradas y características de propiedades (W3C, 2003), etc...

Las tendencias tecnológicas en la Web (ver figura 2.17), se orientan hacia una *Web Semántica* como pronosticó Tim Berners-Lee pero, el desarrollo del estándar OWL (Ontology Web Language) permitirá no sólo la meta descripción del contenido de la Web, sino también su inferencia con operadores de lógica de primer orden (know that+how), convirtiendo la red de redes no sólo en la gran biblioteca, como es conocida en la actualidad, sino en un impensable motor de inferencia, distribuido globalmente.

“La Web Semántica es una extensión de la Web actual en la que la información se le da un significado bien definido, permitiendo mejorar el trabajo en cooperación entre los ordenadores y la gente.” (Fuente: Adaptado de Berners-Lee et al., 2001)

La Web Semántica, con su base ontológica, permitirá, como prevé Berners-Lee, no solo la comunicación entre agentes inteligentes bajo nuestras órdenes, sino también permitirá que nuestros teléfonos pidan al televisor y el equipo estero de música, que bajen el volumen, mientras suena. Pero, para ello, tanto el teléfono como el televisor y el equipo de música deben “entender” o compartir una misma definición de conceptos tales como volumen o bajar. Es ahí, donde entra en juego el papel de la descripción ontológica del modelo que deben compartir el teléfono, el televisor y el equipo estéreo de música. No obstante, hasta hace poco tiempo, y aunque el objetivo de las ontologías estaban originalmente motivadas por la necesidad de compartir y re-usar las bases de conocimiento, el re-uso y el compartir ontologías, estaba (y en muchos casos, está) limitado debido a que los usuarios y otros diseñadores ontológicos no siempre comparten las mismas suposiciones que los diseñadores originales (Gruninger y Lee, 2002).

2.8.2.- E-Services

El comercio electrónico actual evoluciona hacia un nuevo paradigma que se conoce con el nombre de *E-Service*. Este nuevo paradigma plantea cambios significativos en los retos tradicionales, y estos cambios están directamente relacionados con el incremento de beneficios a través de entornos electrónicos (Rust y Lemon, 2001). Asimismo, y en contraposición con lo sucede en el comercio electrónico, que intenta reducir los costes a través de la automatización y el incremento de la eficiencia de los procesos, los *e-services* incrementan los beneficios a través de la extensión de los servicios y la creación de un marco de relación rentable para el cliente.

El resultado natural de la economía basada en redes es el resurgimiento del cliente como punto de enfoque de todos los negocios. Es decir, las tecnologías de la información y comunicaciones han facilitado la personalización en masa que favorecen la personalización de la relación empresa-cliente a múltiples clientes y en tiempo real.

Por otro lado, la rápida evolución de las tecnologías avanzadas, las comunicaciones inalámbricas, el incremento del ancho de banda, la minería de datos y los agentes inteligentes, han contribuido a ofrecer mecanismos de accesibilidad efectivos a los clientes objetivo, y esto a su vez facilita a las organizaciones el suministro de un mayor número de opciones, productos, servicios, en sus transacciones empresariales. Como consecuencia de ello, las organizaciones se encuentran bajo presión para mejorar sus procesos comerciales, dirigidos al desarrollo de nuevos mercados y a mejorar su capacidad competitiva haciendo uso de estas tecnologías. A su vez, esta situación de centrarse en el cliente, provoca una transformación de la economía desde una economía basada en productos a una economía basada en servicios, con una clara tendencia hacia una nueva era económica de *e-services*.

Como argumentan Rust y Kannan (2003), resulta clave que las empresas adopten una orientación hacia los *e-services*, independientemente de que vendan productos o servicios. La razón de ello es debido a que el entorno electrónico basado en redes, demanda un enfoque centrado en el cliente para ser más eficiente y efectivo a la hora de encontrar las necesidades del mercado y mantenerse en él de manera competitiva. La noción de redes electrónicas incluye Internet y las redes inalámbricas, así como cualquier otro mecanismo que permita utilizar el canal de conexión a través de Internet.

2.8.3.- Comunicaciones de datos

En el caso de las comunicaciones, las tendencias parecen dirigirse, básicamente, en tres direcciones: ancho de banda, soporte físico y movilidad. Los trabajos de investigación cuyo objetivo es el incremento del ancho de banda sobre soportes clásico puede representar avances muy significativos en los próximos años, un ejemplo de ello se puede ver en los trabajos sobre el protocolo TCP (JIN *et al.*, 2003). Este tipo de avances puede afectar de forma significativa a la cantidad de datos que se pueden transmitir en la unidad de tiempo. Como consecuencia de ello, los efectos sobre los mercados, las organizaciones y los usuarios pueden ser bastante importantes.

En relación al soporte físico, se observan dos tendencias tecnológicas bien diferenciadas, el PLC (*Power Line Communication*), que permite la transferencia de datos en banda ancha (más de 2Mb/s) a través de la red eléctrica, lo que facilitaría la informática invisible (*ubiquity computing*). Y, por otro lado, otra tendencia tecnológica que ya ha mostrado su relevancia es la comunicación inalámbrica (tecnologías basadas, principalmente, en protocolos 802.XX). El éxito ha estado marcado, por un lado, por la facilidad de instalación de este tipo de redes y, por otro lado, el ancho de banda que se puede obtener para realizar las comunicaciones.

2.8.4.- Informática Móvil

La movilidad de la informática que se conoce en la actualidad como informática móvil, permitirá llevar las capacidades de los ordenadores personales actuales y de las redes informáticas (incluyendo Internet), es decir, ordenadores de grandes capacidades pero de pequeño tamaño, a cualquier lugar donde se desplace cualquier persona. Este hecho marcará las formas en que se utilizan, para la comunicación interpersonal, el procesamiento de datos y las transacciones de los usuarios con otros usuarios u organizaciones.

Una de las fuerzas que ha conducido al incremento y a la evolución de las TICs y la Web, las comunicaciones inalámbricas y los dispositivos portátiles, es el concepto de información en cualquier sitio y en cualquier momento (*information anywhere, anytime*) (Srimani, et al., 2002). La combinación de estas tecnologías, a raíz de los recientes avances en informática y tecnologías de la información, permite llevar la informática a entornos móviles. En la mayoría de los casos, los usuarios requieren de información en tiempo real para la toma de decisiones críticas, por ejemplo, inversiones en bolsa y transacciones comerciales. Por lo que, el tiempo de respuesta se convierte en una medida clave del rendimiento. Por lo tanto, es importante que el usuario tenga conocimiento en cada momento y en cada sitio, cómo establecer comunicación en las redes globales.

Por otro lado, las aplicaciones móviles se podrían clasificar en: aplicaciones verticales y aplicaciones horizontales. En las aplicaciones verticales, los usuarios acceden a los datos dentro de una celda específica y el acceso de otros usuarios fuera de la celda es denegado. Por ejemplo, la situación en la que un usuario quiere obtener información sobre la localización de restaurantes o centros de emergencia dentro de una celda como puede ser un área física perteneciente a una ciudad o un aeropuerto. En las aplicaciones horizontales, sin embargo, los usuarios cooperan en la realización de una tarea y pueden manipular datos distribuidos a través del sistema o la red. Los mercados para este tipo de aplicaciones está en crecimiento, fundamentalmente en dos áreas: las aplicaciones basadas en correo/SMS y los servicios de información.

Como consecuencia de ello, las organizaciones tendrán que utilizar tecnologías como la de *e-services* y tecnologías de agentes para dar servicio a los usuarios móviles. La gestión de la localización del usuario juega un papel fundamental en los sistemas de informática móvil y sus aplicaciones. Como ejemplo de sistemas tenemos los *Geographic Information Systems (GIS)* y su integración con sistemas de marketing o guías turísticas.

Otro aspecto importante a destacar dentro de la telefonía móvil es el procesamiento de consulta a sistemas de información desde dispositivos móviles. Las comunicaciones inalámbricas y los canales que se crean a través de ellas harán posible la diseminación de información a un número ilimitado de clientes móviles. Este hecho requiere del uso eficiente de tecnologías de emisión, planificación, indexación, memoria temporal,...etc. Asimismo, con la informática móvil, los usuarios serán más exigentes en la calidad y el tiempo de respuesta de los sistemas de información, lo que requeriría una mayor utilización de las tecnologías que facilitan la gestión de contenido semántico de los datos.

2.8.5.- Informática Ubicua

Los desarrollos tecnológicos del hardware informático y sus capacidades funcionales, están provocando cambios muy significativos en sus usos y sus aplicaciones. Estos

desarrollos tecnológicos van a provocar que los componentes físicos se conviertan en invisibles (ubicuos), como idearon en sus momentos autores tales como Donald Norman (2000), Mark Weiser y John S. Brown (1997), entre otros. Parafraseando a Mark Wieser (1991), “*Las tecnologías más profundas son aquellas que desaparecen. Se adentran en la fábrica de vida diaria hasta que se hacen indistinguibles.*” Como se puede apreciar en la figura 2.20, ha existido y existe, una evolución en la relación de los seres humanos con los ordenadores o los sistemas informáticos, que en la figura se establecen como fase 1 – *Era Mainframe*; fase 2 – *Era PCs*; Transición - *Internet y la Informática Distribuida*; y fase 3 – *Informática Ubicua o Invisible*.

En la fase 1, la era *Mainframe*, los usuarios de los ordenadores eran expertos que trabajaban detrás de las puertas cerradas de los laboratorios. En esta época, los recursos informáticos eran escasos y, como consecuencia, debían ser negociados y compartidos entre los usuarios de estos *mainframes*. En la actualidad sigue existiendo una informática de *mainframes*, aplicada a tareas tales como: simulaciones físicas o virtuales, que van desde la simulación del tiempo hasta las simulaciones virtuales. Por lo tanto, el modelo *mainframes* implica que un ordenador es compartido por varios usuarios para la realización de múltiples tareas.

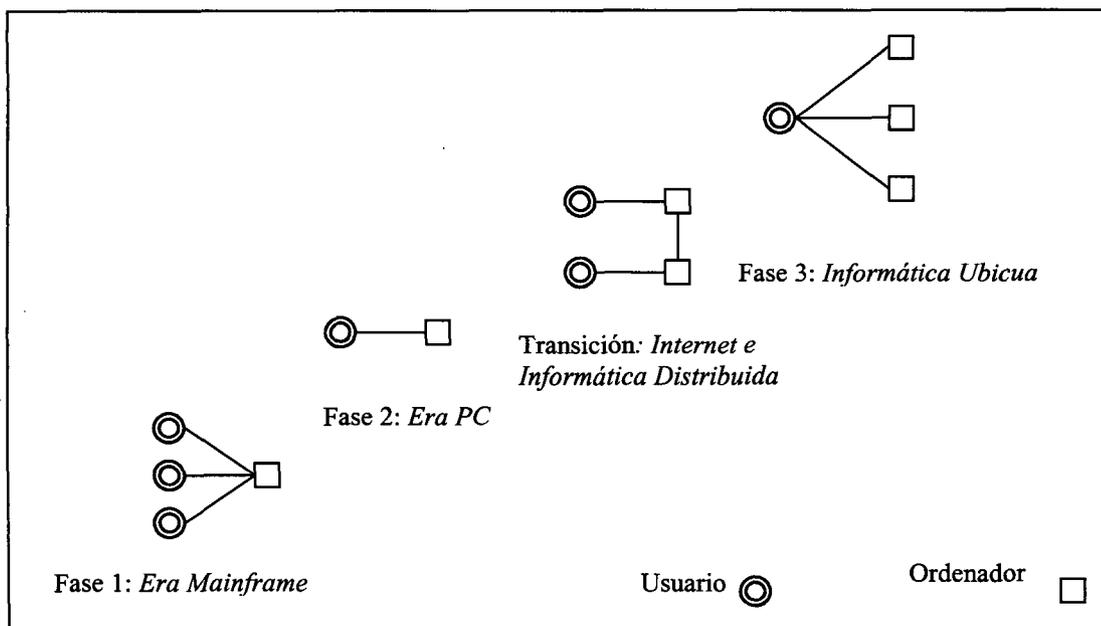


Figura 2.20: Evolución del uso de los sistemas informáticos a través del tiempo.

La fase 2, la era PC, fue considerada la segunda gran tendencia de la era de la informática. Como se publicó en el artículo *Transitions to the information high-way* (IDC, 1995), en 1984, el número de personas que utilizaban ordenadores personales sobrepasaba el número de personas que utilizaban Mainframes. Asimismo, la relación que se presenta con los PCs es personal, incluso íntima. Como es normal, el PC de un usuario contiene sus archivos, su música, su agenda personal, y, como consecuencia, este usuario interacciona de forma directa y profunda con su PC. Además, cuando se interacciona con un ordenador personal se está ocupado, es decir no se realiza ninguna otra tarea. En algunos casos, los usuarios colocan nombre a sus PCs. Este tipo de relación, es muy parecida a la que se tiene con la posesión de un automóvil, que requiere unos altos niveles de atención para su operatividad.

De la misma forma que una persona puede poseer varios coches, un usuario también puede poseer varios PCs, en la casa, en el trabajo,...etc. Además, muchos PDAs (*Personal Digital Assistance*), tales como Zaurus, Pilot o Toshiba, son utilizados en la actualidad como ordenadores personales.

En la Transición, millones de usuarios, sus datos y, en general, sus informaciones, han sido interconectadas. Esta red es dinámica, mundial y está evolucionando de forma continua, 24 horas al día, 365 días al año. A esto hay que añadir sus millones de páginas, procesos y transacciones comerciales, sus millones de servidores y sus millones de usuarios interconectados en tiempo real. Resulta interesante encontrar en Internet elementos de la era Mainframe y de la era PC integrados entre sí. Ésta es una informática basada en el modelo cliente-servidor a escala mundial, que durante su periodo inicial, hacía que los usuarios de los PCs accedieran a los servicios que ofrecían los *mainframes* a través de los programas servidores. En la actualidad, con la aparición de la informática punto-a-punto (*peer-to-peer computing*) este tipo de relación está presente entre los usuarios de las redes. Es decir, un usuario tendrá componentes cliente y servidor a un mismo tiempo, para establecer intercambio de documentos con otros usuarios de este tipo de sistema informático.

Internet se ha convertido en un fenómeno mundial que permite la cooperación e interrelación de múltiples agentes sociales: financieros, consultores, académicos, investigadores, políticos, empresarios, usuarios finales, que permite el intercambio y la

creación de nuevo conocimiento. A su vez, este hecho facilita la rápida y rica evolución del conocimiento. En este sentido, se constata las nuevas formas de negocio, las nuevas organizaciones globales, los nuevos tipos de servicio,....etc.

En la fase 3, era de la Informática Ubicua, como se puede observar en la figura 2.18, ésta era plantea un escenario inverso al de la fase 1, era de *Mainframe* donde los usuarios compartían los recursos informáticos. En esta tercera fase, sin embargo, el usuario es el agente compartido entre muchos ordenadores que, en muchas ocasiones, pueden tener acceso a Internet de forma muy rápida. Estos ordenadores podrán estar embebidos en nuestro entorno diario, es decir, en nuestros espejos, nuestra vestimenta, en nuestros vehículos, en suma, en todo. La identificación en Internet de cada uno de estos ordenadores embebidos es actualmente posible gracias a la tecnología ya desarrollada basada en el protocolo IP versión 6 (Deering y Hinden, 1995). Con éste, cada átomo de la superficie terrestre puede direccionar más de mil dispositivos diferentes. Como consecuencia, el número de dispositivos que pueden estar conectados a Internet puede ser inimaginable.

Por otro lado, la informática ubicua se caracteriza principalmente por la conexión de objetos del mundo físico con la informática y la red. Esta situación va a estar presente a nivel de múltiples escalas (Gabriel, 1995). El impacto social que tendrá este tipo de tecnologías podrá ser semejante a otras dos tecnologías que se han convertido en ubicuas (Weiss, 1997). La primera es la escritura, que se encuentra en todos sitios, es decir, en paredes, libros, billetes,...etc. La segunda es la electricidad, que surge invisible a través de las paredes de cada edificio u objeto. Los seres humanos nos hemos acostumbrado a su uso y nos hemos olvidado del gran impacto que tiene en nuestras vidas diarias. De esta forma, se convertirá la informática ubicua (Norman, 2000).

Una vez analizadas las tecnologías aplicadas a la Gestión del Conocimiento, y con el fin de contribuir a rellenar esa laguna importante, existente actualmente en los trabajos de investigación sobre gestión del conocimiento se presenta un sistema avanzado de gestión del conocimiento aplicado a la competitividad internacional. Asimismo, se presenta una simulación de dicho sistema.

Capítulo 3

**Sistema Avanzado de Gestión del
Conocimiento Aplicado a la
Competitividad Internacional**

*“Estamos ahogándonos en mares de datos y
hambrientos de conocimiento”*

(Adaptado de Jiawei Han y Micheline Kamber)

3.1. FUNDAMENTOS BÁSICOS DEL SISTEMA

Con el fin de satisfacer el objetivo propuesto en esta investigación, y contribuir de esta forma a la eliminación de esa laguna existente en los trabajos de investigación actuales sobre gestión del conocimiento, se presenta en esta tesis, un sistema avanzado de gestión del conocimiento aplicado a la competitividad de los territorios. Dicho modelo, pretende actuar como una plataforma modular, autónoma y abierta para el desarrollo de bases de conocimiento actualizables, que permite la gestión y el descubrimiento automático de nuevo conocimiento sobre la competitividad internacional de los territorios. Asimismo, la propuesta de este modelo actúa como puente de enlace entre dos áreas de conocimiento muy importantes en la actualidad: la competitividad internacional de los territorios y las tecnologías orientadas a la gestión del conocimiento.

En este último capítulo de la tesis, se describe y simula una propuesta para el desarrollo de un Sistema Avanzado de Gestión del Conocimiento aplicado a la Competitividad Internacional de los territorios, en el que se conjugan los componentes para la gestión y generación de conocimiento (know-that), con elementos para la gestión y generación del know-how. Para ello, se inicia esta investigación desde una de las fuentes de información características de esta área de conocimiento un informe de competitividad internacional representativo, para el que se plantea una solución tecnológica integradora, que permita funcionalidades avanzadas, desarrollo ontológicos con posibilidad evolutiva y altos niveles de autonomía. Este tipo de soluciones tecnológicas, requieren de elementos de software estables, de última generación, abiertos, integrables y especializados, en áreas muy específicas de minería de datos, etc....

3.1.1.- Orígenes del Sistema

Este trabajo está inspirado en el documento denominado “*La Competitividad Internacional de Canarias*” elaborado por Rosario Ramos (2002), en el que se analizó la capacidad competitiva de Canarias en el marco internacional. Para ello, se buscó el

cumplimiento de los siguientes objetivos, que ya han sido mencionados en el capítulo uno de la tesis:

1. Determinar el índice de competitividad y posicionamiento competitivo internacional de Canarias con respecto a otros países.
2. Determinar los factores y subfactores que conforman la competitividad de Canarias.
3. Analizar el balance de la competitividad de Canarias, identificar las fuentes de ventajas y desventajas competitivas comparativamente con el resto de territorios, y simular un nuevo índice de competitividad observando que ocurriría si el valor de los factores más débiles fuese sustituido por el valor medio respectivo de todos los países.
4. Determinar el nivel de atracción de empresas foráneas en relación al resto de países en algunos aspectos del estudio.
5. Determinar los mayores desafíos competitivos mundiales y de Canarias para los próximos años con el fin de mejorar su competitividad internacional.

Asimismo, este trabajo se llevó a cabo utilizando como base de investigación dos informes internacionales a escala mundial que analizan y cuantifican los factores que determinan la competitividad de las naciones: el *World Economic Forum* (WEF) y el *Internacional Institute for Management Development* (IMD) y los datos recabados por la investigadora en relación a Canarias. Cada uno de estos estudios han sido descritos previamente en el capítulo uno de la tesis.

Los informes de competitividad basan sus análisis en dos tipos de datos: estadísticos o *hard*, y de percepción directiva o *soft*. Cada uno de estos datos posee una capacidad descriptiva numérica de un aspecto de la competitividad. Los datos se presentan como criterios y están clasificados en factores y en algunos, casos como puede ser el IMD, en subfactores. Un ejemplo de esta clasificación se presenta en los cuadros 3.1 y 3.2.

Actuación Económica	Eficiencia del Gobierno	Eficiencia Empresarial	Infraestructuras
68 criterios	84 criterios	60 criterios	74 criterios
Economía doméstica	Finanzas públicas	Productividad	Infraestructura básica
Comercio Internacional	Política fiscal	Mercado laboral	Infraestructura tecnológica
Inversión internacional	Modelo institucional	Mercados financieros	Infraestructura científica
Empleo	Modelo empresarial	Prácticas de gestión	Salud y medio ambiente
Precios	Educación	Impacto a la globalización	Sistemas de valores

Cuadro 3.1: Clasificación de criterios por Factores y Subfactores del informe de competitividad de IMD.

Como se puede observar en el cuadro 3.1, los datos están clasificados en cuatro factores (Actuación Económica, Eficiencia del Gobierno, Eficiencia Empresarial e Infraestructuras) y, éstos a su vez, en 20 subfactores (Ej. Empleo o Política fiscal), sumando un total de 286 criterios. Esta clasificación no coincide con las de los otros estudios de competitividad (WEF o ULPSC), ni en el número de variables utilizadas, ni en los niveles de abstracción. Asimismo, entre diferentes ediciones de un mismo informe puede incluso existir variaciones en el número de variables, en el ámbito de actuación o en la representación de estas variables o estudios que presentan. Este hecho se acentúa, si hacemos una comparación entre diferentes informes, como es el caso del WEF. A continuación se presenta los factores de competitividad según el informe del WEF.

Factores	Nº Criterios
Indicadores de actuación económica	7
Indicadores de Gobierno y política fiscal	20
Instituciones	17
Infraestructura	22
Recursos humanos	18
Tecnología	13
Finanzas	22
Apertura hacia el comercio exterior y los flujos de capital	12
Competencia doméstica	16
Estrategia y operaciones de la empresa	17
Política medioambiental	13

Cuadro 3.2: Clasificación de criterios por Factores del informe de competitividad de WEF

Como se puede comprobar, en el informe de competitividad del WEF, los criterios solo están clasificados en once factores, no en subfactores, y hacen un total de 177 criterios.

3.1.2.- Estructura de contenido del informe base

En el documento “*La Competitividad Internacional de Canarias*”, utilizado como base de esta investigación, los datos de encuesta (*soft*) se constituyen en 133 ítems y una serie de datos de clasificación que permiten la obtención de información referida al encuestado y a la empresa. A continuación, el cuadro 3.3 muestra la ficha técnica de dicha encuesta.

Procedimiento de recogida de información	<i>Encuesta personal no autoadministrada a través de un cuestionario estructurado</i>
Tipo de escala	<i>Escala de diferencial semántico de seis puntos en la investigación cuantitativa</i>
Ámbito de aplicación	<i>Comunidad Autónoma Canaria</i>
Universo	<i>Agentes sociales con experiencia y conocimiento de la realidad del Archipiélago</i>
Muestra real	<i>140 expertos</i>
Fecha realización Pretest	<i>Octubre-Noviembre 2000</i>
Fecha de realización trabajo de campo	<i>Enero-Marzo 2001</i>
Porcentaje de error muestral (nivel de confianza)	<i>± 8,4% (95,5%)</i>
Técnicas de análisis de información	<i>Análisis univariantes y bivariantes</i>

Cuadro 3.3: Ficha técnica de la investigación de naturaleza primaria (Ramos, 2000; Pág. 101)

Este documento, a su vez, está dividido en dos apartados: el primero adapta la metodología IMD al caso de Canarias, consiguiendo una estructura en sus criterios, que es descrita en el cuadro 3.4.

Actividad Económica	Eficiencia del Gobierno	Eficiencia Empresarial	Infraestructuras
48 criterios	76 criterios	44 criterios	53 criterios
Tamaño <i>Criterios 7</i>	Finanzas públicas <i>Criterios 10</i>	Productividad <i>Criterios 5</i>	Infraestructura básica <i>Criterios 12</i>
Crecimiento <i>Criterios 8</i>	Política fiscal <i>Criterios 10</i>	Costes <i>Criterios 2</i>	Infraestructura tecnológica <i>Criterios 7</i>
Riqueza <i>Criterios 6</i>	Modelo institucional- Banco Central <i>Criterios 7</i>	Mercado laboral – relaciones <i>Criterios 5</i>	Infraestructura científica <i>Criterios 16</i>
Previsiones <i>Criterios 3</i>	Modelo institucional – Eficiencia estatal <i>Criterios 10</i>	Mercado laboral – Disponibilidad de habilidades <i>Criterios 10</i>	Salud y medio ambiente <i>Criterios 13</i>
Comercio Internacional <i>Criterios 12</i>	Modelo institucional – Justicia y seguridad <i>Criterios 4</i>	Mercados financieros – Eficiencia bancaria <i>Criterios 2</i>	Sistemas de valores <i>Criterios 5</i>
Inversión Internacional <i>Criterios 2</i>	Modelo empresarial – apertura o regulaciones de mercado <i>Criterios 5</i>	Mercados financieros – Eficiencia en la Bolsa <i>Criterios 3</i>	
Empleo <i>Criterios 8</i>	Modelo empresarial – Regulaciones a la competencia <i>Criterios 4</i>	Auto-financiación <i>Criterios 1</i>	
Precios <i>Criterios 2</i>	Modelo empresarial – Regulaciones laborales <i>Criterios 3</i>	Prácticas directivas <i>Criterios 11</i>	
	Modelo empresarial – Regulaciones de los mercados de capital <i>Criterios 11</i>	Impacto a la Globalización <i>Criterios 5</i>	
	Educación <i>Criterios 12</i>		

Cuadro 3.4: Clasificación de criterios según metodología IMD

Como se puede comprobar, estos criterios, agrupados en el nivel más bajo por las áreas que representan o subfactores, y en el nivel superior por cuatro factores, hacen un total de 221 criterios. A su vez, en el segundo apartado, es decir, la adaptación de la metodología WEF para Canarias, el estudio recoge un total de 114 criterios como puede observarse en el cuadro 3.5, que se presenta a continuación.

Factores	Nº Criterios
Economía	5
Gobierno	18
Instituciones	9
Infraestructura	14
Recursos humanos	10
Tecnología	10
Finanzas	13
Apertura hacia el comercio exterior y los flujos de capital	6
Competencia	12
Estrategia y operaciones de la empresa	15
Política medioambiental	2
Nº total de criterios	114

Cuadro 3.5: Clasificación de criterios según metodología WEF

Asimismo, este informe de competitividad incluye una serie de estudios específicos sobre diferentes apartados de la competitividad, que se relacionan a continuación:

- Índice de competitividad mundial (según apartado uno)
- Índice de crecimiento de la competitividad mundial
- Índice de la competitividad mundial (según apartado uno)
- Ranking de los factores de la competitividad-factores y subfactores
- Ranking de los factores de la competitividad-factores
- Grafico de estructura de la competitividad de canarias
- Balance de competitividad – fuentes de ventaja competitiva
- Balance de competitividad – fuentes de desventaja competitiva
- Simulación del índice de competitividad
- Atractivo para la localización de actividades empresariales – industria
- Atractivo para la localización de actividades empresariales – I+D
- Atractivo para la localización de actividades empresariales – Servicios y Gestión
- Desafíos para la mejora de la competitividad – de los países
- Desafíos para la mejora de la competitividad – de Canarias

3.1.3.- Ventajas y desventajas del informe base

A continuación, y teniendo en cuenta el aspecto tecnológico, se relacionan las principales ventajas y desventajas del documento base, relacionado con la competitividad internacional de Canarias, y que son utilizadas como referencia para la definición del sistema que se propone en esta investigación. Iniciamos pues el análisis presentando las ventajas del informe base.

1. Ventajas

El informe en el que se basa este trabajo de investigación posee una serie de características o fortalezas que son importantes y lo diferencian del resto de trabajos publicados sobre Canarias. Estas fortalezas deben mantenerse y, si es posible, mejorarse con el sistema de información propuesto. A continuación, se relacionan algunos de los atributos que se pueden observar en este estudio, como pueden ser:

- *Completo*

Este informe tiene un carácter completo porque cubre un amplio abanico de variables descriptivas, además de conocimiento empresarial, de la realidad de una región, en este caso, Canarias, que permite obtener una visión holística de la región en todas sus facetas, tanto económicas, como sociales, gubernamentales etc...

- *Internacional*

Este informe tiene un carácter local y, a su vez, internacional, ya que permite colocar Canarias en niveles de comparación mundial, más allá de las fronteras europeas.

- *Integrador*

Este informe es integrador, porque integra información *hard* con información *soft*, información *local* con información *global*, información de diferentes sectores sociales, información de muchos países, información de varios informes internacionales de competitividad, información de muchas fuentes

diferentes etc...

- *Validado*

Este informe, no sólo ha validado los datos que manejó a través de la verificación de sus fuentes y a través del equipo de investigación, sino que identificó errores, confirmados por las fuentes, utilizadas en los dos informes que le precedieron (WEF e IMD). Algo insólito, teniendo en cuenta el nivel de precisión, de validación y de experiencia (más de diez años de existencia) que tienen estos documentos con proyección internacional.

- *Innovador* (región versus países)

Este estudio es innovador, porque nunca antes se había hecho un estudio como este en Canarias, y según se tiene constancia, es la primera vez que se realiza una comparativa de una región europea, como es el caso, con más de 50 países. Esta característica ha provocado el interés internacional de muchos investigadores por permitir la conexión de la dimensión micro y macro en el área de economía.

- *Único*

Este trabajo es único, ya que no existe en la actualidad ningún trabajo de investigación sobre Canarias, tan preciso, que contenga un estudio comparativo con más de 300 variables para cada país (ver Anexo II), de una lista de más de 50 países (ver Anexo I), y con la precisión que tiene el estudio sobre "*La Competitividad Internacional de Canarias*".

Una vez, presentadas las ventajas encontradas en el estudio base, a continuación se analizan las principales debilidades encontradas en dicho estudio, y que las TICs permitirán eliminar.

2. Desventajas

El informe base presenta una serie de desventajas que deben tenerse en cuenta. Muchas de estas debilidades están ligadas al hecho de que el informe se presenta como los documentos físicos, con las implicaciones que esto conlleva. Entre las debilidades más destacadas se encuentran:

- Presentación

La información, como se comentó en el párrafo anterior, se presenta en un documento físico. Esto trae algunas implicaciones muy importantes, desde el punto de vista del usuario así como de la información en sí. Entre estas implicaciones se incluyen:

- *Presentación única*, es decir, la información se presenta de una sola forma, en el documento y, en muchas ocasiones, el usuario podría precisar de otra forma de presentación que requeriría un proceso laborioso.
- *Presentación unipersonal*, es decir, sólo el usuario que posea la publicación tendrá acceso a la información que en ella se contenga y sólo una persona a la vez, podrá tener acceso a la información.
- *Presentación sólo para personas*, es decir, si la información tuviera que ser transmitida a un programa informático, requeriría de un proceso bastante costoso, con el fin de que éste pudiera procesarla.

- Accesibilidad a la Información

Otro de los aspectos básicos que presenta ciertas desventajas del informe, es la accesibilidad que tiene un usuario. Entre estas desventajas se encuentran:

- *Accesibilidad a la publicación*, es decir, son muy pocas las personas que tienen acceso al libro físico, podríamos decir que es exclusivo. Esto limita las posibilidades a cualquier usuario potencial de enriquecerse de la transferencia de información y conocimiento que de este se pueda derivar.
- *Accesibilidad a la información*, es decir, la estructura estática y perenne que tiene la publicación puede provocar que si no se es consciente de donde se encuentra un simple dato, éste sea ignorado por un usuario, e incluso que sea bastante difícil relacionar datos que pueden estar muy interrelacionados para un usuario.
- *Accesibilidad a las fuentes de información*, es decir, si un usuario quiere tener acceso a la fuente de información de algunos de los criterios que

en el informe se presentan, no tiene forma de identificarlas.

- Portabilidad de la información

En muchas ocasiones, es de suma importancia poder extraer datos o estudios presentes en el informe, con el fin de poder analizar o elaborar nuevos análisis, o simplemente integrarlos en otras aplicaciones de interés y que podrían ser claves para la toma de decisiones para un usuario u organización. Esta tarea es bastante difícil realizarla, con un documento físico como el que se analiza en esta tesis.

- Actualización y evolución

Los datos, el modelo general y los estudios que están incluidos en este informe, tienen una naturaleza dinámica, aunque el libro físico no permite este dinamismo, si no es a través de nuevas ediciones futuras, como es el caso de los otros informes.

- Corrección de errores

La experiencia enseña que aun teniendo altos niveles de control de calidad en este tipo de informes, sus características podrían provocar la existencia de errores menores, que serían difíciles de enmendar debido al carácter estático del documento físico.

- Verticalidad de los datos (no tiene histórico)

Este informe presenta una descripción de Canarias en el marco internacional en un momento determinado de la historia de Canarias, podríamos decir que es como una foto de Canarias donde se analizan aspectos muy importantes sobre ésta. Pero este informe carece de datos históricos, por ejemplo de los criterios que se presentan, que podrían ser de gran ayuda para realizar análisis de tendencias, predicciones etc...

- Compacto (no modular)

El carácter compacto de este informe puede ser un inconveniente ya que no permite intercambiar secciones o apartados de este que podrían enriquecer el nivel de información y conocimiento que ya aporta.

- Cerrado (No abierto o Integrador)
El hecho de que el informe tenga la característica anterior (compacto – no modular), implica que carece de la *propiedad de ser abierto*, es decir, de poder ser integrado o conectado con otros informes o sistemas de información.
- Conocimiento estático
Toda la información y conocimiento disponible en este informe están soportados sobre un medio estático, lo que imposibilita el análisis e inferencia artificial del conocimiento, es decir, a través de motores de inferencia u otro procedimiento artificial disponible en la actualidad.
- Dependiente
El carácter de objeto físico estático, hace necesaria la mano humana e imposible la automatización autónoma de tareas que permitan realizar operaciones que provoque el incremento de información y conocimiento, todavía sin descubrir y que permanece oculto en el contenido del informe.

Con el objetivo de fortalecer las ventajas y superar las debilidades del informe base, a continuación se describe la metodología del modelo tecnológico propuesto.

3.2. METODOLOGÍA DE DESARROLLO DEL SISTEMA

Para realizar el desarrollo del sistema propuesto se debe considerar la tipología del desarrollo y el objetivo que se propone, para poder entender qué tipo de metodología o metodologías, son necesarias para el desarrollo del sistema avanzado de gestión del conocimiento, aplicado a la competitividad internacional. El sistema debe cumplir el objetivo descrito en el próximo subapartado de este documento. Por esta razón, se debe considerar la metodología de la parte del desarrollo del sistema, y por otro, la metodología para la implementación de modelos ontológicos basados en el sistema propuesto.

3.2.1.- Introducción a la metodología

El desarrollo del sistema aplicado a la competitividad internacional de los territorios se va realizar en su primera parte, siguiendo un ciclo de vida utilizan un modelo clásico de cascada (*waterfall model*), propuesta inicialmente por Winston Royce (1970). El *waterfall model*, como se puede comprobar en la figura 3.1, está compuesto de una fase de análisis, una fase de diseño, una fase de implementación o desarrollo, una fase de depuración, una fase de implantación y por último, una fase de mantenimiento. Aunque el modelo clásico de cascada presenta problemas de adaptación a muchas situaciones reales, se reconoce su importancia como modelo base para otros modelos de ciclo de vida más efectivos, como los propuestos en esta tesis, para la segunda parte del desarrollo del sistema. Sin embargo, la razón por la que se ha elegido el *waterfall model* estriba en la naturaleza misma del tipo de aplicación, ya que sus objetivos iniciales permanecen claros y estables en el medio plazo.

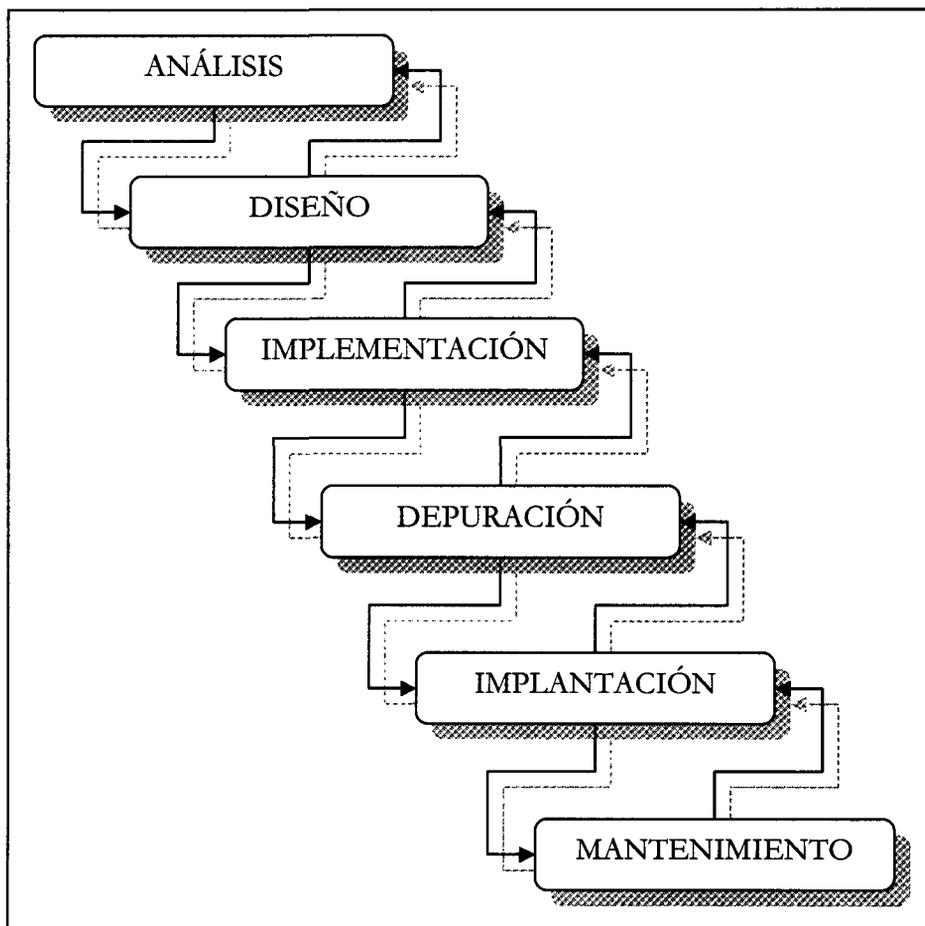


Figura 3.1: Fases del modelo de desarrollo en cascada (*waterfall*)

Como el objetivo de esta investigación consiste específicamente en definir un sistema para crear y gestionar, de forma eficiente, el conocimiento sobre la competitividad de los territorios, el desarrollo y gestión de ontologías, y sus posteriores aplicaciones relacionados directamente con la creación y gestión del conocimiento de la competitividad de los territorios, se encuentran fuera del objetivo de esta investigación, y por lo tanto no serán desarrollados en la tesis. Asimismo, cabe añadir que la mayor dificultad de este tipo de aplicaciones está en la creación en sí del sistema avanzado de gestión del conocimiento, ya que servirá como plataforma de futuros desarrollos y gestión del conocimiento sobre la competitividad internacional. Además, el sistema que se propone, abre un abanico de posibilidades para estudios de competitividad de empresas, sectores, regiones o bloques regionales, como la Unión Europea. Y, más

importante, pretende contribuir a cubrir la laguna existente en la actualidad en las investigaciones sobre la gestión del conocimiento en las organizaciones.

Con el fin de simplificar los aspectos más técnicos del desarrollo del sistema avanzado de gestión del conocimiento se omitirán algunos apartados como planning, estudios de viabilidad, entre otros, menos significativos para el objetivo descriptivo de este capítulo. Por consiguiente, a continuación se presentan los apartados más significativos del desarrollo del sistema.

Análisis del Sistema

El análisis del informe base "*La Competitividad Internacional de Canarias*", considerando sus fortalezas y sus debilidades, y las reuniones de trabajo mantenidas con la responsable del estudio, han permitido trazar las necesidades del nuevo sistema avanzado de gestión del conocimiento aplicado a la competitividad de los territorios. Por ello, en este apartado se describe algunas de las tareas básicas de la fase de análisis, como son: objetivo general del sistema, requerimientos del sistema y propuesta del nuevo sistema.

Objetivo general del sistema

“Definir y desarrollar desde un punto de vista teórico-conceptual, con una perspectiva tecnológica, un sistema software basado en los modelos de competitividad internacional, que cubra y extienda sus fortalezas, que elimine sus debilidades, y que funcione como plataforma modular, autónoma y abierta para el desarrollo de bases de conocimiento actualizable, y que permita, a su vez, el descubrimiento automático de nuevo conocimiento para la competitividad internacional de los territorios”

Requerimientos

Una vez realizado un análisis detallado del objetivo general del nuevo sistema y las fortalezas y debilidades del informe en el que se fundamenta, se relacionan los requerimientos del sistema en el cuadro 3.6. Como se puede comprobar, dieciséis son los requerimientos que el sistema propuesto pretende cumplir, para satisfacer el objetivo de esta investigación, y alcanzar de esta forma un mayor nivel de conocimiento sobre la competitividad internacional de los territorios.

Entre estos requerimientos, se incluyen aspectos tales como, que el sistema sea completo, internacional, integrador, validado, innovador, único, modular, abierto, autónomo, entre otros. A su vez, cada uno de estos requerimientos se define por su importancia y complejidad, cuyo nivel difiere en función del tipo de requerimiento.

REQUERIMIENTOS DEL SISTEMA	
Descripción del requerimiento	
1.	<i>Completo:</i> Debe permitir cubrir un número indeterminado de variables descriptivas de la competitividad de regiones
2.	<i>Internacional:</i> Debe permitir la comparativa internacional de cualquier región de la que posea información
3.	<i>Integrador:</i> Debe permitir gestionar información y conocimiento de diferentes tipos de fuentes de información.
4.	<i>Validado:</i> Debe facilitar la validación de la información y el conocimiento que contenga o descubra.
5.	<i>Innovador:</i> Debe permitir realizar las comparativas presentes en el informe en el que se basa, y además permitir el descubrimiento de nuevo conocimiento.
6.	<i>Único:</i> Debe mantener la singularidad como sistema para el objetivo que debe cumplir.
7.	<i>Presentación:</i> Debe ser variable en la presentación de su información, permitir el acceso multiusuario y permitir el acceso a la información y conocimiento desde programas de software.
8.	<i>Accesibilidad:</i> Debe permitir la accesibilidad no dependiendo de un soporte físico como un libro, variable en función del tipo de usuario y a las fuentes disponibles de información de la información.
9.	<i>Portabilidad:</i> Debe permitir exportar e importar la información que requiera o contenga.
10.	<i>Actualización y evolución:</i> Debe permitir la actualización de los datos y la evolución del modelo o modelos que contenga.
11.	<i>Corrección de errores:</i> debe facilitar y simplificar la corrección de errores.
12.	<i>Horizontalidad:</i> Debe permitir el estudio de información longitudinal para la creación de estudios tales como predicciones, tendencias etc.
13.	<i>Modular:</i> Debe permitir el acceso modular a la información, a los modelos y al conocimiento.
14.	<i>Abierto:</i> Debe permitir la integración de sus modelos, su información y su funcionalidad en otros sistemas.
15.	<i>Conocimiento dinámico:</i> Debe permitir la evolución del conocimiento a través del tiempo.
16.	<i>Autónomo:</i> Debe permitir el procesamiento de datos y el mantenimiento autónomo del sistema.

Cuadro 3.6: Requerimientos del sistema propuesto.

3.2.2.- Metodologías para la fase de implementación

En una segunda parte, y una vez establecido el modelo del sistema avanzado de gestión del conocimiento aplicado a la competitividad internacional de los territorios, se ejecutarían las fases de depuración, implantación y mantenimiento, para la que se propondría una metodología más dinámica, ágil y apropiada, para el estudio de la competitividad, es decir, una metodología que facilite el rápido desarrollo de sistemas ontológicos específicos para la competitividad, con soporte en *e-services* y agentes. Por esta razón, sería ideal hacer uso de alguna de las metodologías que cumplen los principios del Agile Manifesto (Beck et al., 2001), para el desarrollo rápido de aplicaciones. Entre estas metodologías se encuentran (AgileAlliance, 2003):

- *AD - Agile Database Techniques*: Esta metodología, promovida por Scott W. Ambler, tiene como meta aportar a los profesionales de las TICs un amplio abanico de soluciones, que faciliten el trabajo conjunto en los aspectos relacionados con los datos de los sistemas informáticos.
- *AM - Agile Modeling*: Es una colección de principios, valores y prácticas para el modelado de software, que pueden ser aplicados en el desarrollo de proyectos de software, de forma efectiva y poco costosa.
- *ASD - Adaptive Software Development*: Esta metodología esta dirigida por la creencia de la adaptación continua. En ASD, el ciclo de vida estático es sustituido por un ciclo de vida especulativo, colaborativo y de aprendizaje. Es un ciclo de vida dedicado al aprendizaje continuo y orientado al cambio, la reevaluación, la mirada hacia un futuro incierto, y la intensa colaboración entre desarrolladores, directivos, y clientes.
- *Crystal*: Esta es una familia de metodologías de desarrollo de software, adaptable, ultraligera y “accionada por el humano”. Es decir, tienen como meta conseguir el éxito del proyecto, extendiendo las capacidades de los profesionales envueltos en el proyecto, reduciendo la burocracia y el trabajo en papel, y reduciendo el

volumen de trabajo incluso en pequeños proyectos.

- **FDD - *Feature Driven Development***: Esta metodología empieza con la creación de un modelo de los objetos del área de estudio, en colaboración con los expertos del área. Utilizando la información de la actividad del modelado y de cualquier otra actividad requerida que haya tenido lugar, los desarrolladores continúan creando la lista de atributos. Posteriormente, se realiza un plan aproximado y se asignan responsabilidades. A continuación, pequeños equipos formados, de forma dinámica, desarrollan los atributos repitiendo el diseño y construyendo iteración con una duración no superior a dos semanas y, en muchas ocasiones, en más cortos periodos de tiempos.
- **DSDM - *Dynamic Systems Development Method***: El DSDM, más que un método, es un marco de trabajo. El proceso del proyecto tiene siete fases, que se repiten durante el ciclo de vida del proyecto, haciendo el DSDM un método incremental e iterativo. Las fases son:
 1. Fase pre-proyecto: asegura que sólo los proyectos correctos son iniciados y son activados correctamente.
 2. Estudio de viabilidad: se valora si el DSDM es el enfoque adecuado para el proyecto y una definición del problema, se valora los costes y la viabilidad técnica de producir un sistema.
 3. Estudio del negocio: El primer foco de atención está en los procesos del negocio y sus necesidades de información. A través de una serie de talleres de trabajo se consigue un rápido consenso como prioridades del desarrollo.
 4. Iteración del Modelo Funcional: refinamiento de los aspectos del sistema basados en el negocio.
 5. Iteración de Diseño y Construcción: es donde se fragua el sistema.
 6. Implementación: cubre desde el entorno de desarrollo al entorno operativo.
 7. Post-proyecto: Esta fase cubre la solución operando de forma eficiente.
- **Lean Software Development**: El pensamiento *Lean Thinking* ha tenido un tremendo impacto en la economía global y ha transformado la forma en la que muchas industrias operan para producir riqueza, eliminando desechos, permitiendo cualquier inversión en personas y recursos productivos dirigidos a producir más

valor, mientras se suministra a los participantes, trabajos significativos y que les satisfagan. Este planteamiento ha sido aplicado al desarrollo de software, y permite guiar a los equipos de desarrollo de software indicando qué y cuando hacer las tareas.

- **Scrum:** Considerado como un proceso ágil y ligero que puede ser utilizado para gestionar y controlar el desarrollo de software y productos. Scrum genera los beneficios producidos por el desarrollo ágil, con las ventajas de una simple implementación.
- **TDD - *Test-Driven Design*:** Se presenta como el oficio de producir test automatizados para la producción de código, dirigiendo el diseño y la programación. Para cada pequeña parte funcional en la producción de código, se desarrolla, en primer lugar, un test que especifica y valida la función que realizará dicho código. Y, de esta forma, se genera el código que exactamente se necesita para pasar dicho test.
- **XBreed:** Es el resultado de desarrollar múltiples aplicaciones y componentes compartidos, de forma tan rápida como humanamente sea posible. XBreed es el producto de la combinación de SCRUM, XP y las ideas Alexanderianas.
- **XP - eXtreme Programming:** Es una disciplina de desarrollo de software basada en los valores de simpleza, comunicación, retroalimentación y coraje. Funciona agrupando a todo el equipo junto a simples prácticas, con suficiente retroalimentación y capacitándolo para ver donde están y adaptar las prácticas a su situación específica.

3.3. PROPUESTA TECNOLÓGICA

La propuesta representada en la figura 3.2, intenta cubrir, en la mejor medida posible, los requerimientos expuestos en el cuadro anterior. Por consiguiente, se define en base al objetivo del sistema y a las necesidades existentes, cubriendo los siguientes

principios:

1. Modularidad (basado en componentes).
2. Integración de componentes.
3. Gestión de conocimiento basado en ontologías.
4. Inferencia de conocimiento mediante ontologías.
5. Autonomía de tareas (basado en modelos de agentes).
6. Creación de nuevo conocimiento (*Knowledge Discovery*).
7. Fuentes de datos estándar.

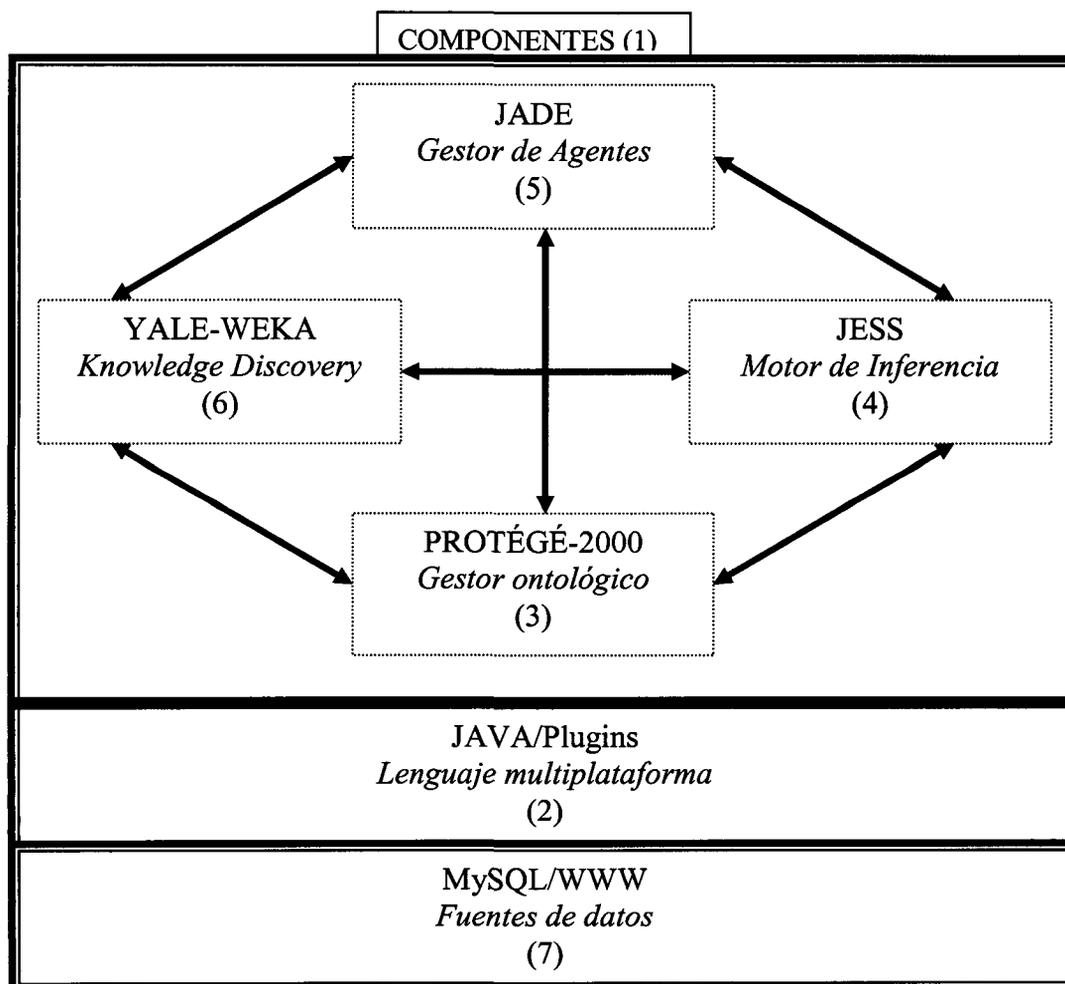


Figura 3.2: Propuesta del Sistema

La figura 3.2 representa los componentes modulares que cumplen cada uno de los principios relacionados anteriormente. Así, la modularidad se consigue gracias al uso de las diferentes herramientas tecnológicas de última generación (Protégé-2000, Jess, JADE, Yale-WEKA), que en nuestro sistema avanzado de gestión del conocimiento aplicado a la competitividad de los territorios, actúan como componentes integrables, creando una sinergia debido a la combinación de sus capacidades funcionales. El principio relacionado con la integración de los componentes se cumple gracias al uso del lenguaje multiplataforma JAVA, formatos Web para los ficheros de trabajo y software adicional que facilita la integración entre algunos de los componentes (*plugins*).

Asimismo, el principio de gestión de conocimiento basado en ontologías se cumple a través del uso de la herramienta de desarrollo Protégé-2000. Por otro lado, el principio de inferencia de conocimiento mediante ontologías se cumple mediante el uso de motor de inferencias Jess. El principio de autonomía de tareas (basado en modelos de agentes) se cumple con el uso de la herramienta de desarrollo de sistemas de agentes JADE. Y en lo que se refiere al principio de creación de nuevo conocimiento (*Knowledge Discovery*), este se cumple gracias al uso del entorno para el desarrollo de experimentos de *machine learning*, YALE-WEKA. Y finalmente, para el principio de fuentes de datos estándar se utiliza la base de datos MySQL y la Web. Cada una de estas tecnologías son descritas posteriormente, donde se puede comprobar que son estables, de última generación, abiertas, integrables y especializadas, como se propuso en la introducción de este capítulo, su integración puede permitir, de una forma innovadora, sinergias nunca antes conseguidas, y que se perfila en esta investigación.

Por otro lado, cabe añadir que cada uno de estos principios del sistema avanzado de gestión del conocimiento aplicado a la competitividad de los territorios, ha requerido, por parte del investigador, la realización de un trabajo preciso, exhaustivo y detallado en la búsqueda de las herramientas que permitieran cumplir tales principios. En este sentido, durante el transcurso de esta investigación fueron muchas las herramientas, de difícil localización, que tuvieron que ser identificadas, analizadas y comprobada su idoneidad por si mismas, y su capacidad de integración entre si. Además, las capacidades especializadas de cada una de estas herramientas tecnológicas y las

sinergias funcionales que se crean con su integración, son idóneas en la aplicación para el objetivo de esta investigación.

Con el fin de facilitar la comprensión de las posibilidades que ofrece esta integración, probada ya su viabilidad técnica, a continuación se presenta una breve descripción de cada uno de los componentes principales del sistema.

3.3.1. Componentes básicos del sistema

Las herramientas propuestas para este sistema han sido seleccionadas después del estudio de un gran número de herramientas de última generación, o en vías de desarrollo, de centros de investigación de relevancia internacional, tales como el *Sandia National Laboratories, Livermore, Canada* o el *Stanford Medical Informatics de la universidad de Stanford*. En esta selección, se ha considerado los requerimientos del nuevo sistema, la consistencia de las distintas herramientas, la confianza de los grupos de desarrollo, el nivel de integración entre todas las herramientas etc...

Cada una de las herramientas propuestas, cubre un área fundamental en el nuevo sistema, pero asegura la interdependencia entre cada una de las herramientas de tal forma que puedan ser sustituidas en el futuro, por otras que puedan ofrecer una mejor solución para cada situación.

3.3.2. Protégé-2000

Protégé-2000 es un entorno de desarrollo de ontologías basado en la idea de Mark Musen, de una meta herramienta de sistemas basados en el conocimiento, propuesta en 1987, y orientada a la creación de herramientas para la adquisición de conocimiento de planificación médica en el *Stanford Medical Informatics* (GENNARI, MUSEN, FERGERSON, GROSSO, CRUBÉZY, ERIKSSON, NOY y TU 2002). Protégé-2000, se ha convertido en una herramienta de software integrado, dirigida a desarrolladores de sistemas y expertos en áreas de conocimiento, para la creación de bases de conocimiento. A nivel general, se podría decir que Protégé-2000 permite

básicamente:

- 1.- La creación de ontologías de área de conocimiento
- 2.- Personalizar formularios para la adquisición de conocimiento
- 3.- Introducir nuevo conocimiento.

Esta herramienta tiene tres posibles formas de acceso, la primera como aplicación software, facilitando al usuario toda la funcionalidad y base de conocimiento de la que dispone; la segunda como plataforma, integrando un conjunto de herramientas de diferentes equipos de desarrolladores y facilitando al usuario el acceso a estas utilidades externas; y, por último como librería, dando acceso desde otras aplicaciones externas a las funciones y a la base de conocimiento.

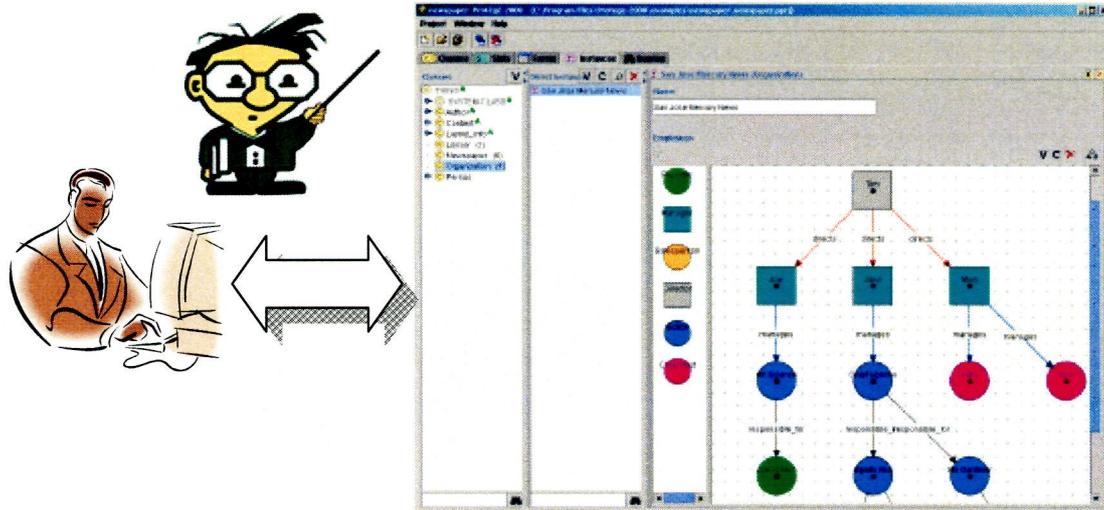


Figura 3.3: Protégé-2000 en su primera forma de acceso

(Adaptado de <http://protege.stanford.edu/index.html>)

En la figura 3.4 se presenta un ejemplo de Protégé-2000, con componentes externos (plugins) como puede ser JessTab, un componente que permite la integración del motor de inferencia Jess con Protégé-2000; TopicMap, otro componente que facilita la exportación de las ontologías a XTM; Ontology Bean Generator Tab, un plugin que genera código fuente Java, representando ontologías que pueden ser utilizadas por agentes desarrollados desde Jade (figura 3.5).; etc...

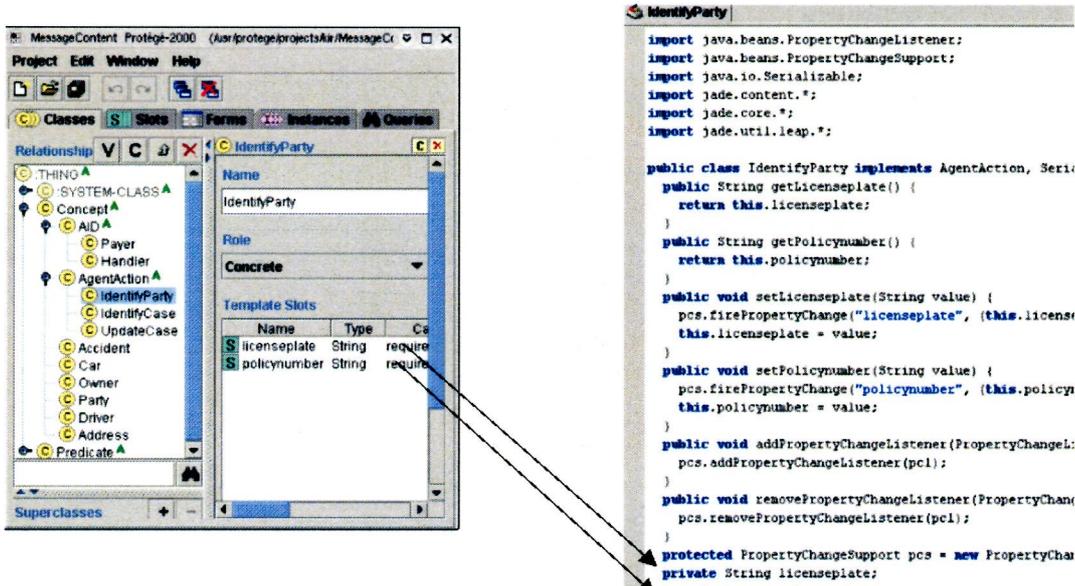


Figura 3.4: Protégé-2000 con el plugin Ontology Bean Generator
(Fuente: <http://gaper.swi.psy.uva.nl/beangenerator/content/main.php>)

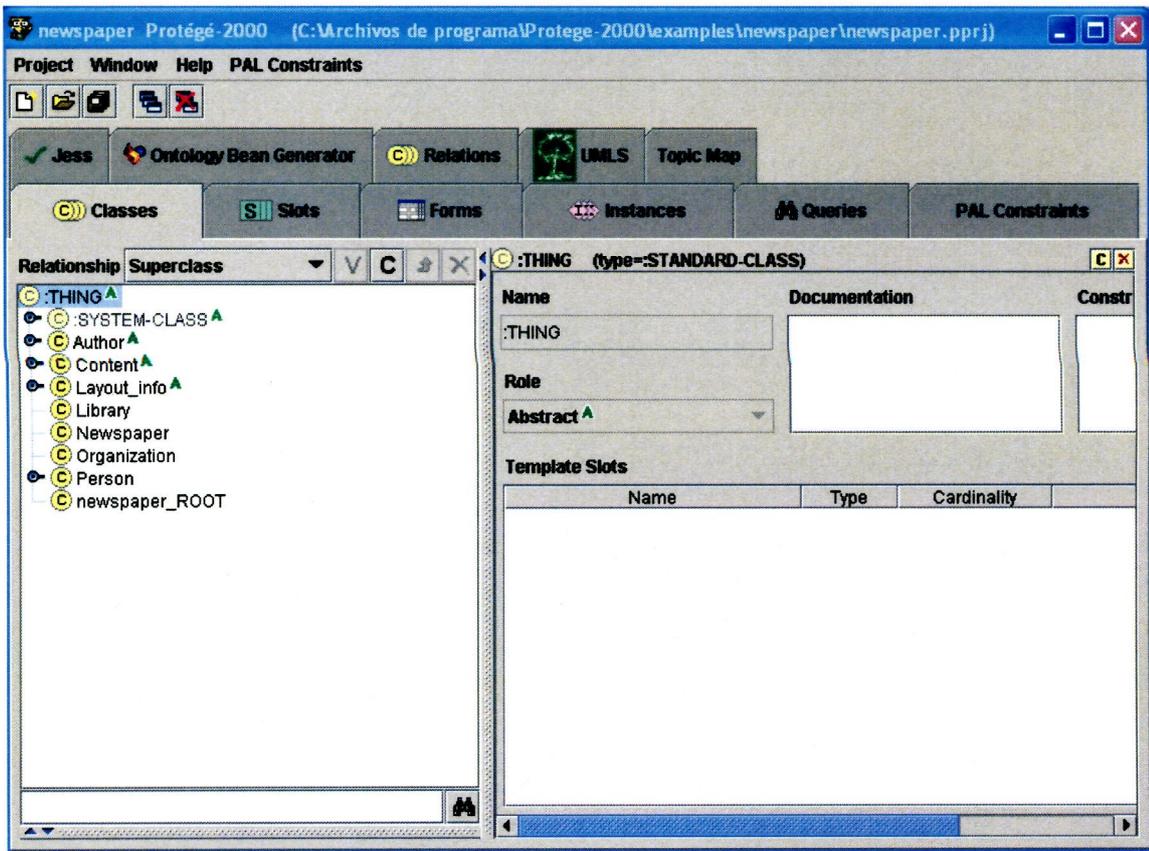
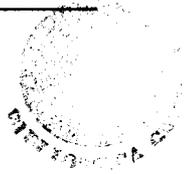


Figura 3.5: Protégé-2000 en su segunda forma de acceso



Protégé-2000, en su tercera forma, como librería, permite acceder a toda su funcionalidad, desde otras aplicaciones desde el fichero librería suministrado en lenguaje Java (Protégé-2000 API). Esto es posible gracias a que este software ha sido desarrollado en JAVA y se presenta en forma de librería JAVA. A continuación, en la figura 3.6, se muestra una representación gráfica mostrando esta interacción entre diferentes aplicaciones y la librería de Protégé-2000.

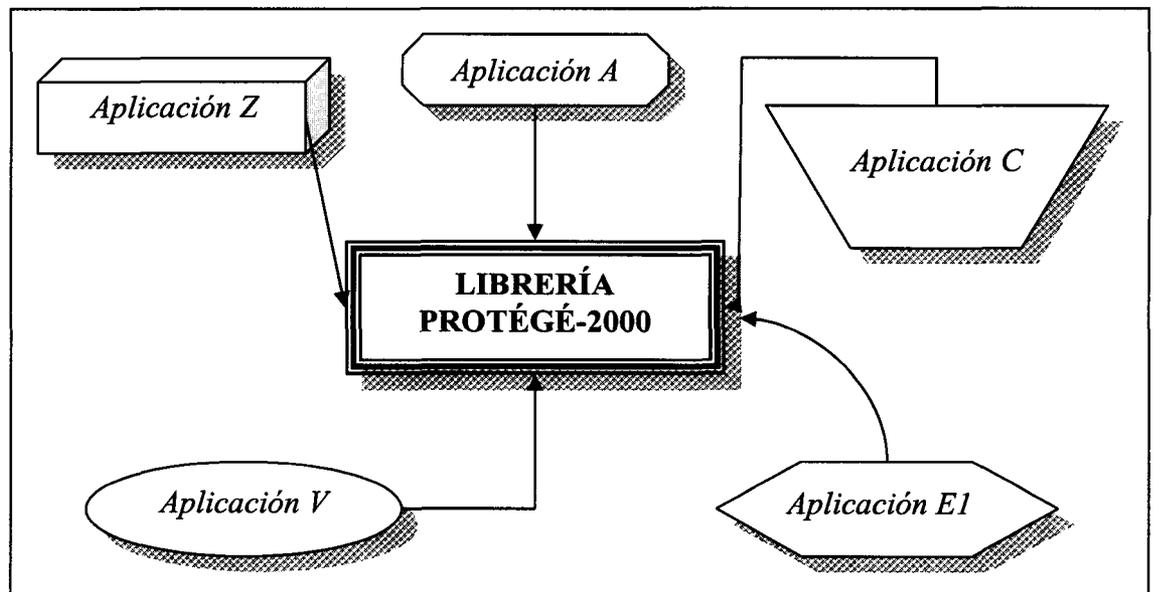


Figura 3.6: Representación de Protégé-2000 en su tercera forma de acceso

Este software es abierto (open-source) lo que promueve su evolución continua y la participación de muchos equipos de desarrollo e investigación a nivel internacional. En su versión 1.9, Protégé-2000 permite la gestión de modelos ontológicos expresados en ficheros de texto estándar, Base de datos JDBC, DAML+OIL, RDF Schemas, XML DTD y XML Schema (véase figura 3.7).



Figura 3.7: Cuadro de Diálogo de entrada de Protégé-2000 en su primera forma de acceso

3.3.3. JESS

Jess (Java Expert System Shell) es una herramienta y lenguaje script para la implementación de sistemas expertos basados en reglas, y ha sido desarrollado íntegramente bajo el lenguaje multiplataforma JAVA de Sun Microsystems, por Ernest Friedman-Hill, en *Sandia National Laboratories*, Livermore, Canada. Este lenguaje fue inspirado en otra popular herramienta para el desarrollo de sistemas basados en reglas, denominado CLIPS (C Language Integrated Production System; Relay, 2003), de bastante aplicación en desarrollos tecnológicos gubernamentales, industriales y del mundo académico, y originado en 1984 en el Jonson Space Center de la NASA. Esta inspiración queda patente, cuando se observa la compatibilidad de los programas desarrollados en CLIPS sobre Jess. Asimismo, Jess permite la inferencia de conocimiento explícito representado en forma de reglas declarativas. Jess es pequeño y rápido, y basa su procesamiento de reglas en un algoritmo de comparación de patrones llamado Rete (FORGY 1989), desarrollado por Charles Forgy de la Universidad Carnegie-Mellon. A su vez, la relación que mantiene CLIPS y Jess con el lenguaje de programación LISP (List Processing; Steele, 1990), queda reflejada en las similitudes en sintaxis muy representativas de LISP. En el siguiente ejemplo, se representa la expresión *Madrid es la capital de España*.

(capital-de Madrid España)

Cuadro 3.7: Ejemplo de una *fact* expresada en LISP

Las aplicaciones en las que Jess ha sido utilizado son muy variadas, ejemplos de estas aplicaciones se pueden encontrar en el razonamiento automatizado de documentos legales sobre XML (Leff, 2001), en un sistema experto difuso para la WWW (Sedbrook, 1998), en la automatización de contratos comerciales mediante agentes (Grosf y Poon, 2003), en un sistema de monitorización de riesgos financieros mediante agentes inteligentes (Wang *et al.*, 2002), en sistemas multiagentes para mensajería personal (Meech *et al.*, 2000), en el razonamiento dinámico de sistemas basados en el conocimiento con reglas de producción difusas (Song y Lee, 2002), etc..

Jess posee un conjunto de extensiones desarrolladas por otros equipos de trabajo, que amplifican, en gran medida, las capacidades de esta herramienta. Entre estas extensiones cabe destacar: JessTab (véase figura 3.8), un plugin para Jess que permite su integración al entorno Protégé-2000 (Ericsson, 2002); FuzzyJess, una librería desarrollada en JAVA por el *Institute for Information Technology del National Research Council of Canada's*, que permite la manipulación de conceptos difusos e inferencia difusa; y por último, XML Integration, que permite la conversión de documentos escritos en XML, en *facts* de Jess.

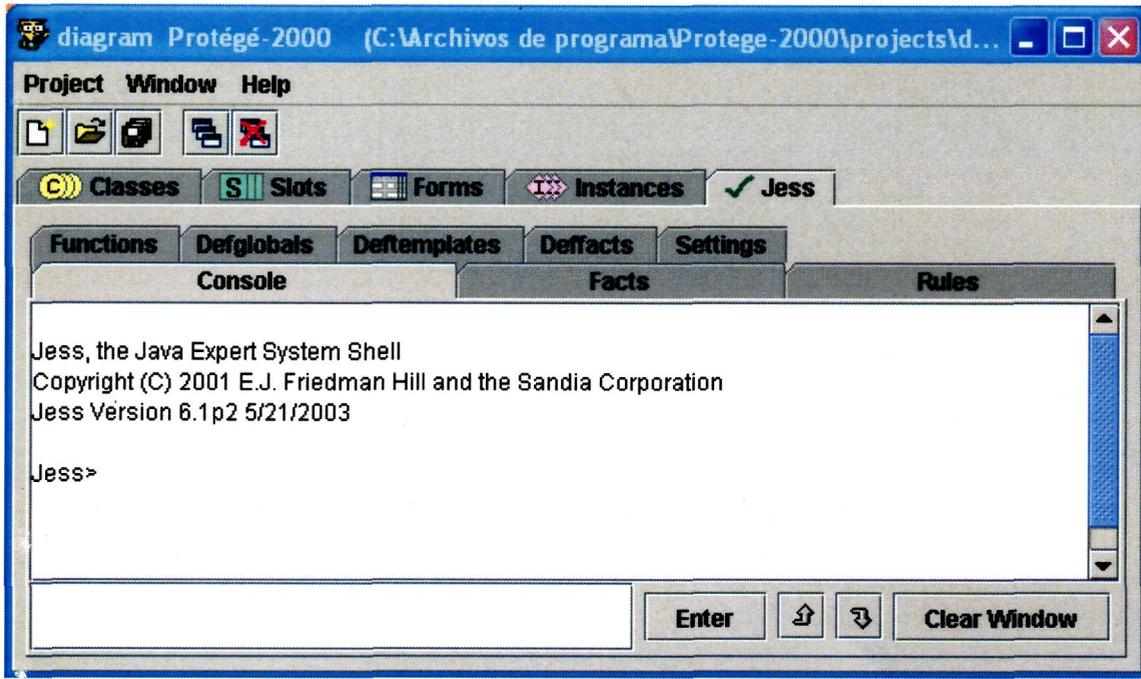


Figura 3.8: Ventana de la consola de Jess en el entorno de Protegé-2000

3.3.4. JADE

JADE (Java Agent DEvelopment Framework), es un entorno de desarrollo enteramente en JAVA, que facilita la implementación de sistemas multiagentes a través de un conjunto de herramientas software, y se argumenta, que JADE cumple la especificación del estándar FIPA (Foundation for Intelligent Physical Agents). La plataforma puede ser distribuida a través de ordenadores con diferentes sistemas operativos, y el control de los agentes y las plataformas puede realizarse de forma remota, a través de un GUI (véase figura 3.9). La configuración de los agentes puede modificarse en tiempo real y moverlos entre ordenadores. Uno de los últimos avances de este entorno ha sido su integración al proyecto europeo LEAP (Lightweight Extensible Agent Platform, IST-1999-10211), permitiendo el desarrollo de servicios móviles para JAVA.

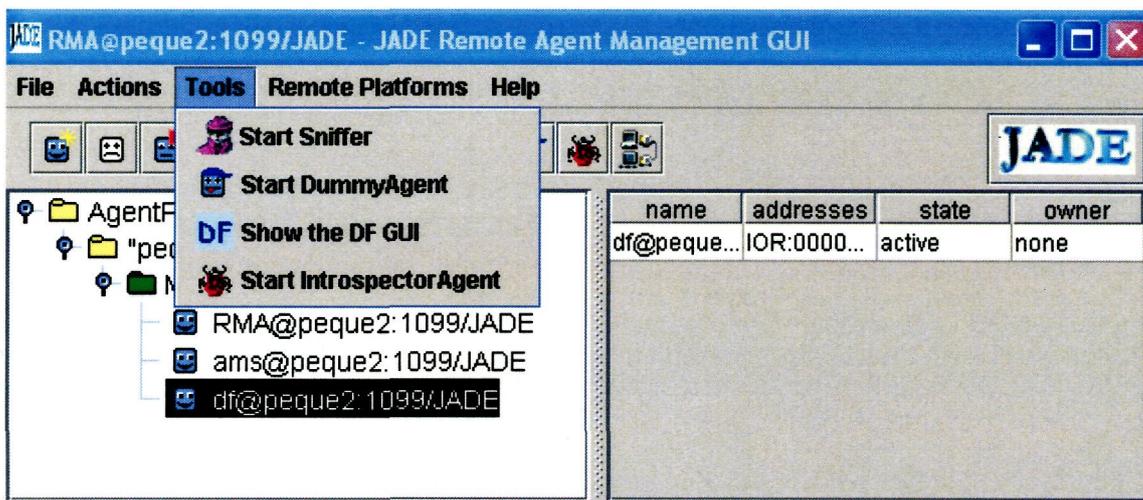


Figura 3.9: Interfaz gráfica de usuario de JADE.

Son muchos las organizaciones públicas y privadas que utilizan JADE a nivel internacional, algunas de estas se relacionan a continuación:

- CSELT, NHK, Imperial College de Londres y IRST dentro del WP A1 FACTS, para crear un Sistema Audio-Visual de Entretenimiento basado en agentes.
- BT, FT, y Broadcom, dentro del WP A3 of FACTS, para crear un Asistente Personal de Viaje.
- La Universidad de Helsinki, en el marco del proyecto Finlandés MONADS.
- CSELT, KPN, y Starlab, dentro del marco del proyecto DICEMAN, para crear un mercado basado en agentes para el comercio.
- CSELT, dentro de un proyecto interno para la creación de un Agente para Planificación de Reuniones y un Agente Organizador de Películas de Cines, para un dispositivo iPaq.
- Algunos proyectos internos de la Universidad de Parma.

- La Universidad de Brescia, para crear servicios universitarios basados en agentes.
- Algunos proyectos internos de la Universidad de Bologna.
- Sonera, quien está evaluando JADE frente a otros entornos de desarrollo basados en agentes.
- INRIA (Instituto nacional Francés de investigación en Informática y Control), Nice-Sophia-Antipolis, equipo de investigación ACACIA, y la agencia ATOS Sophia Antipolis, dentro del marco del proyecto Europeo CoMMA, para el desarrollo de un sistema de información multi-agente para ayudar en la consulta de la memoria corporativa basada en XML.
- Imperial College de Londres, Philips, Queen Margaret University College y Domus Academy han utilizado JADE dentro del proyecto WP4 Living Memory (LiMe), para crear un sistema multi-agente para comunidades conectadas.
- The Business & Technology Research Laboratory de La Universidad de Newcastle (Australia), están evaluando JADE para aplicaciones de eBusiness, utilizando Java Server Pages.
- El proyecto LEAP está dirigido al desarrollo de infraestructuras abiertas y servicios, para el soporte de empresas móviles y dinámicas. Para este propósito, en el proyecto se está desarrollando un nuevo entorno ligero runtime para JADE, que puede ser aplicado en Internet y entornos inalámbricos. El consorcio LEAP está compuesto de los siguientes miembros: Motorola (coordinador), ADAC, Broadcom, BT, TILAB (ex CSELT), Universidad de Parma, y Siemens.
- Con el proyecto *Knowledge on Demand* (KOD), Proyecto IST 12503, se ha diseñado y se está construyendo, una plataforma multi-role de aprendizaje personalizado utilizando agentes de colaboración como los componentes de

integración y para la función de búsqueda Knowledge Package (KP) del sistema. Se utilizan diferentes tecnologías existentes, basada en las especificaciones ínter operativas de aprendizaje, para otros aspectos del proceso de personalización.

- El Multi-Agent System Group (GruSMA) de URV (University Rovira i Virgili in Tarragona, Spain), recibió dos premios en la competición de la AgentCities Agent Technology Competition. Su sistema llamado "Deployment of agent-based health care services", ha sido desarrollado íntegramente con Jade 2.61 y el plugin de seguridad Jade-S.
- Pellucid (Una plataforma para empleados públicos móviles) es un proyecto de investigación y desarrollo dentro del programa IST de la Comisión Europea, que esta utilizando JADE para el desarrollo de una plataforma flexible de software muti-agente orientada a la gestión de conocimiento para asistir a los trabajadores móviles de niveles medios y superiores, de las organizaciones del sector público.

El desarrollo de servicios Web (e-services), denominado por algunos investigadores, el nuevo paradigma de los negocios en el entorno electrónico (Rush y Kannan, 2003), la evolución de la Web Semántica, y el desarrollo de la informática móvil, con la combinación de los PDAs (Personal Digital Assistant) y la telefonía móvil, provoca la necesidad de aplicar sistemas de agentes, que extiendan las capacidades operativas de forma distribuida, no sólo de los usuarios, sino también de los sistemas informáticos integrados en la Web. Como se puede observar, el entender la importancia de estas tecnologías será un aspecto estratégico clave para la supervivencia de muchas organizaciones. Un ejemplo de ello lo encontramos en las empresas como Amazon Inc. o Yahoo Inc., supervivientes de la burbuja tecnológica, y que han incorporado tecnologías de Web Semántica en sus sistemas Web, permitiendo obtener ventajas competitivas y diferenciación tecnológica frente a sus competidores, haciendolas a su vez, más eficientes y aumentando el beneficio de sus clientes.

3.3.5. YALE – WEKA

Yale (Yet another learning environment) es un entorno para el desarrollo de experimentos de minería de datos. Desarrollado con el lenguaje multiplataforma JAVA en la Universidad de Dortmund, Alemania, por un equipo de desarrolladores entre los que se encuentran Simon Fischer, Ralf Klinkenberg, Ingo Mierswa y Oliver Ritthoff. Se puede considerar una evolución de otro entorno llamado WEKA (Waikato Environment for Knowledge Análisis).

Son muchas las aplicaciones en las que se puede hacer uso de este entorno, desde investigación hasta aplicaciones industriales o empresariales, entre otras. Yale posee un GUI (Graphical User Interface – Interfaz Gráfico de Usuario), figura 3.10, que facilita la elaboración del experimento de minería de datos.

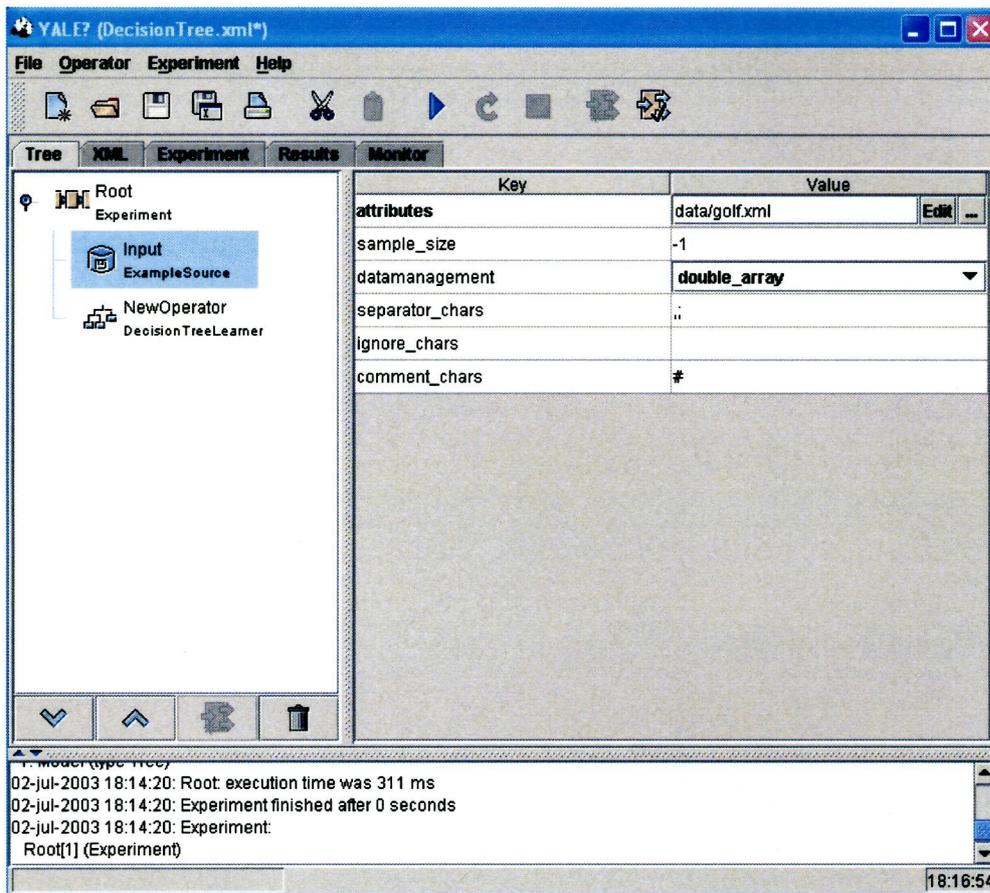


Figura 3.10: Interfaz gráfico de usuario de Yale.

Entre los operadores que son incluidos en Yale se pueden encontrar:

1. *Aprendizaje máquina (machine learning)*: SVM (Support Vector Machines) aplicados a tareas de regresión y clasificación, aprendizaje por árboles de decisión (C4.5 y otros), algoritmos de clustering, y un integrador para todos los operadores clasificadores y de clustering de WEKA. Es decir, WEKA esta incluido como una parte más de Yale.
2. *Selección y generación de atributos*: selección hacia delante, eliminación hacia atrás y uso de varios algoritmos genéticos.
3. *Pre- procesamiento de datos*.
4. *Evaluación de rendimiento*: esquemas de evaluación y validación cruzada, varios criterios de rendimiento para clasificación y regresión, operadores para la optimización de parámetros en operadores cerrados y cadenas de operadores para el registro y presentación de resultados.
5. *Entrada y salida de datos*: varios operadores flexibles para la entrada-salida de datos, soporte para la (re) disposición y uso de metadatos (opcional).

Asimismo, WEKA representa una colección de algoritmos de aprendizaje máquina implementados en JAVA. WEKA incluye un amplio número de esquemas de aprendizaje para la clasificación y regresión (predicción numérica), como inductores de árboles de decisión, aprendizaje de reglas, SVM, aprendizaje basado en ejemplos, naive Bayes, perceptrones multicapa etc, y métodos básicos de evaluación, como validación cruzada y bootstrapping.(Efron y Tibshirani, 1993). WEKA, también posee algunos operadores para pre-procesamiento de atributos, así como tres esquemas básicos de selección de atributos, llamados *enfoque basado en la correlación de atributos* (Hall, 1999), *un enfoque wrapper* (Kohavi y John, 1999), y *el enfoque de filtro* (Kira y Rendell, 1992); e incluye, además, meta clasificadores tales como bagging (Breiman, 1996) y boosting (Schapire, 1999). Sin embargo, WEKA no posee: ni la posibilidad de componer y analizar cadenas de operadores complejos, consistentes en diferentes pasos enlazados de pre-procesamiento, aprendizaje y evaluación; ni generadores de atributos, para la introducción de nuevos atributos, entre otros

aspectos.

Por otro lado, Yale, en contraposición a WEKA o a otras herramientas, ofrece la oportunidad de realizar complejas operaciones de minería de datos, en las que se incluya una secuencia o combinación de varios métodos de aprendizaje máquina y pre-procesamiento de datos. Cada uno de los métodos en Yale, es conocido como *operador*. Una secuencia de operadores, se conoce con el nombre de *cadena de operadores*. Y de nuevo, una *cadena de operadores* es un *operador* (figura 3.11). La posibilidad de enlazar *operadores* o *cadena de operadores* hace posible la creación de procedimientos complejos de experimentación (Fischer *et al.*, 2002).

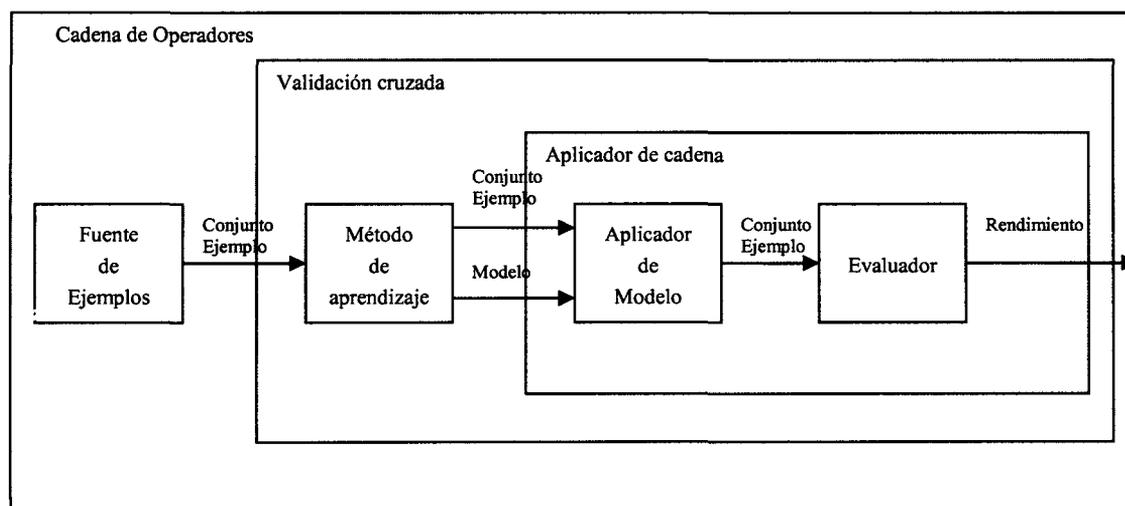


Figura 3.11: Simple ejemplo de una cadena de operadores enlazados en Yale.

(Fuente: Adaptado de Fischer *et al.*, 2002; pag. 13)

Otra característica de gran importancia y que permite la diferenciación con otras herramientas, es el uso de formatos de almacenamientos como es XML. Esto permite la estandarización, la integración con tecnologías Web y la compatibilidad con otras herramientas. Y por último, YALE permite el uso de meta conocimiento especialmente para las tareas de generación de atributos. Esto permite a su vez, la integración de los experimentos de minería de datos con modelos ontológicos, para áreas específicas de conocimiento, como es el caso de esta investigación. Para cada estructura de datos, existe una ontología que describe las jerarquías relacionales entre los diferentes tipos.

En el cuadro 3.8 se presenta la relación de métodos disponibles en Yale, agrupados

por el tipo de algoritmo. Con todo ello se puede observar las posibilidades que ofrece este tipo de herramientas que le hacen una candidata ideal para el Sistema Avanzado de Gestión del Conocimiento aplicado a la Competitividad Internacional.

TIPO DE ALGORITMO	ALGORITMO
Machine Learning	<ul style="list-style-type: none"> • Learner • ModelApplier • SVMLearner • SVMApplier • SVMLightLearner • SVMLightApplier • C45Learner • RuleSetApplier • ID3Learner • DecisionTreeLearner • DecisionTreeApplier • NeuralNetLearner • NeuralNetApplier • MultiClassLearner • MultiModelApplier • WekaLearner • WekaApplier
Clustering	<ul style="list-style-type: none"> • Clusterer • WekaClusterer • ClusterWrapper
Validación y evaluación del rendimiento	<ul style="list-style-type: none"> • PerformanceEvaluator • ValidationChain • FixedSplitValidationChain • XValidation • MethodValidationChain • RandomSplitMethodValidationChain • MethodXValidation • ParameterOptimization
Generación y selección de atributo	<ul style="list-style-type: none"> • FeatureOperator • BruteForce • FeatureSelection • GeneticAlgorithm • GeneratingGeneticAlgorithm • DirectedGeneratingGeneticAlgorithm • YAGGA • FeatureGeneration • MissingValueReplenishment

Cuadro 3.8: Métodos disponibles en el entorno Yale

A continuación, y una vez descritas cada uno de los componentes del sistema se presentan los diferentes tipos de integración que el investigador previamente ha comprobado que pueden aplicarse en la práctica.

3.3.6. Métodos de integración

Son muchas las formas de integración que se pueden realizar con este sistema, como ejemplo de ello, el siguiente cuadro 3.9 se puede observar algunas de las posibles combinaciones y el método de integración que se puede realizar.

COMBINACIÓN	MÉTODO DE INTEGRACIÓN
Protégé-2000 + Jess	<ul style="list-style-type: none"> • JessTab • Línea de comando Jess • Programa JAVA
Protégé-2000 + JADE	<ul style="list-style-type: none"> • BeanGenerator • Programa JAVA
Protege-2000 + Jess + Yale-WEKA	<ul style="list-style-type: none"> • JessTab • Línea de comando Jess • Programa JAVA
Protégé-2000 + Jade + Yale-WEKA	<ul style="list-style-type: none"> • BeanGenerator • Programa JAVA
Protégé-2000 + Jess + JADE + Yale-WEKA	<ul style="list-style-type: none"> • JessTab • Línea de comando Jess • Progrema JAVA + BeanGenerator
Jess + JADE	<ul style="list-style-type: none"> • Línea de comando Jess • Programa JAVA
Jess + Yale	<ul style="list-style-type: none"> • Línea de comando Jess • Programa JAVA
Jess + JADE + Yale-WEKA	<ul style="list-style-type: none"> • Línea de comando Jess • Programa JAVA
JADE + Yale-WEKA	<ul style="list-style-type: none"> • Programa JAVA

Cuadro 3.9: Métodos de acceso disponibles en el entorno del sistema

Como se puede observar en el cuadro 3.9, los métodos de integración se realizan, como mínimo, desde dos componentes del sistema. Por ejemplo, se pueden realizar integraciones directas entre componentes tales como, el gestor ontológico Protégé-2000 y el motor de inferencia Jess. Esta integración se puede realizar vía el plugin de Jess para Protégé, JessTab; vía la línea de comando de Jess, o directamente desde la llamada realizada por un programa informático desarrollado en JAVA. Este tipo de integración se produce de manera parecido en la mayoría de las otras formas de integración Si bien, en algunas situaciones como el caso de la integración entre JADE y Yale-Weka, la única forma de realizar dicha integración en este momento es a través de un programa desarrollado en JAVA.

Una vez presentados los aspectos metodológicos del sistema, a continuación se presenta una simulación del sistema propuesto.

3.4. SIMULACIÓN PRÁCTICA DEL SISTEMA

A continuación se describe una simulación del sistema avanzado de gestión del conocimiento aplicado a la competitividad internacional, como demostración de la capacidad funcional e integradora de dicho sistema para los estudios de la competitividad.

Tomando como base el cuadro 3.6, relacionado con los requerimientos del sistema propuesto, y el cuadro 3.9, referido a los métodos de acceso disponibles del sistema presentados en este capítulo, el siguiente ejemplo pretende mostrar, a través de los componentes del sistema (Protégé-2000, Jess, JADE, Yale-Weka) y el estudio base de la competitividad internacional de Canarias, una muestra de las posibilidades funcionales que puede ofrecer el sistema avanzado de gestión del conocimiento propuesto en esta tesis para el estudio de la competitividad internacional de los territorios. Hay que tener en cuenta que las posibilidades que ofrece el sistema son ilimitadas, ya que incluye potentes herramientas tecnológicas, convertidas aquí en componentes, que por sí solas e integradas entre ellas abre un mundo de posibilidades para gestionar y crear conocimiento sobre la competitividad internacional de los territorios.

Asimismo, el contexto sobre el que parte esta demostración, es el de una situación en la que no existe modelo ontológico alguno y sólo se posee un conjunto de libros en Excel que contienen los datos del informe base de competitividad. Iniciamos entonces el proceso de transferencia de los datos contenidos en los libros Excel, a una base de datos sobre MySQL. La razón por la que se realiza este proceso previo, radica en las enormes posibilidades que ofrece MySQL, como base de datos relacional. A continuación, se presentan los procesos utilizados para la demostración práctica del sistema avanzado de gestión del conocimiento.

3.4.1. - Proceso de conversión de datos a modelo ontológico

En el caso específico del modelo de competitividad objeto de estudio, se importa uno de los conceptos básicos (países), utilizado en el modelo de competitividad y almacenados previamente en la base de datos MySQL (nominada *phd*). Este proceso, se realiza desde el componente Protégé-2000, a través de un *plugin* denominado DataGenie, que permite conectar Protégé-2000 con el servidor MySQL, y que a su vez, permite realizar una selección de la tabla o tablas para ser incorporadas como Clase o Clases, con el objetivo de crear un modelo ontológico parcial, basado en el informe base de competitividad internacional. Es decir, la transferencia desde MySQL a Protégé-2000, se realiza utilizando como base del diseño ontológico, la estructura de la base de datos del informe de competitividad, que ha sido creada en MySQL. Los datos del estudio de competitividad, almacenados en la base de datos, son incorporados a Protégé-2000 junto al modelo ontológico creado, convirtiéndose en casos (instancias), generados a raíz de las tablas de datos. De esta forma, el proceso 1 satisface el requerimiento 3 del sistema relacionado con la integración, ya que permite integrar Protégé-2000 con MySQL.

En la figura 3.12, se puede observar la ventana del Protégé-2000, con el interface del plugin *Dataginie* (véase ①). En esta ventana, se puede apreciar que la fuente de datos utilizada (Data Source) es la llamada *phd* (véase ②), es decir, la base de datos en MySQL. Asimismo, al utilizar una base de datos multiusuario, se podría solicitar el nombre de usuario (User Login) y la clave de acceso de dicho usuario (Password), para realizar la autorización de acceso a la base de datos necesaria en

estos casos (véase 3). Justo debajo de estos campos de entrada de datos, se encuentran dos botones que permiten, después de entrar la información de la base de datos, con la que se desea realizar la conexión, establecer la conexión (Connect), o solicitar información de la base de datos (Info) (véase 4). Por su parte, en la sección denominada *Data Tables*, se relaciona las tablas existentes en la base de datos que contiene los datos del informe de competitividad (véase 5). En este caso, el usuario puede seleccionar la tabla que desea convertir en clase del modelo ontológico. Por otro lado, en el apartado *Import Methods* se permite la selección de un método de importación que puede ser: importando todas las tablas de la base de datos de competitividad (Straight-thru dump – all tables above), importando la tabla o tablas seleccionadas, que en este caso es la tabla países (Straight-thru dump – selected tables), o utilizar un método heurístico para realizar el proceso de importación (Heuristic) (véase 6). Cuando la decisión haya sido tomada, se pulsa el botón (Begin Importing) y se comenzará el proceso (véase 7). Además, en la sección de la derecha, se puede observar el contenido de la tabla seleccionada en este caso, incluyendo por columna el nombre del campo de la tabla, como por ejemplo *PaisnNomEn* (véase 8). El número de registros de la tabla (filas) que se desea visualizar, puede ser establecido en el campo *Number of rows* (Número de Filas). El valor por defecto es 50 filas, aunque colocando el valor 0, se puede obtener el total de registros que posea la tabla.

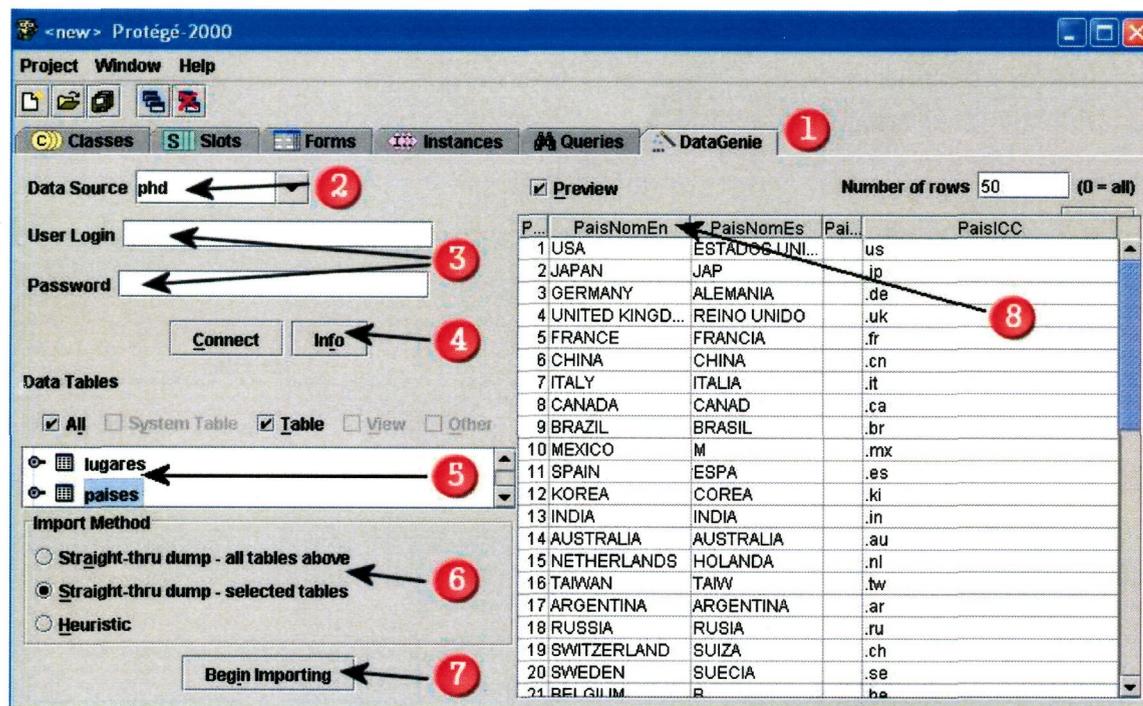


Figura 3.12: Proceso 1 – Conversión de datos a modelo ontológico

3.4.2. - Administración del modelo ontológico

Como consecuencia de la realización del *proceso 1*, relacionado con la conversión de datos al modelo ontológico, se crea, de forma automática, en el modelo ontológico gestionado desde el Protégé-2000, la clase *países* (véase figura 3.13 **1**), con sus referencias (España, Italia, Francia,...etc.). Así, este segundo proceso, *Administración del modelo ontológico*, se inicia con la modificación de la definición del concepto *países*, descrito a través de sus elementos (slots) (véase figura 3.13 **2**), como pueden ser: nombre del país, dirección url (dirección web oficial del país), nombre del país en inglés (puede incluir cualquier otro idioma), localización geográfica, entorno regional al que pertenece (ejemplo Unión Europea). La modificación, por ejemplo, consiste, en este caso, en la eliminación del Slot que hace referencia a la denominación del país en español. De esta forma, el proceso 2 satisface el requerimiento 10 relacionado con la actualización de datos y la evolución del modelo ontológico (véase figuras 3.13 y 3.14, procesos 2a y 2b). En este caso, la modificación la ha realizado directamente el usuario, desde el GUI denominado *Classes*, de Protégé-2000.

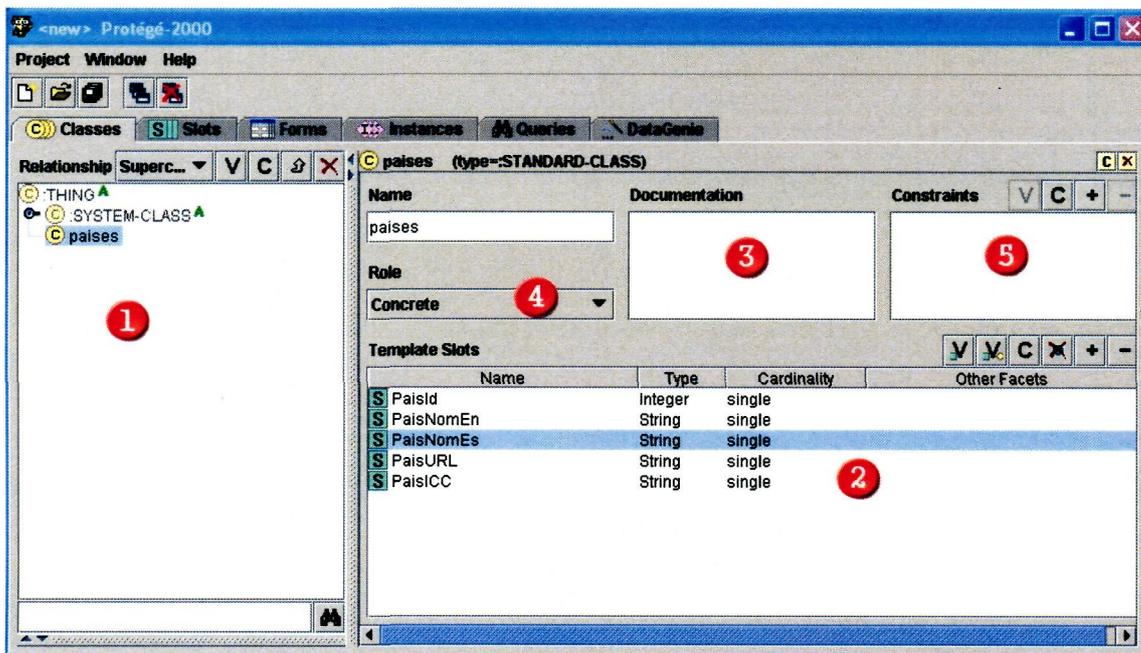


Figura 3.13: Proceso 2a – Modificación del modelo ontológico.

Como se puede observar en la sección de la derecha de la figura 3.13, el nivel de detalle que puede poseer la clase es completo. La sección *Documentation* (véase 3) está dirigida a la descripción del campo o spot que se encuentre seleccionada en cada momento. En el apartado de *Role* (véase 4), se puede definir el papel de la clase, es decir, si es concreta o abstracta, en el caso de que sea concreta, dicha clase puede contener *Instances* de forma directa, pero en caso de que sea abstracta, dicha clase no puede contener de *Instances* de forma directa. La sección *Constraint* (véase 5), permite la definición de limitaciones de la base de conocimiento, que puede ser expresada de forma programática mediante el lenguaje de PAL (*Protege Axiom Language*), es decir, un súper conjunto de lógica de predicado de primer orden, que permite, no sólo definir las limitaciones o restricciones sobre la base de conocimiento, sino también permite definir consultas lógicas, sobre el contenido de la base de conocimiento.

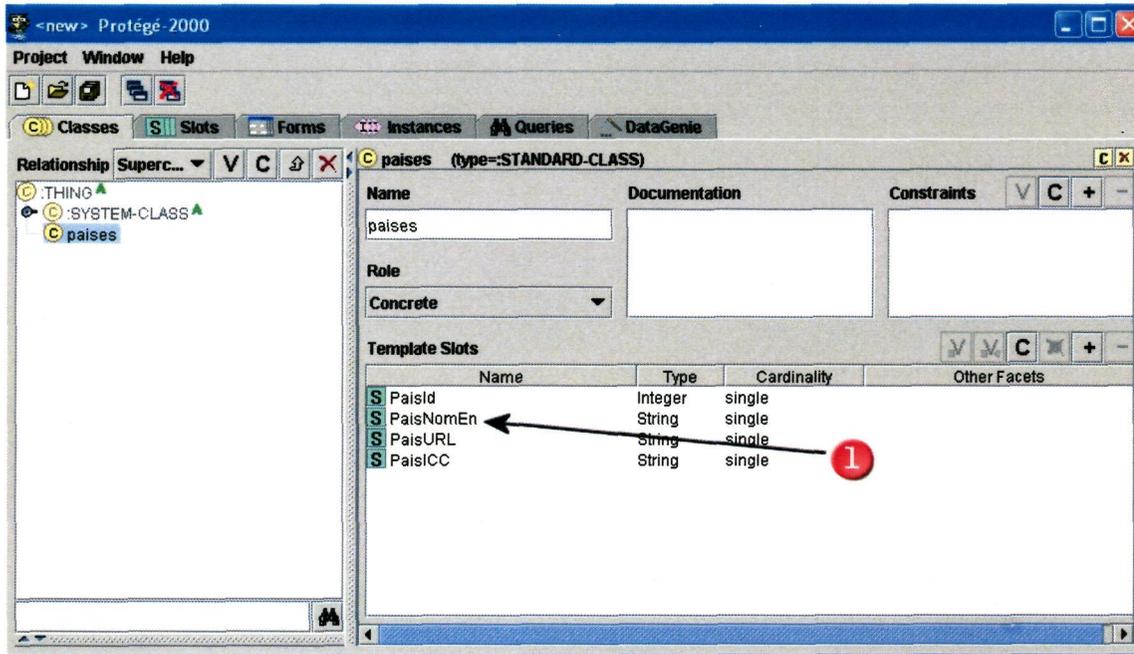


Figura 3.14: Proceso 2b

En esta figura 3.14, se puede observar que el *Slot PaisNomEs* situado debajo del *Slot PaisNomEn* (véase figura 3.13 1), es decir, el nombre del país en español, ha sido eliminado de la clase país y a su vez, de la base de conocimiento. Quedando como *Slots* de la clase país, *PaisId*, *PaisNomEn*, *PaisURL* y *PaisCC*. No obstante, es importante resaltar que aunque se haya producido este cambio en la estructura de la clase, este cambio no afecta en ningún sentido al contenido de la base de datos de competitividad desarrollado en MySQL.

3.4.3. - Integración del motor de inferencia

Con el fin de satisfacer los requerimientos 13 y 14, referenciados en el apartado 3.2.1. de este capítulo, y relacionados con la Modularidad y el carácter abierto del sistema, en este cuarto proceso se realiza la integración del componente Jess (motor de inferencia) con Protégé-2000, componente para la administración del modelo

ontológico, a través del plugin JessTab (véase figura 3.15 **1** y **2**).

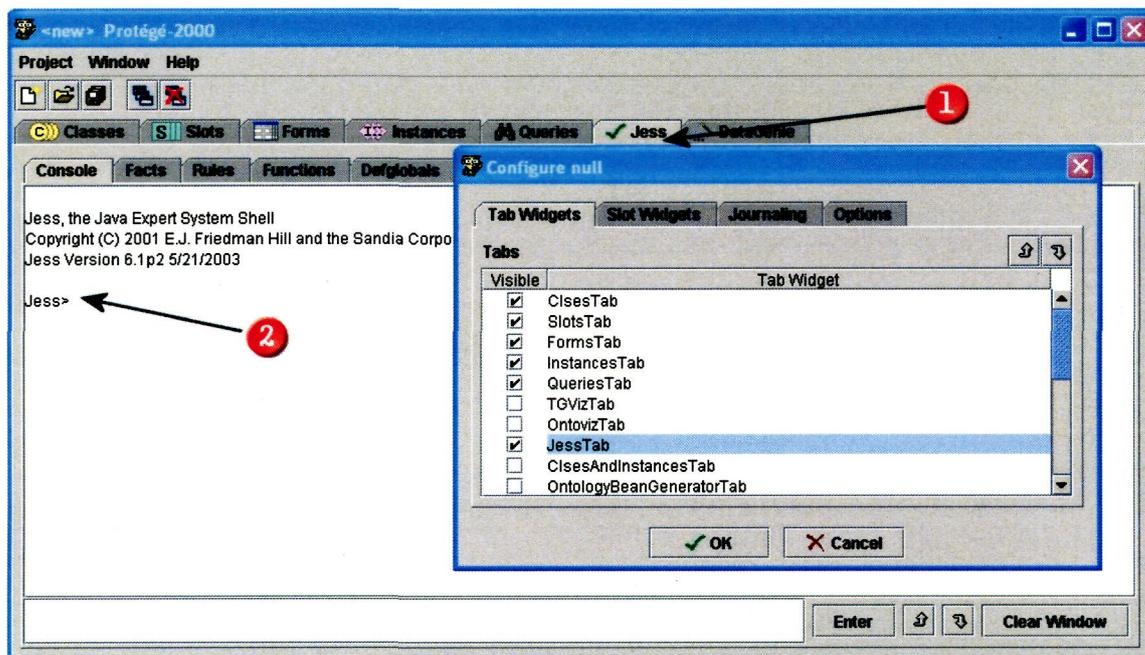


Figura 3.15: Proceso 3 – Integración de Jess a Protégé-2000

Con el proceso 3 se abre un abanico de posibilidades para la evolución del modelo ontológico que se defina y por consiguiente la generación de nuevo conocimiento simbólico (basado en símbolos) sobre la competitividad de los territorios y mediante la inferencia de conocimiento.

3.4.4. - Vinculación de las *facts* al modelo ontológico

En este tercer proceso de simulación, se transfiere la clase países y sus ejemplos (*instances*) desde Protégé-2000 a Jess (véase figura 3.16 **1** , proceso 4a) mediante la orden (*mapclass países reactive*), y enlazados en tiempo real, de forma que cualquier modificación del modelo de conocimiento explicitado en Protégé-2000 será actualizada de forma automática en Jess y viceversa (véase figuras 3.17, 3.18 y 3.19, procesos 4b, 4c y 4d). El proceso cuatro satisface los requerimientos 10 y 14, referenciados en el apartado 3.2.1., de este capítulo, y relacionados con la actualización y el carácter abierto del sistema. Asimismo, este proceso permite integrar

el conocimiento de competitividad y su evolución entre los componentes del sistema.

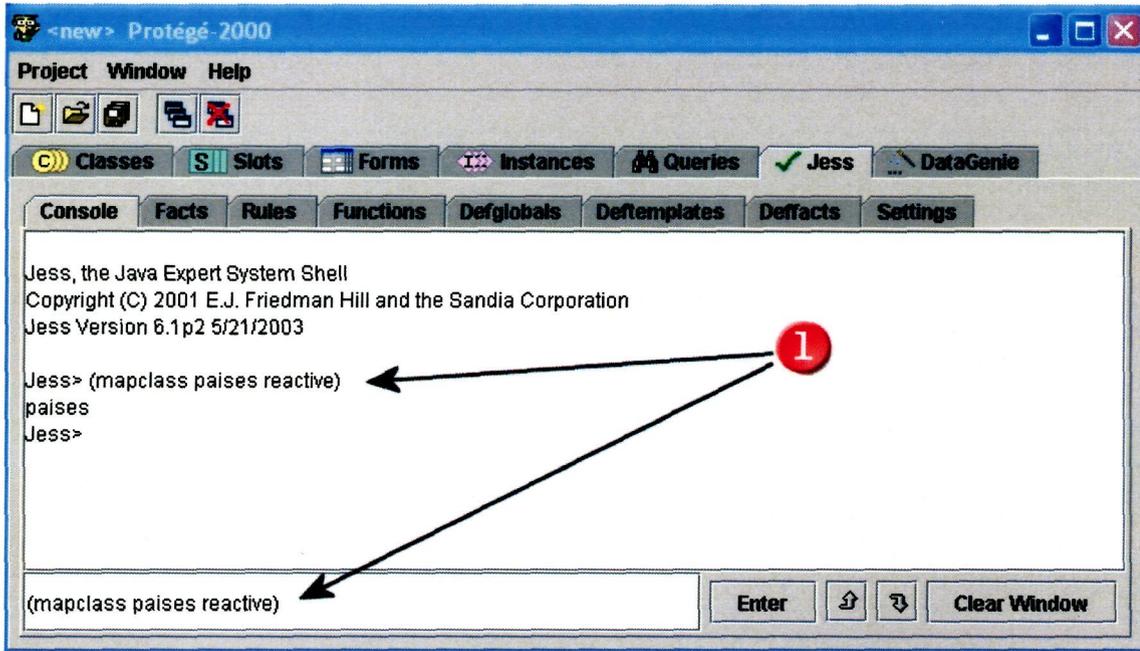


Figura 3.16: Proceso 4a – Portabilidad desde el Protégé-2000 a Jess

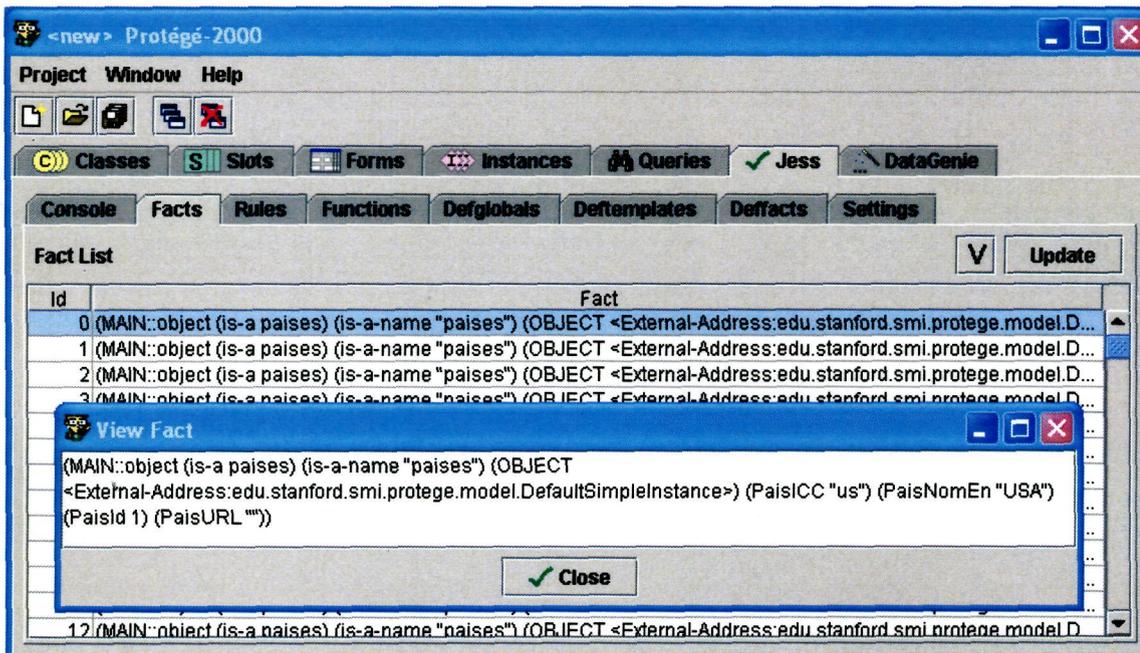


Figura 3.17: Proceso 4b

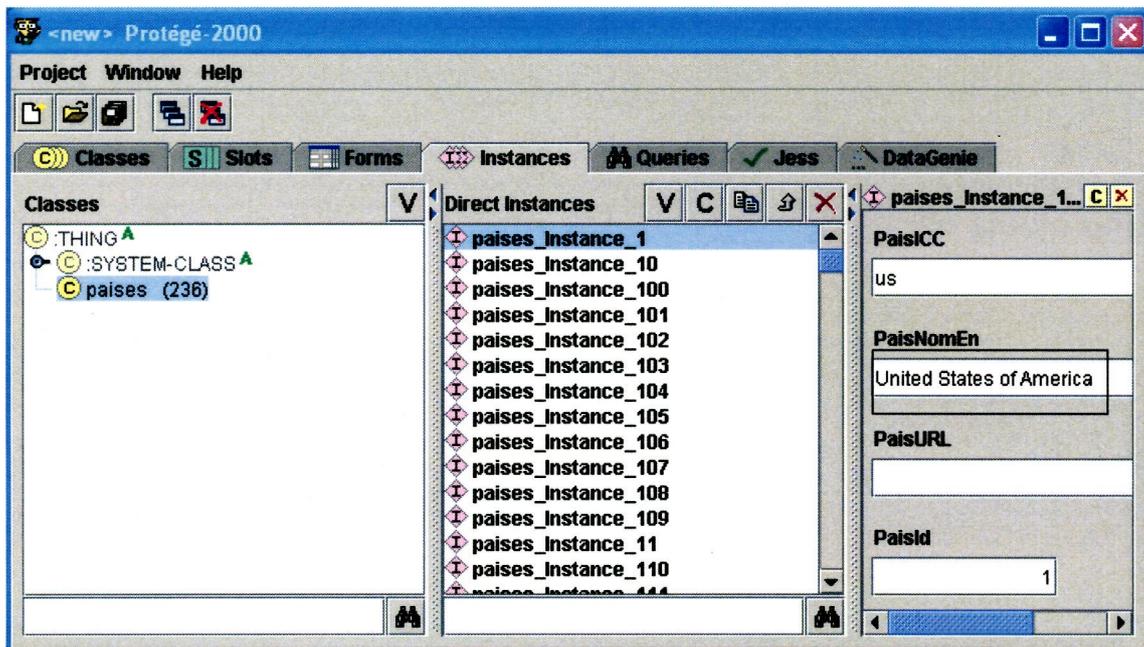


Figura 3.18: Proceso 4c – Modificación del nombre USA a United States of America desde Protégé-2000

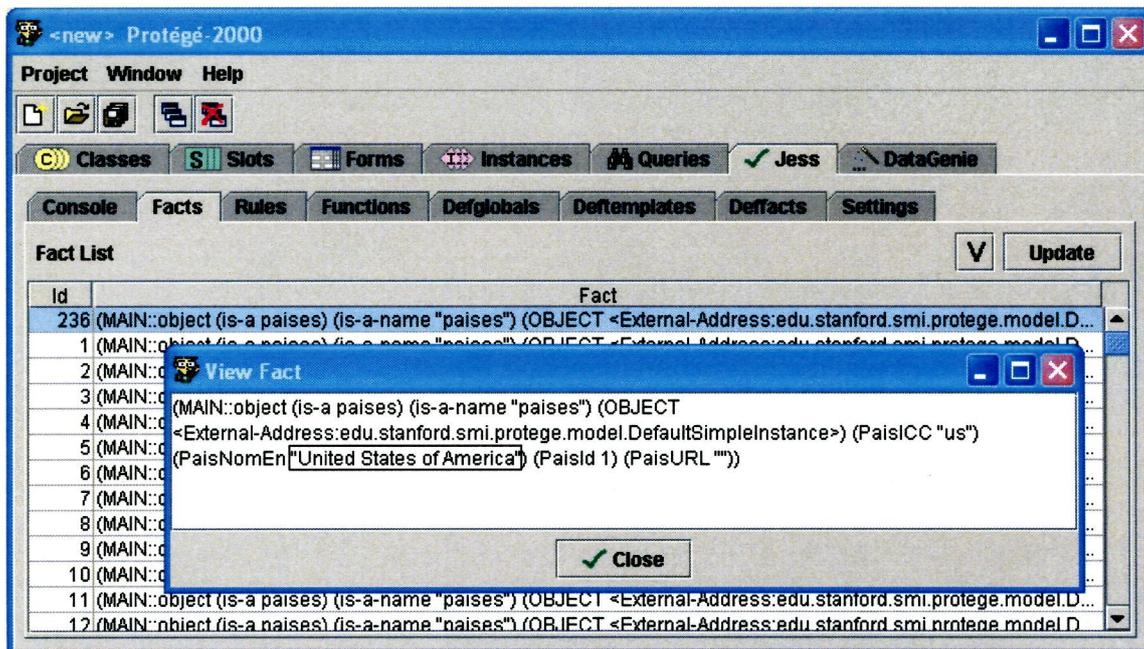


Figura 3.19: Proceso 4d – Confirmación del cambio automático del nombre en Jess

3.4.5. - Producción de código ontológico de agentes

En este quinto proceso de la simulación, se realiza la integración del componente JADE con el modelo ontológico de conocimiento, que se está desarrollando (utilizando la clase *países*), en Protégé-2000. Se lleva a cabo utilizando el plugin para Protégé-2000, llamado Ontology Bean Generador, que genera, de forma automática, el código fuente en JAVA, basado en la ontología definida hasta el momento, para su posterior uso en el componente JADE o en cualquier otro programa específico.

De esta forma, permitimos adaptar los agentes de JADE al conocimiento del modelo ontológico de competitividad, para permitir la realización de tareas automáticas y de transferencia del conocimiento ontológico de competitividad, de forma artificial. Así, se pueden realizar de forma artificial tareas complejas como, por ejemplo, la comunicación o actualización de los datos de competitividad en tiempo real.

Como consecuencia de este proceso, se satisface los requerimientos 10, 15 y 16, referenciados en el apartado 3.2.1 de este capítulo, y correspondientes a la actualización del modelo ontológico de competitividad, el carácter dinámico del conocimiento y la autonomía de las tareas (véase figuras 3.20 y 3.21, procesos 5a y 5b).

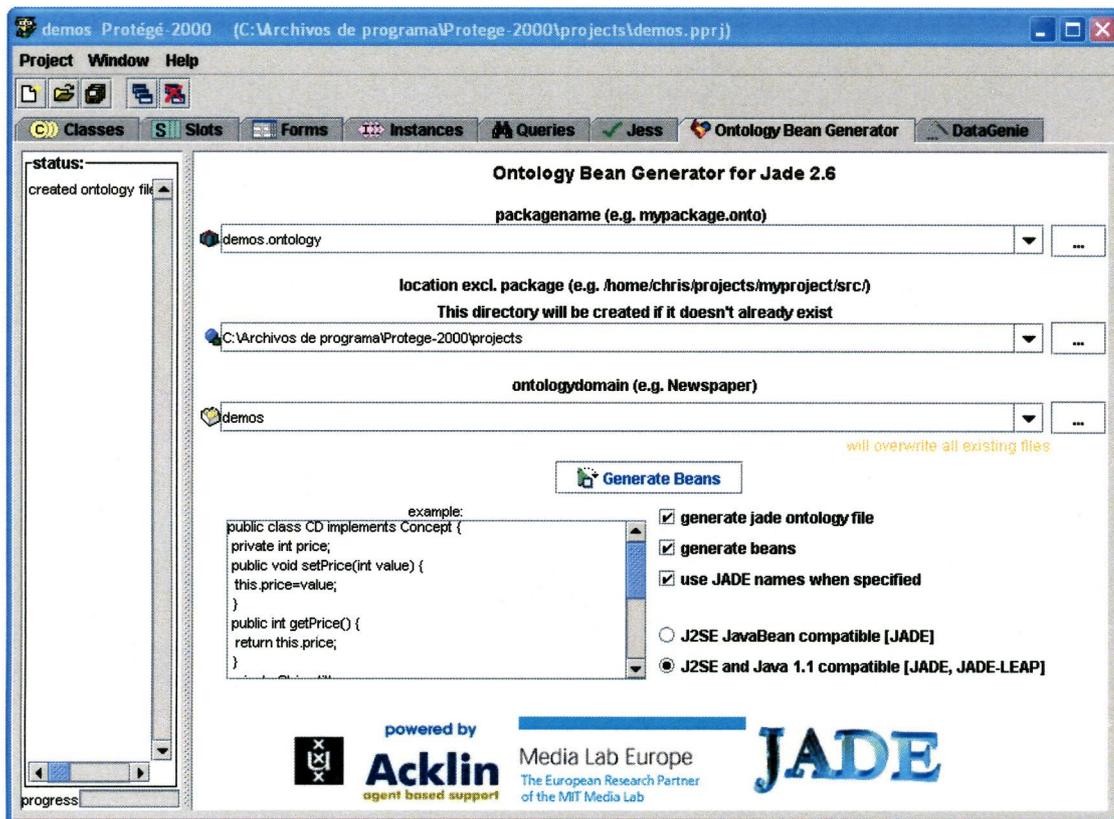


Figura 3.20: Proceso 5a – Generación código JAVA para JADE desde Protégé-2000

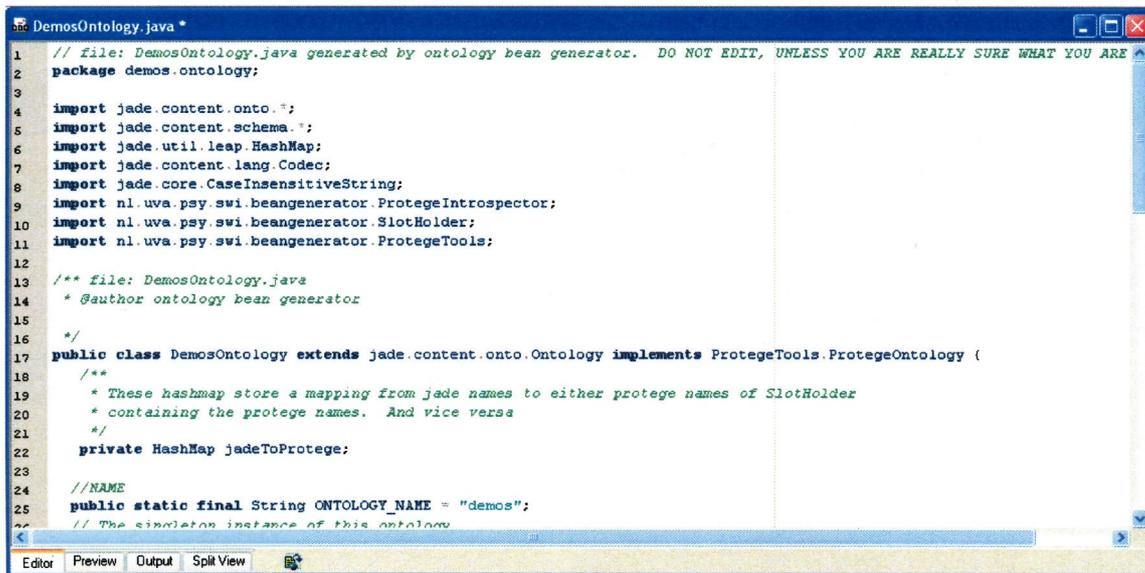


Figura 3.21: Proceso 5b – Código JAVA generado desde Protégé-2000

3.4.6. - Descubrimiento de un árbol de decisión

Partiendo de la tabla de datos generada a partir del índice de competitividad mundial, el ranking de los factores de competitividad, y el PIB-PPC (Producto Interior Bruto a paridad de poder de compra) del informe base, el proceso sexto se inicia realizando un experimento con el componente de *knowledge discovery* (Yale-Weka). Este experimento, resultará en la generación automática de un árbol de decisión sobre la competitividad internacional de los territorios.

El cuadro 3.10 presenta la tabla de datos para generar el árbol de decisión en el que se incluyen, el posicionamiento de competitividad internacional de cada país objeto de estudio, el nombre del país, los rankings de actividad económica, eficiencia del gobierno, eficiencia empresarial, e infraestructuras; así como el PIB-PPC de cada país. En este punto, y con el fin de simplificar la clasificación, se dividieron en tres niveles el posicionamiento competitivo: muy competitivo, los países comprendidos entre la posición 1 y 16 del ranking (desde EE.UU. hasta Israel), competitivos los encontrados entre la posición 17 y 32 del ranking (desde Bélgica hasta Portugal), y poco competitivos, los encontrados entre la posición 33 y 50 del ranking (desde Italia a Indonesia).

Cuadro 3.10: Tabla de datos para generar el árbol de decisión (Fuente: Ramos, 2002)

Indice Competitividad	País	Actuación Económica	Eficiencia del Gobierno	Eficiencia Empresarial	Infraestructura	PIB (PPC) Billones \$
1	EE.UU.	1	12	1	1	9332
2	Finlandia	31	2	2	3	126
3	Singapur	2	1	8	9	90
4	Holanda	8	8	3	10	400
5	Suecia	17	13	4	2	209
6	Luxemburgo	3	5	7	15	20
7	Suiza	15	6	11	4	200
8	Irlanda	6	3	10	19	107
9	Hong Kong	4	7	6	23	164
10	Australia	33	4	14	8	486
11	Canadá	16	10	12	11	839
12	Alemania	5	16	15	5	2007
13	Islandia	41	14	5	6	8
14	Dinamarca	27	11	9	12	142
15	Austria	29	9	13	13	210
16	Israel	19	15	16	18	119
17	Bélgica	9	23	19	17	270
18	Reino Unido	7	22	20	22	1354
19	Noruega	18	25	22	14	131
20	Nueva Zelanda	37	18	21	21	75
21	Taiwán	26	24	18	20	494
22	Chile	34	17	17	31	137
23	Francia	10	29	26	16	1387
24	España	24	20	24	24	741
25	Japón	12	34	32	7	3205
26	Estonia	28	19	25	25	13
27	Hungría	20	21	27	28	122
28	Malasia	13	31	29	32	202
29	Grecia	43	26	28	30	169
30	Brasil	25	39	23	37	1232
31	Corea	23	36	36	26	801
32	Portugal	40	27	33	34	166
33	Italia	22	40	30	29	1315
34	China	11	41	42	35	4898
35	República Checa	32	30	45	27	137
36	Rep. Eslovaca	50	32	37	33	58
37	México	35	28	40	43	857
38	Filipinas	39	38	34	48	294
39	Sudáfrica	49	35	35	41	387
40	Tailandia	21	37	44	46	386
41	Turquía	45	45	31	39	440
42	Argentina	46	41	42	38	447
43	Eslovenia	36	46	39	36	33
44	India	30	44	38	44	2377
45	Islas Canarias	38	33	50	47	30,27

Índice Competitividad	País	Actuación Económica	Eficiencia del Gobierno	Eficiencia Empresarial	Infraestructura	PIB (PPC) Billones \$
46	Colombia	47	43	43	49	246
47	Rusia	14	50	49	42	1177
48	Venezuela	44	48	48	40	134
49	Polonia	42	47	46	50	343
50	Indonesia	48	49	47	45	620

Cuadro 3.10 (cont.):Tabla de datos para generar el árbol de decisión
(Fuente:Ramos, 2002)

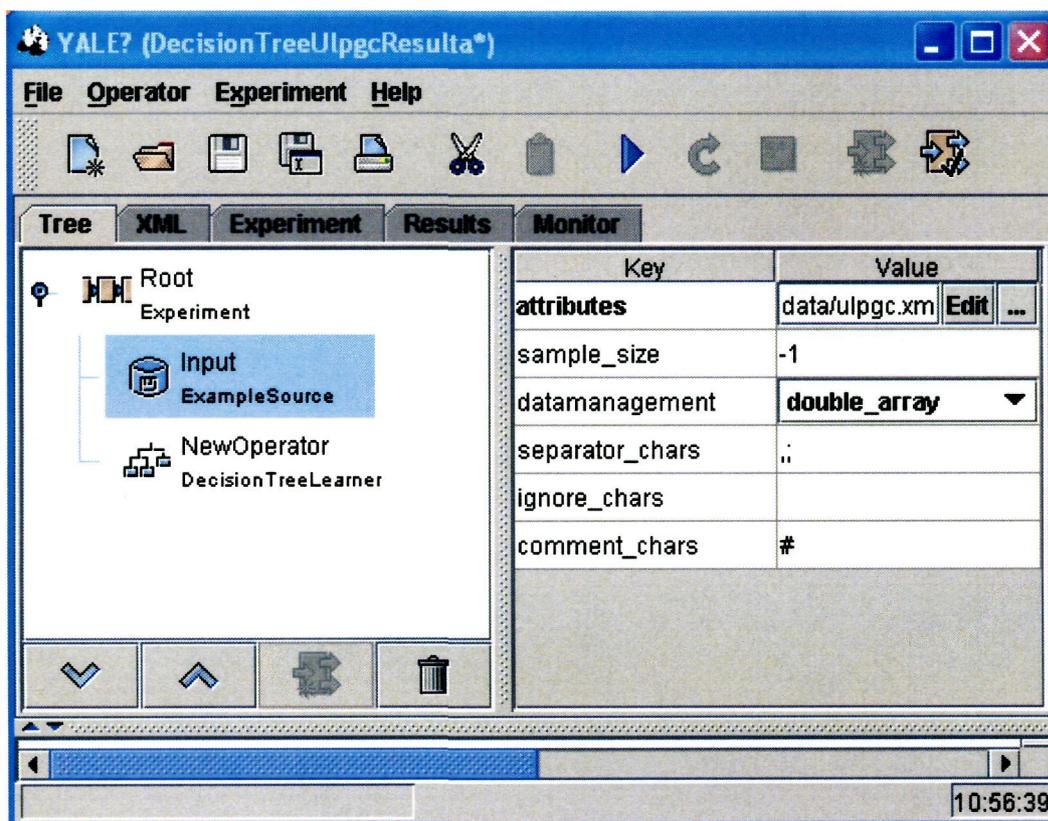


Figura 3.22: Proceso 6a – Árbol de la cadena de operador del experimento

La figura 3.22 (proceso 6a) muestra la cadena de operadores del experimento en Yale-Weka, donde se pueden apreciar los parámetros utilizados en el experimento. Asimismo, la figura 3.23 (proceso 6b), presenta la descripción del experimento en formato XML. Esto permitiría la transferencia y manipulación de los experimentos a través de la Web. Como se puede observar el método de clasificación utilizado, en este caso es un simple generador de árbol de decisión soportado en este componente denominado *DecisionTreeLearner*, un método que extiende las capacidades del método ID3, y soporta este tipo de experimentos. La figura 3.24 (proceso 6c) por su parte, ilustra la representación gráfica de la cadena de operador de este experimento en Yale-Weka.

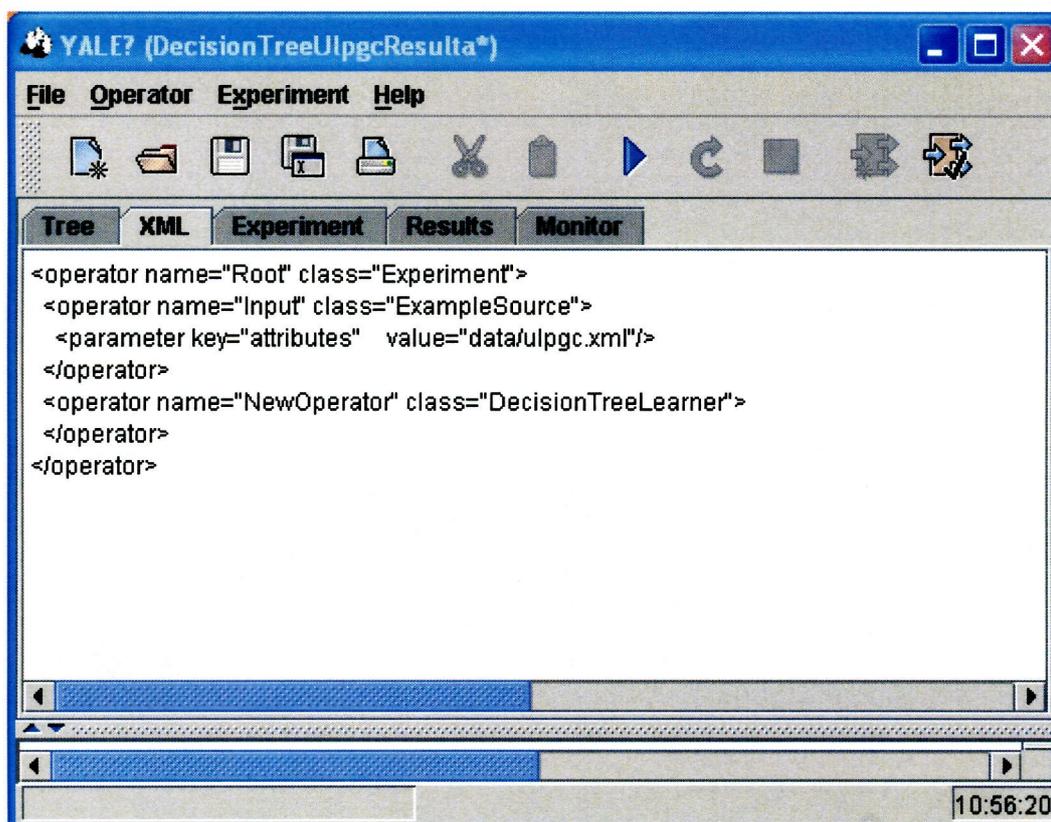


Figura 3.23: Proceso 6b – Fichero XML descriptor del experimento

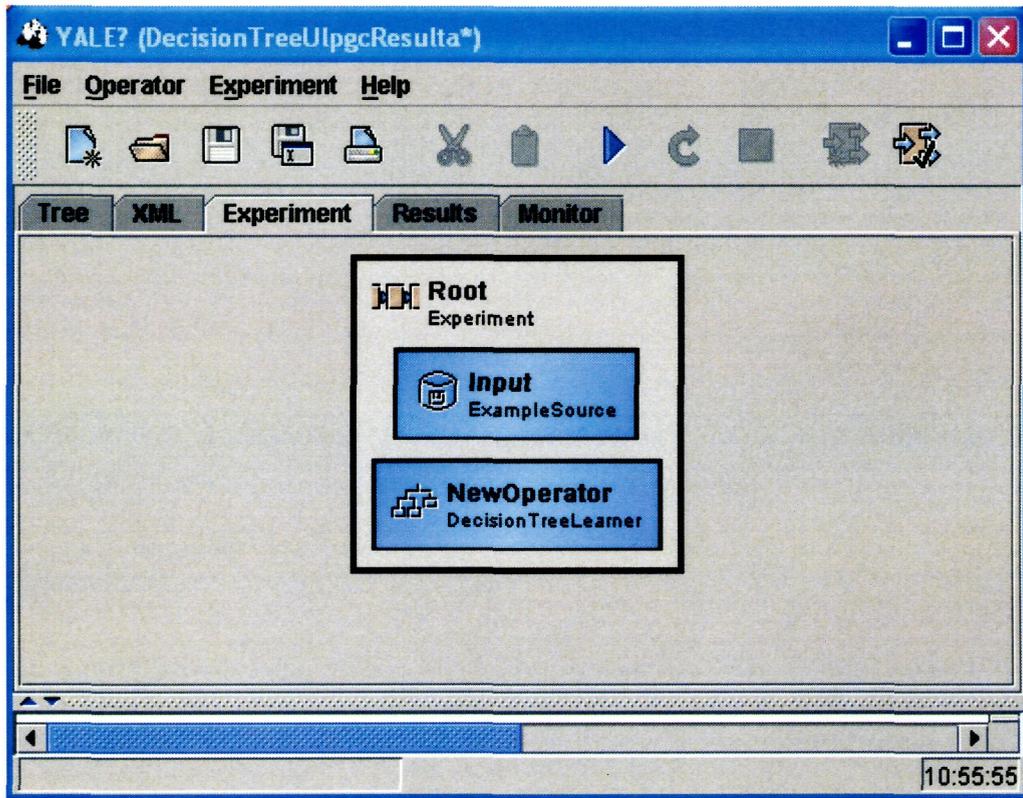


Figura 3.24: Proceso 6c – Representación gráfica del experimento

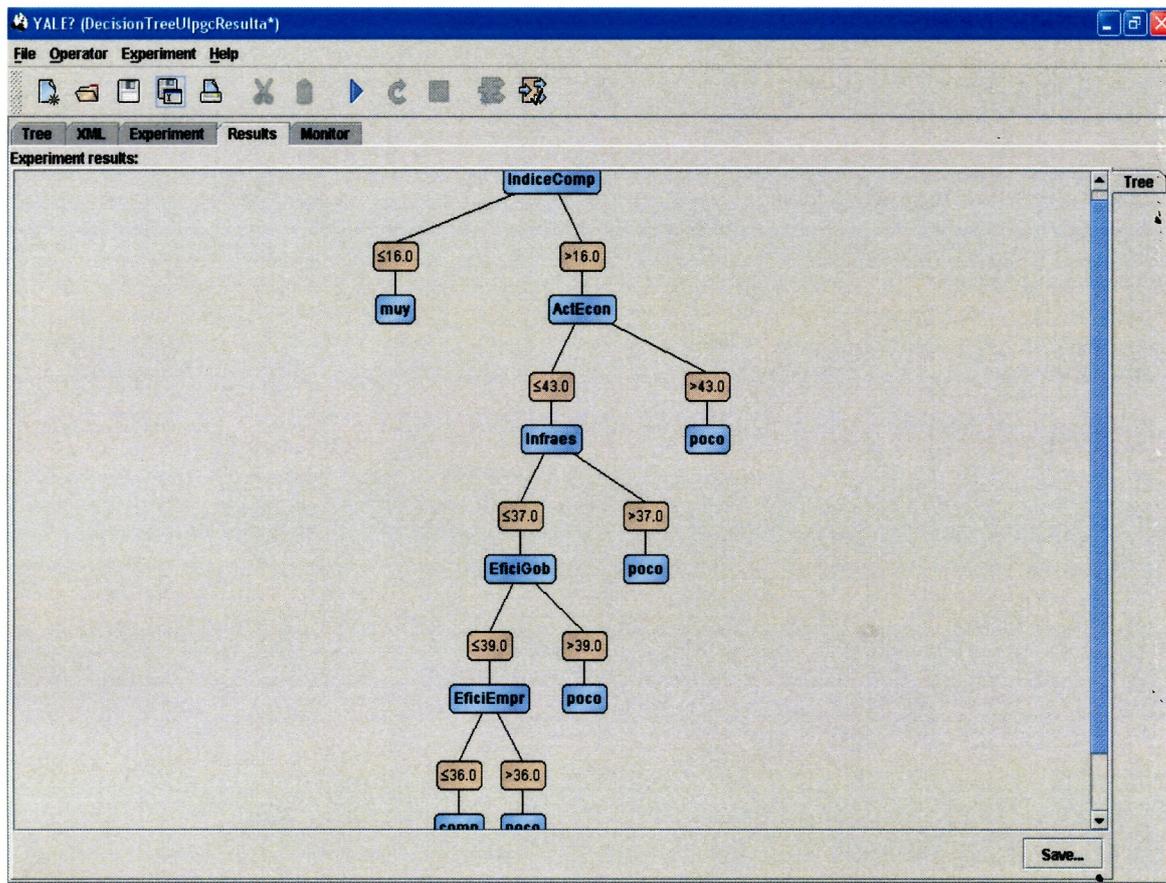


Figura 3.25: Proceso 6d – Árbol de decisión resultante del experimento.

En la figura 3.25 (proceso 6d), se presenta el árbol de decisión resultante, generado en el experimento, de forma automática, con una duración de menos de 2 segundos, utilizando el cuadro 3.10 como fuente de datos. En el árbol, se puede observar que el primer nodo está representado por el índice de competitividad (columna 1 del cuadro 3.10). En el segundo nivel del árbol, se clasifica el valor del índice de competitividad internacional en dos grupos, los índices de posición menor o igual a 16, indican que el país es muy competitivo y la decisión final es tomada (país muy competitivo). Este hecho, coincide con la clasificación realizada anteriormente, es decir, no se aporta nuevo conocimiento. Sin embargo, en el segundo grupo (países con posicionamiento mayor que 16) el árbol continúa la toma de decisiones y analiza el posicionamiento del país en su Actividad Económica (columna 3 del cuadro 3.10). Aquellos países, cuyo posicionamiento dentro de este ranking es mayor de 43, son poco competitivos independientemente del posicionamiento que tengan en otros factores y la decisión

final es tomada (país poco competitivo). En este caso, el árbol ha tomado la decisión de analizar, en este punto, primero el factor de Actividad Económica y clasifica los países poco competitivos, bajo el criterio establecido de forma automática, es decir no aportada por el usuario. De esta forma, el sistema aporta nuevo conocimiento en la toma de decisión.

En el caso de que el posicionamiento del país en el ranking de Actividad Económica sea menor o igual a 43, provoca que se analice el posicionamiento del país en el ranking de Infraestructuras (columna 6 del cuadro 3.10). Si su posicionamiento en este factor es mayor que 37, la decisión tomada es que el país es poco competitivo. En cambio, si su posicionamiento es menor o igual a 37, continúa el árbol de decisión y se analiza el posicionamiento del país en el ranking del factor de Eficiencia del Gobierno (columna 4 del cuadro 3.10). De esta forma, el sistema aporta nuevo conocimiento en la toma de decisión.

Continuando la secuencia anterior, cuando el posicionamiento en el ranking del factor de Eficiencia del Gobierno, es mayor que 39 entonces la decisión tomada es que el país es poco competitivo. Sin embargo, si es menor o igual a 39, entonces se clasifica la competitividad en función al ranking del factor de Eficiencia Empresarial (columna 5 del cuadro 3.10). En este caso, si la posición del país en el ranking internacional en el factor de Eficiencia Empresarial es mayor que 36, entonces la decisión tomada es que el país es poco competitivo, y sino el país es competitivo. Un aspecto a considerar, es que en ningún momento del proceso de toma de decisión en árbol resultante se hace uso de la columna 7 del cuadro 3.10 que contiene los valores correspondientes al PIB-PPC de los países. Además, la clasificación resultante de este árbol de decisión puede ser fácilmente contrastada con el cuadro 3.10.

3.4.7. - Conversión de árbol de decisión a regla de decisión

En este último proceso de la simulación, y utilizando el sistema avanzado de gestión del conocimiento aplicado a la competitividad internacional de los territorios, se toma el árbol de decisión resultante del proceso 6, y se transcribe en reglas de decisión a

través del JessTab y el plugin de Jess para Protégé-2000 (véase cuadro 3.11 del proceso 7). Esto permite realizar inferencias sobre el conocimiento almacenado en el modelo ontológico, definido en Protégé-2000 (véase figura 3.26, proceso 7). De esta forma, estamos integrando el componente Yale-Weka a los componentes Jess y Protégé-2000 del sistema.

Este tipo de integración, puede facilitar la evolución del conocimiento representado en el modelo ontológico de la competitividad internacional de los territorios, desarrollado con el componente Protégé-2000. Este método de integración puede realizar modificaciones de forma automática, sobre el modelo ontológico de la competitividad

```

defrule muycompetitivo
  (pais ?Pais ?PosCompGlob&:(<= ?PosCompGlob 16) ?PosActEcon ?PosEfGob ?PosEmpr ?PosInf ?PIBPPC)
  =>
  (printout t ?Pais " es un pais muy competitivo" crlf)

(defrule pococompetitivo
  (or (pais ?Pais ?PosCompGlob ?PosActEcon&:(> ?PosActEcon 43) ?PosEfGob ?PosEmpr ?PosInf ?PIBPPC)
      (pais ?Pais ?PosCompGlob ?PosActEcon ?PosEfGob ?PosEmpr ?PosInf&:(> ?PosInf 37) ?PIBPPC)
      (pais ?Pais ?PosCompGlob ?PosActEcon ?PosEfGob&:(> ?PosEfGob 39) ?PosEmpr ?PosInf ?PIBPPC)
      (pais ?Pais ?PosCompGlob ?PosActEcon ?PosEfGob ?PosEmpr&:(> ?PosEmpr 36) ?PosInf ?PIBPPC))
  =>
  (printout t ?Pais " es un pais poco competitivo" crlf)

(defrule competitivo
  (pais ?Pais ?PosCompGlob&:(> ?PosCompGlob 16) ?PosActEcon&:(<= ?PosActEcon 43) ?PosEfGob&:(<=
  ?PosEfGob 39) ?PosEmpr&:(<= ?PosEmpr 36) ?PosInf&:(<= ?PosInf 37) ?PIBPPC)
  =>
  (printout t ?Pais " es un pais competitivo" crlf)

```

Cuadro 3.11: Proceso 7 – Reglas de decisión transcritas a JessTab

Con este último proceso, se demuestra la forma de generación de nuevo conocimiento a través del componente Yale-Weka y la integración del mismo en el modelo ontológico de competitividad que se desarrolle en Protégé-2000. Asimismo, este último paso del funcionamiento del sistema nos permite satisfacer los requerimientos 14, 15 y 16, referenciados en el apartado 3.2.1 de este capítulo, y relacionados con el carácter abierto, conocimiento dinámico y la autonomía, y además cumple el requerimiento número 4 de validación ya que nos permite validar en este caso, el árbol de decisión con los datos del cuadro 3.10.

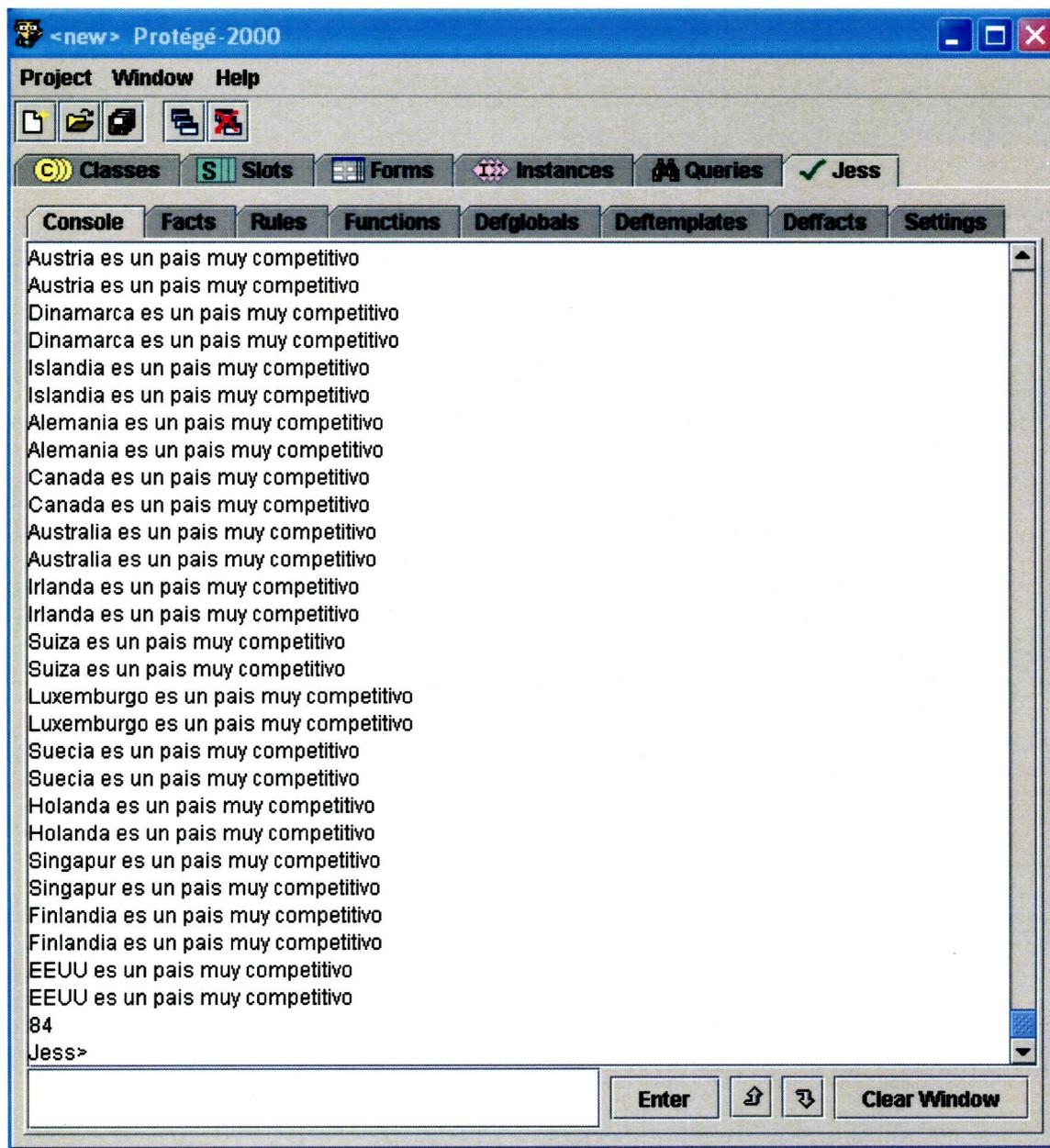


Figura 3.26: Proceso 7– Resultados de la inferencia en JessTab

3.5. ACCESOS Y USUARIO DEL SISTEMA

El sistema avanzado de gestión del conocimiento propuesto en esta tesis, pretende servir de plataforma para crear y gestionar de forma eficiente el conocimiento sobre la competitividad internacional de los territorios, y extender las posibilidades existentes en los estudios de competitividad actuales. Teniendo en cuenta la creciente importancia que está teniendo los trabajos de investigación sobre la competitividad de los territorios, y su utilidad para la toma de decisiones en mercados globales, este sistema, abre un abanico de posibilidades a los diferentes usuarios potenciales, independientemente del nivel de conocimiento técnico que posea. Asimismo, el sistema permite satisfacer diferentes tipos de necesidades cognitivas, en relación a la aplicación que se le quiera dar.

3.5.1. - Modalidades de acceso

Como se ha mencionado anteriormente, en el apartado métodos de integración, son muchas las formas que pueden ser utilizadas para acceder a la base de conocimiento descrita a través del modelo ontológico, así como al propio modelo ontológico. Si bien, estas modalidades de acceso al sistema pueden ser clasificadas en:

- Interfaces gráficos (GUIs) de los componentes, es decir, en función de la operación que se desea realizar existe la posibilidad de acceder a la información y modelo ontológico desarrollado a través del GUI correspondiente del componente.
- A través de algún *plugin*, como puede ser JessTab, relacionado con alguno de los componentes tales como Protégé-2000.
- A través de algún programa informático desarrollado en Java.
- A través de la Web mediante alguna aplicación de gestión de ontologías. Este

tipo de aplicaciones puede ser desarrolladas con herramientas software, tales como KAON (<http://www.kaon.org>), una infraestructura de código abierto, dirigida a la gestión y creación de ontologías, así como aplicaciones ontológicas para la Web.

El nivel de desarrollo que permite este sistema, favorece la creación de nuevas formas de acceso, que dependerá del objetivo que se intente cubrir.

3.5.2. - Tipología de usuarios del sistema

El acceso a este sistema difiere, lógicamente, en función del nivel de conocimiento que tenga el usuario y en dos aspectos clave: el conocimiento sobre estudios de competitividad internacional y el nivel de conocimiento del sistema y de los componentes del sistema. Básicamente, se podrían clasificar en tres niveles: nivel usuario final, nivel del editor de conocimiento, y, por último, nivel avanzado o de desarrollo. A continuación se describe cada uno de estos tipos de usuarios.

El primer nivel, usuario final es básico, es decir, el nivel de usuario final no necesita tener conocimiento ni sobre el sistema ni del conocimiento que contiene. Es por ello, que los interfaces que debe utilizar este tipo de usuario para acceder al sistema, deben ser fáciles y directos al conocimiento y los servicios que se le ofrezca.

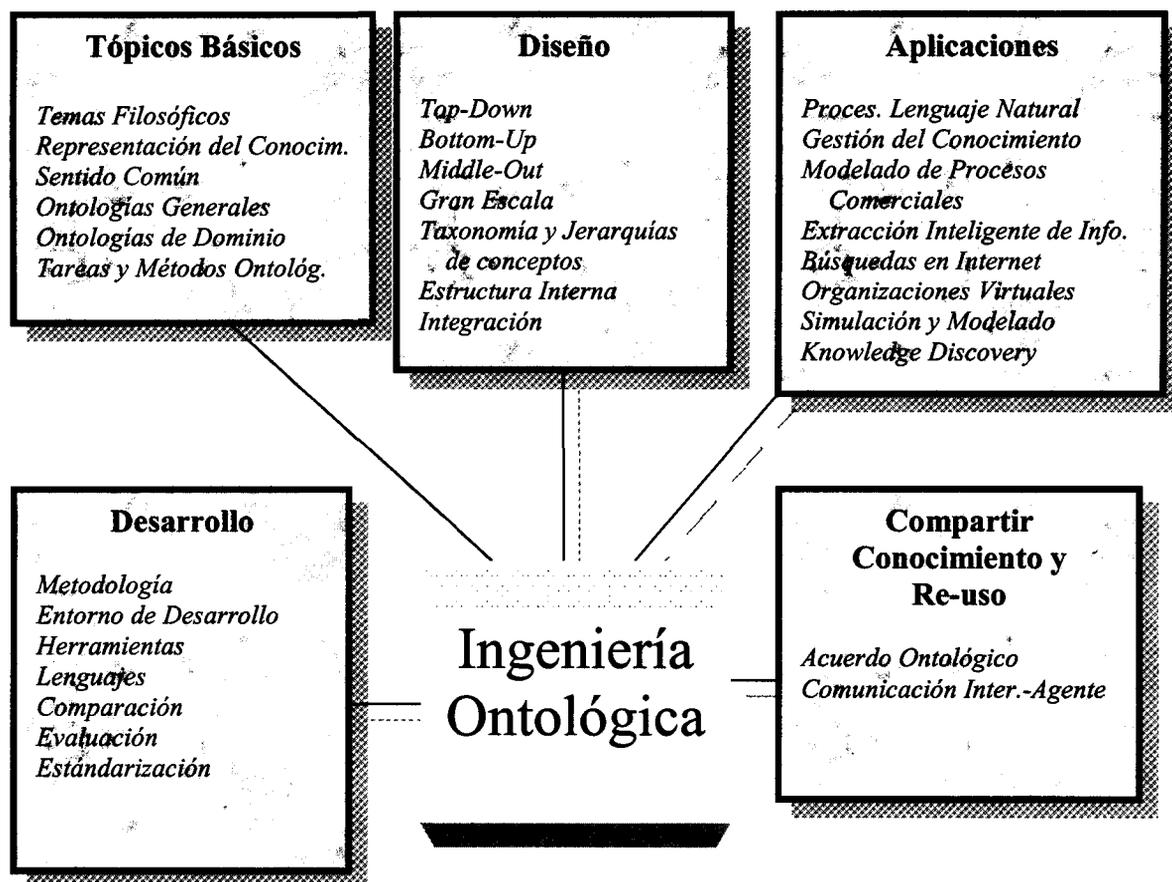


Figura 3.27: Áreas de influencia de la Ingeniería Ontológica
(Fuente: Adaptado de Devedzic, 2002)

Para el segundo nivel, editor de conocimiento, se ofrece la posibilidad de editar modelos ontológicos sobre competitividad, requiriendo dos aspectos clave: 1. Conocimiento sobre los estudios de competitividad; 2. Conocimiento sobre la funcionalidad de los componentes del sistema y otros. Un ejemplo de estas áreas de conocimiento que pueden ser necesarias en el trabajo ontológico, desde el punto de vista de la edición, se puede encontrar en la figura 3.27. Como se observa en esta figura, el conocimiento que debe poseer el editor incluyen: los Conceptos Básicos, Diseño, Aplicaciones, Desarrollo, Compartir y Re-usar Conocimiento. Específicamente, dentro de Tópicos Básicos, se cubre áreas de conocimiento tales como: temas filosóficos, representación del conocimiento, conocimiento del sentido común, ontologías genéricas, ontologías basadas en un área de conocimiento u ontologías específicas y métodos ontológicos. En el área de Diseño se incluyen: diseño *top-down*,

bottom-up, *middle-out*, diseño a gran escala, taxonomías y jerarquías de conceptos, estructuras internas e integración. Asimismo, dentro del área de Aplicaciones se integran las siguientes posibles aplicaciones: procesamiento de lenguaje natural, gestión del conocimiento, modelado de procesos comerciales o de negocio, extracción inteligente de información, búsquedas en Internet, organizaciones virtuales, simulación y modelado y *knowledge discovery*. En el área de Desarrollo se incluyen: metodologías, entorno de desarrollos, herramientas, lenguajes, comparación, evaluación y estandarización. Y, finalmente, en el área donde se comparte de Conocimiento y Re-uso se encuentran los acuerdos ontológicos y las comunicaciones Inter-agentes.

Finalmente, el nivel de desarrollo o avanzado, podría tener acceso total, no sólo desde el punto de vista del uso, sino de desarrollo y extensión del sistema. Para ello se requieren: 1. Conocimiento avanzado sobre competitividad; 2. Técnicos sobre las diferentes tecnologías que sirven o puedan servir de base o de extensión de este sistema.

3.5.3. - Enfoques aplicados al diseño ontológico

El diseño del modelo ontológico sobre competitividad internacional de los territorios que se desee desarrollar, debe considerarse como una pieza fundamental para el buen uso y aplicación de este sistema. La complejidad de esta tarea, viene en función de nivel de diseño que se pretende conseguir así como su aplicación, y la audiencia a la que va dirigida. Es decir, si por ejemplo la audiencia que va a utilizar una ontología es amplia, entonces se requiere de un amplio acuerdo en la definición ontológico. En este caso, y con el fin de simplificar el objetivo de cualquier diseño ontológico de la competitividad internacional, se utilizará la perspectiva utilizada por Gruber (1995), para la definición de ontología como la especificación explícita de una vista simplificada y abstracta de la parte del mundo que se desea representar. De esta forma, se pueden definir los términos y definiciones para los conceptos y sus relaciones, presentes en el conocimiento sobre la competitividad internacional de los territorios.

En la creación y aplicación de ontologías, es aconsejable distinguir entre la ontología en sí, la cual posee la descripción de los conceptos que existen y cuyas relaciones son ciertas, y por el otro lado, las *facts* (hechos) acerca de estos conceptos y relaciones. Es decir, los diferentes valores que pueden tomar estos conceptos o relaciones. Estas *facts*, no son la ontología, aunque estén estructuradas por esta. Estas *facts* están a su vez, sujetas al contexto, la observación, la evaluación o la modificación. Por ejemplo, el concepto *país* es parte del modelo ontológico, pero *España pertenece al continente europeo*, que es un ejemplo de conocimiento, no es parte del modelo ontológico de la competitividad internacional. Por otro lado, el acuerdo o compromiso ontológico es fundamental para la efectividad del modelo ontológico que se plantee sobre la competitividad internacional. Este acuerdo de establecerse no sólo entre los expertos y usuarios de esta área de conocimiento (la competitividad internacional), sino también entre este tipo grupo y el o los editores de conocimiento. La falta de acuerdo ontológico entre las partes puede provocar la reducción de beneficio que puede obtenerse de la manipulación ontológica y del conocimiento que representa o modela. No obstante, existen algunos trabajos de investigación que proponen procedimiento para gestionar las posibles ambigüedades u excepciones en los modelos ontológicos, por ejemplo el trabajo propuesto por Tamma y Bench-Capon (2001). Pero, como ya planteó Gruber (1995), un criterio importante para evaluar la calidad del diseño ontológico consiste en tender hacia el nivel mínimo necesario de acuerdo ontológico, es decir, que el modelo ontológico que se proponga requiera el mínimo nivel de acuerdo. Por lo tanto, a un mayor nivel de acuerdo, una mayor probabilidad de reducción de la calidad del modelo ontológico.

Para realizar el diseño ontológico, se pueden encontrar cinco posibles enfoques básicos: inspiración, inducción, deducción, síntesis y colaboración. Esto no quiere decir que puedan existir enfoques híbridos, en el que se combinen varios enfoques básicos (Holsapple y Joshi, 2002).

ENFOQUES EN EL DISEÑO ONTOLÓGICO	
Enfoque	Descripción básica
Inspiración	Puntos de vista individualizado del área de conocimiento
Inducción	Caso específico dentro del área de conocimiento
Deducción	Principios generales del área de conocimiento
Síntesis	Conjunto de ontologías existentes, cada una de las cuales aporta una caracterización parcial del área de conocimiento.
Colaboración	Punto de vista acerca del área de conocimiento por múltiples individuos, posiblemente unido a una ontología inicial como un ancla.

Cuadro 3.12: Relación de enfoques en el diseño ontológico

(Fuente: Adaptado de Holsapple y Joshi, 2002)

A continuación se describen brevemente los diferentes tipos de enfoques propuestos es el cuadro 3.12.

1. *Enfoque basado en la inspiración:* En este caso, el desarrollador responsable del modelo ontológico empieza desde una premisa acerca de la ontología que se necesita. A través de la imaginación personal, la creatividad, visión personal que tenga del área de conocimiento, en este la competitividad internacional de los territorios, o algún aspecto más específico dentro de esta área, este editor o desarrollador ontológico procede a realizar el diseño ontológico con el objetivo de cumplir las necesidades definidas con anterioridad. Salvo situaciones excepcionales, el acuerdo ontológico para el diseño ontológico resultante es muy limitado.
2. *Enfoque basado en la inducción:* Con este tipo de enfoque, el diseño ontológico se desarrolla a través de la observación, examen y análisis de casos específicos del área de conocimiento de interés. El diseño resultante de este proceso, es,

posteriormente, aplicado a otros casos dentro de la misma área o dominio. El grado de acuerdo, en este caso, viene en función del ámbito de aplicación de otras organizaciones o entornos.

3. *Enfoque deductivo*: En este tipo de enfoque, el procedimiento está relacionado con el establecimiento de algunos principios generales y su adaptación de forma progresiva hacia la construcción ontológica de un caso específico. Es un enfoque, que requiere de un filtrado o extracción de las nociones generales, de tal forma que puedan ser personalizadas al subárea de interés.
4. *Enfoque sintético*: En este caso, el editor ontológico toma como base un conjunto de ontologías, independientes entre sí, y, en ocasiones, se incorporan conceptos relacionados con el área de conocimiento en estudio, y de esta forma, se crea un diseño unificado. Como consecuencia de la incorporación de ontologías preestablecidas inicialmente, este tipo de enfoque puede promover una interacción más coherente entre sus participantes.
5. *Enfoque colaborativo*: En este tipo de enfoques, el diseño ontológico se consigue a través de la participación de varios participantes, que ofrecen sus conocimientos y experiencias para cooperar de forma intencionada en la producción de un diseño ontológico enriquecedor. El nivel de aceptación se verá extendido si estas personas contribuyen de forma diversa y diferenciada en sus propuestas. De esta forma, se puede conseguir un mayor nivel de estabilidad del diseño, así como de riqueza. El punto de inicio del proceso de diseño puede estar marcado por un diseño ontológico base, a partir del cual se empieza la colaboración de los participantes. Es por ello necesario, que se respeten unos criterios de consenso entre las partes involucradas en el desarrollo.

Este último enfoque, *enfoque colaborativo*, podría ser el que mejores resultados pueda producir, pero como ya se indicó anteriormente, requiere de tres aspectos que son necesarios y enriquecedores: un rico y productivo grupo de participantes, el uso de un buen diseño ontológico base y, por último, unos criterios y procedimientos de consenso, eficientes, que permitan una evolución estable y enriquecedora.

En este último capítulo de la tesis, se ha incluido una simulación de la aplicación del sistema propuesto en el que se ha ilustrado como dicho sistema avanzado permite cumplir los requerimientos del sistema, propuesto en este capítulo. Asimismo, se describe cómo se puede gestionar el conocimiento a través de las tecnologías y las capacidades de creación de conocimiento, vitales para la toma de decisiones basadas en los niveles de competitividad de los territorios.

Conclusiones

y

Recomendaciones

En este último apartado de la tesis, se exponen las conclusiones más significativas que han resultado de este trabajo de investigación, así como las principales implicaciones académicas y prácticas y las limitaciones encontradas en esta investigación. Finalmente, se proponen algunas recomendaciones y posibles extensiones para el área objeto de estudio.

Esta investigación ha contribuido a cubrir la laguna existente en la actualidad, relacionada con la aplicación de las tecnologías de la información y el conocimiento en la gestión del conocimiento explícito. En particular, se dirigió a potenciar las capacidades de análisis de los estudios de competitividad internacional, a través de la aplicación de tecnologías avanzadas para la gestión del conocimiento.

Con el fin de cumplir el objetivo principal de esta tesis, relacionado con la definición de un sistema avanzado de gestión del conocimiento aplicado a la competitividad internacional, se han establecido las bases teóricas de las áreas que apoyan esta investigación. En este sentido, en el capítulo 1 – *El estudio de la competitividad internacional* –, se realizó una revisión de los modelos teóricos de competitividad y se presentaron los modelos empíricos de mayor relevancia para medir la competitividad de los territorios. Asimismo, en el capítulo 2 – *Tecnologías aplicadas a la gestión del conocimiento* –, se realizó una revisión teórica de las tecnologías que pueden ser aplicadas a la gestión del conocimiento explícito, más relevantes para el estudio de la competitividad. Además, en este segundo capítulo se introdujo un apartado sobre las tendencias tecnológicas para los próximos años, que podrían contribuir a marcar las directrices de los futuros trabajos de investigación.

Tomando como referencia estos dos primeros capítulos de la tesis, en el capítulo 3 – *Sistema avanzado de gestión del conocimiento aplicado a la competitividad internacional* –, se desarrolló una propuesta tecnológica para la gestión de la competitividad y se presentó una simulación del sistema propuesto. Las principales conclusiones obtenidas a raíz de la elaboración de este último capítulo, se presentan a continuación.

Principales resultados

El resultado más importante de esta investigación ha sido satisfacer con éxito el objetivo propuesto. Este objetivo, consistía en definir y desarrollar desde un punto de vista teórico-conceptual, con una perspectiva tecnológica, un sistema software basado en los modelos de competitividad internacional, que cubriera y extendiera sus fortalezas, que eliminara sus debilidades, y que funcionara como plataforma modular, autónoma y abierta para el desarrollo de bases de conocimiento actualizable, y que permitiese, a su vez, el descubrimiento automático de nuevo conocimiento para la competitividad internacional de los territorios.

Para ello, se desarrolló un sistema avanzado de gestión del conocimiento aplicado a la competitividad internacional. A raíz de la naturaleza del problema a tratar, y desde un punto de vista metodológico, se estableció la metodología de desarrollo en cascada o *waterfall model*, que permitió definir los pasos metodológicos hasta la construcción del modelo de sistema. Posteriormente, para la fase de implementación, se propuso la aplicación de metodologías ágiles que permitiese un rápido desarrollo de aplicaciones finales, basadas en el modelo de sistema propuesto en esta investigación. Por otro lado, se debe argumentar en este punto, que los posibles futuros desarrollos tecnológicos basados en este sistema no están sujetos a una perspectiva única de desarrollo. En este sentido, el sistema propuesto abre un abanico de posibilidades que van desde la definición directa de modelos ontológicos de competitividad, sobre el componente Protégé-2000, hasta portales Web de competitividad internacional, basados en *e-services* y agentes inteligentes.

En este sentido, el sistema propuesto satisface el objetivo principal de la tesis, comentado anteriormente. Para definir los requerimientos del sistema propuesto, se analizó el informe de competitividad internacional, y se identificaron las principales ventajas y desventajas de dicho informe. Posteriormente, y tras un análisis meticuloso de las tecnologías existentes en la actualidad, se planteó una propuesta tecnológica viable y de bajo coste basada en cuatro componentes básicos. Un gestor de ontologías denominado Protégé-2000, que permitiera presentar el conocimiento explícito basado en estándares internacionales y abiertos. Un segundo componente clave en este sistema, denominado Jess, y desarrollado en el lenguaje Java, que

realizase las funciones de motor de inferencias y permitiera inferir el conocimiento explícito representado en el gestor ontológico Protégé-2000.

Además, un tercer componente, dirigido a la automatización de tareas y a la movilidad de agentes inteligentes, se integró en el sistema. Este componente, llamado JADE, integrable con los dos anteriores, permitía extender las capacidades del sistema en los aspectos de automatización y creación de sistemas multiagentes. Finalmente, un cuarto y último componente clave denominado Yale-WEKA, fue integrado al sistema. Este último componente, permitía el descubrimiento de nuevo conocimiento explícito, mediante un conjunto de herramientas, desarrolladas con el lenguaje Java, y basadas en técnicas de minería de datos, que incluyen, además, la resolución de problemas mediante redes neuronales artificiales, algoritmos genéticos, entre otras otras.

Un aspecto primordial de un sistema basado en componentes, es su nivel de integración. En este sentido, en esta investigación y en relación a la selección de los componentes, se tuvo especial cuidado en el nivel de integración de los componentes propuestos. De esta forma, el sistema presentado permite: 1) gestionar conocimiento a través de técnicas estándar de representación simbólica del conocimiento explícito; 2) realizar inferencias sobre dicho conocimiento basado en modelos ontológicos; 3) gestionar de forma avanzada tareas automáticas y distribuidas; y, por último, 4) descubrir conocimiento, mediante tecnologías basadas en representación no-simbólica del conocimiento explícito. Todo ello, y gracias a este nivel de integración, se ha conseguido una sinergia funcional que va más allá de la suma de las funcionalidades individuales de cada uno de los componentes.

La validación del sistema propuesto frente a los requerimientos expresados en el cuadro 3.6, *Requerimientos del sistema propuesto*, y establecidos en la fase de análisis, se realizó en el apartado 3.4.- *Simulación práctica del sistema* –, del capítulo 3 de esta tesis. En esta simulación, se satisfacen dichos requerimientos mediante una serie de procesos que se resumen en:

- La conversión de datos a modelo ontológico mediante la interconexión del gestor ontológico Protégé-2000, con la base de datos en MySQL, donde se almacenan los datos del informe base.

- La administración del modelo ontológico generado de forma automática en el proceso anterior. De esta forma, el modelo ontológico resultante, así como los datos que representa dicho modelo, pueden ser gestionados de forma versátil a través de este componente.
- La integración del motor de inferencia Jess, con el gestor de ontologías Protégé-2000. Este proceso se hace posible a través de una adaptación software (interface), que permite establecer un lazo de interconexión entre ambos componentes.
- La vinculación de las *facts* almacenadas en el motor de inferencia con el conocimiento explícito, basado en el modelo ontológico, administrado desde gestor de ontologías. De esta forma, cualquier cambio experimentado por el conocimiento explícito a través de modificaciones realizadas desde el gestor ontológico, pasarían de forma directa a la base de conocimiento mantenida en el motor de inferencias y viceversa.
- La producción de código Java, basado en el modelo ontológico establecido, y que definen el comportamiento de agentes inteligentes gestionados desde el componente JADE. Esta producción se realiza desde el gestor ontológico Protégé-2000. Este es un ejemplo más del nivel de integración que soportan los componentes del sistema propuesto.
- El descubrimiento de un árbol de decisión sobre el nivel de competitividad internacional de los territorios, haciendo uso del componente Yale-WEKA del sistema propuesto, y partiendo de la tabla de datos generada a partir del índice de competitividad mundial, el ranking de los factores de competitividad, y el PIB-PPC (Producto Interior Bruto a paridad de poder de compra) del informe base.
- Y, por último, la conversión del árbol de decisión generado en el proceso anterior, a regla de decisión que puede incrementar el conocimiento almacenado en la base de conocimiento en Jess.

En este sentido, se puede argumentar que el objetivo de esta tesis ha sido alcanzado y contrastado a través de la simulación práctica presentada en el último capítulo de esta tesis.

Por otro lado, en el apartado 3.5 – *Accesos y usuarios del sistema* -, se presentan los diferentes métodos de acceso que pueden ser utilizados para este sistema, así como la tipología básica de los posibles usuarios y las posibilidades funcionales de los mismos.

Asimismo, también se debe considerar, que a raíz de la realización de esta investigación, se han obtenido otros resultados, como son, por un lado, la demostración de cómo se puede aplicar tecnologías de vanguardia para la gestión de información y conocimiento, dirigidas a gestionar conocimiento sobre competitividad internacional de los territorios. Y, por otro lado, el sistema propuesto abre una puerta de mucho interés para la realización de investigaciones aplicadas, en el área de organización de empresas, donde se integran tecnologías de vanguardia como gestores ontológicos, motores de inferencia de conocimiento, sistemas de agentes y sistemas de *knowledge discovery*, basados en minería de datos. Con todo ello, se ha pretendido aportar un método para demostrar cómo se puede cubrir la laguna existente en las investigaciones relacionadas con la gestión del conocimiento, donde se ha estado infravalorando el papel de las tecnologías en esa área de conocimiento.

Por último, el sistema propuesto en esta tesis, va a permitir avanzar en el conocimiento sobre competitividad internacional, más allá de lo que se puede conseguir en la actualidad a través de los informes de competitividad.

Implicaciones académicas

Desde una perspectiva académica, se podría afirmar que ésta es la primera vez que se integran todos estos componentes tecnológicos (Protégé-2000, Jess, JADE, Yale-WEKA), de forma conjunta, como un solo sistema. Lo que podría infundir interés en futuras investigaciones en el área de integración de componentes tecnológicamente

avanzados y con aplicaciones diferentes. Se debe recordar que uno de los aspectos más destacados en la actualidad, responsable de una parte importante al gasto empresarial en TICs, es la integración tecnológica. Asimismo, esta investigación describe una forma de gestionar el conocimiento explícito sobre competitividad, utilizando tecnologías abiertas y en sintonía, permitiendo la fusión del sistema propuesto con Internet.

Por otro lado, esta investigación abre una nueva vía de aplicación de conocimiento sobre competitividad, facilitando el desarrollo de modelos ontológicos de esta área y la realización de investigaciones parciales o específicas de competitividad (país, sector, empresa).

Implicaciones prácticas

Entre los resultados más significativos que se han obtenido de esta investigación, cabría destacar los siguientes:

- La posibilidad de crear y gestionar modelos ontológicos de competitividad estándar y específicos en varios formatos estándar. Esto convierte los informes de competitividad en sistemas avanzados dinámicos y abiertos, de información y conocimiento, sobre competitividad internacional.
- La gestión de bases de datos y conocimiento sobre competitividad internacional, con características que no tienen precedentes, por ser: completa, incremental, longitudinal, actualizable de forma automática, versátil y multiusuario.
- Permitir un nivel de estudio de la competitividad internacional muy superior a los informes actuales debido a las capacidades de los diferentes componentes que conforman el sistema, y que facilitan el análisis de datos.
- Permite trabajar desde diferentes perspectivas: por niveles de usuario, por tipología descriptiva (datos, información, conocimiento), de forma local, remota y distribuida, por tipos accesos, etc..

- Permite automatizar una gran variedad de tareas y la evolución automática del conocimiento mediante el componente de *knowledge discovery* (Yale –WEKA), o mediante inferencia de conocimiento a través del motor de inferencias.
- Permite ofrecer *e-services* a través de los agentes definidos en JADE, en combinación con ciertos estándares existentes en Internet e incluso con extensión a la telefonía móvil.

Limitaciones de la investigación

Entre las limitaciones más importantes que se pueden observar se encuentran, por un lado, el nivel de especialización necesario para maximizar el rendimiento de este sistema, ya que las áreas de conocimientos necesarias son de diferentes naturalezas. En primer lugar, el alto nivel del conocimiento sobre la competitividad internacional y los modelos cognitivos que permitan describir dicha competitividad. En segundo lugar, el conocimiento técnico específico, necesario para gestionar, de forma eficiente, cada uno de los componentes que conforman el sistema propuesto en la tesis, así como el conocimiento sobre la tecnología que cubre cada componente. Es decir, para utilizar de forma eficiente JADE, el componente que permite la creación de sistemas multiagentes, es necesario saber programar en Java, conocer los estándares utilizados para sistemas de agentes, tales como FIPA o ACL, qué protocolos de comunicación son aplicables con este componente, conocer la teoría fundamental de agentes, tener un buen nivel de conocimiento en lenguajes aplicados a la Web semántica, tales como XML o OWL, etc... Otro claro ejemplo se puede encontrar con el motor de inferencias. Es decir, no sólo es necesario conocer cómo se utiliza JESS, sino también, aspectos tales como los tipos de inferencias posibles y las formas de representación de conocimiento mediante *facts* y reglas.

Por otro lado, la exigencia que se impone en la calidad de los componentes que la forma, y su nivel de integración. Además, a través de este trabajo se ha observado el poco desarrollo existente en esta área, lo que ha dificultado de forma importante la investigación.

Recomendaciones para futuras investigaciones

Entre las recomendaciones para la extensión de este trabajo de investigación y su aplicación en estudios de la competitividad internacional, sería necesario clasificar cinco grandes áreas de estudio: la gestión de datos sobre competitividad, la ontología y la Web para estudios de competitividad, los sistemas de agentes, el motor de inferencia para el desarrollo de estructuras know-how de competitividad y, por último, *knowledge discovery* para el área de competitividad.

En relación a la gestión de datos, sería de interés estudiar la aplicación de ORACLE como gestor de base datos sobre competitividad, ya que ofrece posibilidades para la aplicación de minería de datos y la gestión de datos, desde formatos tales como el XML. No obstante, se deberá buscar los mecanismos de interconexión entre el sistema avanzado de gestión del conocimiento aplicado a la competitividad internacional y ORACLE, para que no se produzca una reducción de la estabilidad de interconexión que en la actualidad está presente en MySQL.

La gestión ontológica que definen los modelos de conocimiento explícito sobre la competitividad internacional puede ser tratada desde un lenguaje de gestión de ontologías que permita cubrir con facilidad el conocimiento *know-that* y el conocimiento *know-how*, como es posible con el OWL. Para que dicha aplicación pueda realizarse con éxito, sería necesaria la coordinación con los equipos de desarrollo internacional que cooperan en la implementación y evolución de los diferentes componentes que conforman el sistema propuesto en esta tesis, con el fin de seguir manteniendo el nivel de compatibilidad y estabilidad entre dichos componentes, algo bastante difícil de conseguir debido a la constante evolución tecnológica, no siempre sincronizadas, de las tecnologías utilizadas en este caso. Esta extensión podría suponer una mayor flexibilidad para la publicación en la Web, no solo del conocimiento explícito que se pueda poseer y el que se pueda descubrir, sino del conocimiento "know-how" que se pueda obtener después de aplicar el sistema propuesto. Además, sería de gran interés internacional que se pudieran estandarizar descripciones semánticas de conceptos fundamentales sobre competitividad internacional y los modelos ontológicos.

En lo referido a la gestión de los sistemas de agentes, cabe destacar la automatización de la actualización y mantenimiento de los datos, la gestión automática de tareas, el

incremento de canales de integración de Yale–Weka a través de los agentes y la evolución supervisada y automática de los modelos ontológicos.

En definitiva, esta investigación, con el desarrollo del sistema propuesto, abre un abanico de posibilidades en la gestión y descubrimiento de conocimiento sobre competitividad internacional, que será de interés para el mundo académico, organismos públicos, locales, nacionales e internacionales, y para el sector privado. En el poco tiempo que lleva de vida, ha demostrado unas posibilidades que deben ser desarrolladas para permitir el enriquecimiento del conocimiento que en estos días se tiene, sobre la competitividad de los territorios.

Bibliografía

- ABOWD, J., F. KRAMARZ, T. LEMIEUX y D. MANGOLIS** (1999). "Minimum Wages and Youth Unemployment in France and the United States", National Bureau of Economic Research.
- ADLER, M., J. BRADSHAW, M. MAHAN y N. SURI** (2001). "Applying Mobile Agents to Enable Dynamic, Context-Aware Interaction for Mobile Phone Users". *Mobile Agents for Telecommunication Applications: Third International Workshop, MATA 2001, Montreal, Canada, August 14-16, 2001. Proceedings*, Springer-Verlag Heidelberg. 2164: 184--194.
- AGILEALLIANCE** (2003). "Roadmap to Agile methods and tools", Agile Alliance. 2003.<http://www.agilealliance.com/programs/roadmaps/Roadmap/index.htm>.
- AGRAWAL, R., J. GEHRKE, D. GUNOPULOS y P. RAGHAVAN** (1998). "Automatic subspace clustering of high dimensional data for data mining applications." 94--105.
- ALONSO, J. A.** (1992). "Ventajas comerciales y competitividad, aspectos conceptuales y empíricos." *Información Comercial Española* 705: 38-76.
- ANDERSON, J. R.** (1983). "The Architecture of Cognition", Harvard University Press, Cambridge.
- ANKERST, M., M. BREUNING, H. KRIEGEL y J. SANDER** (1999). "Optics: Ordering Points to Identify the Clustering Structure". *International Conference on Management of Data*, ACM: 49--60.
- BALA, J., S. BAIK, A. HADJARIAN, B. K. GOGIA y C. MANTHORNE** (2002). "Application of a distributed data mining approach to network intrusion detection". *Proceedings of the first international joint conference on Autonomous agents and multiagent systems. @AAMAS02-EDITOR*. Bologna, Italy, ACM Press: 1419--1420.
- BANDINI, S. y S. MANZONI** (2001). "Application of fuzzy indexing and retrieval in case based reasoning for design". *Proceedings of the 2001 ACM symposium on Applied computing*, ACM Press: 462--466.
- BARRO, R.** (1991). "Economic Growth in a Cross Section of Countries." *Quarterly Journal of Economics* 106: 407-444.

- BARRO, R. y X. SALA-I-MARTIN** (1995). "Economic Growth". New York, McGraw Hill.
- BASDOGAN, C., C.-H. HO, M. A. SRINIVASAN y M. SLATER** (2000). "An experimental study on the role of touch in shared virtual environments." *ACM Transactions on Computer-Human Interaction (TOCHI)* 7(4): 443--460.
- BECK, K., M. BEEDLE, A. V. BENNEKUM y E. AL.** (2001). "Principles behind the Agile Manifesto", AgileAlliance. 2003.<http://agilemanifesto.org/principles.html>.
- BERENDT, B., A. HOTHO y G. STUMME** (2002). "Towards Semantic Web Mining". *International Semantic Web Conference*, Conference: 265--278 Location.
- BERGHOFF, J., O. DROBNIK, A. Y. LINGNAU y C. MÖNCH** (1996). "Agent-based configuration management of distributed applications". *International Conference on Configurable Distributed Systems*. Annapolis, MD: 52--59.
- BERNERS-LEE, T.** (2000). "Weaving the Web", TEXERE Publishing Limited.
- BERNERS-LEE, T., J. HENDLER y O. LASSILA** (2001). "The Semantic Web". *Scientific American*.
- BEYER, H.** (1992). "Some Aspects of the 'Evolution Strategy' for Solving TSP-Like Optimization Problems Appearing at the Design Studies of a 0.5 TeV e+e--Linear Collider". *Parallel problem solving from nature 2*. R. MÄNNER y B. MANDERICK. Amsterdam, North-Holland: 361--370.
- BHATTACHARYYA, S.** (2000). "Evolutionary algorithms in data mining: multi-objective performance modeling for direct marketing". *Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining*, ACM Press: 465--473.
- BORIS, R. V.** (2003). "KAON SERVER A Semantic Web Management System". *Proceedings of the 12th World Wide Web, Alternate Tracks - Practice and Experience*.
- BOSENIUK, T.** (1990). "Travelling Salesman Problem: Optimization by Evolution of Interacting Tours". *Evolution and Optimization '89*. H.-M. VOIGT, H. MÜHLENBEIN y H.-P. SCHWEFEL. Berlin, Akademie-Verlag Berlin: 189--198.

- BOSTAN, B. y A. YAZICI** (1998). "A fuzzy deductive object-oriented database model". *The 1998 IEEE International Conference on Fuzzy Systems Proceedings, 1998. IEEE World Congress on Computational Intelligence*, Conference: 1361-1366 Location.
- BOWDEN, R., P. KAEWTRAKULPONG y M. LEWIN** (2002). "Jeremiah: the face of computer vision". *Proceedings of the 2nd international symposium on Smart graphics*, ACM Press: 124--128.
- BRAY, T., J. PAOLI, C. M. SPERBERG-MCQUEEN y E. MALER** (2000). "Extensible Markup Language {{XML}} 1.0 (Second Edition)", W3C Recommendation.
- BREIMAN, L.** (1996). "Bagging predictors." *Machine Learning* 2(13): 30-37.
- BROEKSTRA, J., M. C. A. KLEIN, S. DECKER, D. FENSEL, F. HARMELEN y I. HORROCKS** (2001). "Enabling knowledge representation on the Web by extending {RDF} schema". *World Wide Web*: 467-478.
- BUSCH, P. y D. RICHARDS** (2001). "Graphically defining articulable tacit knowledge". *Selected papers from the Pan-Sydney workshop on Visualisation*, Australian Computer Society, Inc.: 51--60.
- BUSCH, P. A., D. RICHARDS y C. N. G. K. DAMPNEY** (2001). "Visual mapping of articulable tacit knowledge". *Australian symposium on Information visualisation*, Australian Computer Society, Inc.: 37--47.
- BUSH, V.** (1945). "As We May Think". *The Atlantic Monthly*. <http://www.ps.uni-sb.de/~duchier/pub/vbush/vbush0.shtml>.
- CARDELLI, L. y A. D. GORDON** (2000). "Anytime, anywhere: modal logics for mobile ambients". *Proceedings of the 27th ACM SIGPLAN-SIGACT symposium on Principles of programming languages*, ACM Press: 365--377.
- CASTELLS, M.** (2001). "La Galaxia Internet", Plaza & Janés Editores S.A.
- CATTONI, R., T. COIANIZ, S. MESSELODI y C. MODENA** (1998). "Geometric layout analysis techniques for document image understanding: a review."
- CHANK, R.** (1995). "Information is Surprises, chap. Chapter 9". *The Third Culture: Beyond the Scientific Revolution*, Simon & Schuster.

CHEESEMAN, P., J. KELLY, M. SELF, J. STUTZ, W. TAYLOR y D. FREEMAN (1988). "Autoclass: A Bayesian Classification System". *Fifth Int. Conf. On Machine Learning*: 54-64.

CHEESEMAN, P. y J. STUTZ (1996). "Bayesian Classification (AutoClass): Theory and Results". *Advances in Knowledge Discovery and Data Mining*: 153-180.

CHIU, K.-S. y A. HUNTER (1999). "GA design of crisp-fuzzy logic controllers". *Proceedings of the 1999 ACM symposium on Applied computing*, Conference: 238--242 Location.

CHO, D. S. (1994). "A Dynamic Approach to International Competitiveness: The case of Korea." *Journal of Far Eastern Business* 1(1): 17:36.

CHO, D. S. y H. C. MOON (2000). "From Adam Smith to Michael Porter, Evolution of Competitiveness Theory", Theory. World Scientific Publishing Co. Pte. Ltd.

COCKBURN, A. (2002). "Agile software development", Addison-Wesley Longman Publishing Co., Inc.

COE, D., E. HELPMAN y A. HOFFMAISTER (1995). "North-South R & D Spillovers". NBER Working Paper 5048.

COE, D. y D. SNOWER (1997). "Policy Complementaries: The Case for Fundamental Labour Market Reform".

COHEN, R. y P. BACDDAYAN (1994). "Organizational routines are stored as procedural memory: evidence from a laboratory study." *Organization Science*(Diciembre): 554-568.

COHEN, S. S. y J. ZYSMAN (1987). "Manufacturing Matters." New York, Basic Books.

COHN, A. G. (1995). "The Challenge of Qualitative Spatial Reasoning." *ACM Computing Surveys* 27(3): 323--325.

COMMUNITIES, C. O. T. E. (2000). "Report From the Comission to the Council and The European Parliament in Accordance with Article 30 of Regulation (EEC). No. 1601/92". Bruselas.

- CONRADI, R. y T. DYBAA** (2001). "An empirical study on the utility of formal routines to transfer knowledge and experience". *Proceedings of the 8th European software engineering conference held jointly with 9th ACM SIGSOFT international symposium on Foundations of software engineering*, ACM Press: 268--276.
- DANIELS, M.** (1999). "Integrating Simulation Technologies With Swarm". *Agent Simulation: Applications, Models and Tools*. U. O. CHICAGO.
- DARWIN, C.** (1859). "On the Origin of Species". London, John Murray.
- DEERING, S. y R. HINDEN** (1995). "IPv6 Specification".
<http://ds.internet.net/rfc/rfc1883.txt>.
- DEVEDZIC, V.** (2002). "Understanding ontological engineering." *Communications of the ACM* 45(4): 136--144.<http://doi.acm.org/10.1145/505248.506002>.
- DONG, G., X. ZHANG, L. WONG y J. LI** (1999). "CAEP: Classification by Aggregating Emerging Patterns". *Discovery Science*: 30-42.
- EFRON, B. y R. TIBSHIRANI** (1993). "An Introduction to the bootstrap". New York, USA, Chapman & Hall.
- EIBEN, A. E., T. J. EUVERMAN, W. KOWALCZYK, E. PEELEN, F. SLISSER y J. A. M. WESSELING** (1996). "Comparing Adaptive and Traditional Techniques for Direct Marketing". *Proceedings of the 4th European Congress on Intelligent Techniques and Soft Computing*. H.-J. ZIMMERMANN, Verlag Mainz: 434--437.
- EL-GERESY, B. A. y A. I. ABDELMOTY** (2001). "Qualitative representations in large spatial databases". *2001 International Symposium on Database Engineering & Applications*, Conference: 68-75 Location.
- ERIKSSON, H.** (2002). "The JessTab approach to Protégé and Jess integration". *Proceedings of Intelligent Information Processing (IIP 2002)*, Conference: 237-248 Location.
- ERIKSSON, R. y B. OLSSON** (1997). "Cooperative Coevolution in Inventory Control Optimisation". *Proceedings of the Third International Conference on Artificial Neural Networks and Genetic Algorithms*. G. D. SMITH, N. C. STEELE y R. F. ALBRECHT. University of East Anglia, Norwich, UK, Springer-Verlag.

ESTER, M., H.-P. KRIEGEL, J. SANDER y X. XU (1996). "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise". *Second International Conference on Knowledge Discovery and Data Mining*. E. SIMOUDIS, J. HAN y U. FAYYAD. Portland, Oregon, {AAAI} Press: 226--231.

ESTY, D. y M. PORTER (2000). "Measuring national environmental performance and its determinants", Oxford University Press.

EZEALA-HARRISON, F. (1999). "Theory and Policy of International Competitiveness". Westport, Connecticut London.

FALEY, L. y L. PRUSAK (1998). "The eleven deadliest sins of knowledge management." *California Management Review* 40(3): 256-277.

FENSEL, D., F. V. HARMELEN, M. KLEIN, H. AKKERMANS, J. BROEKSTRA, C. FLUIT, J. V. D. MEER, H. SCHNURR, R. STUDER, J. HUGHES, U. KROHN, J. DAVIES, R. ENGELS, B. BREMDAL, F. YGGE, T. LAU, B. NOVOTNY, U. REIMER y I. HORROCKS (2000). "On-to-knowledge: Ontology-based tools for knowledge management.", <http://www.cs.vu.nl/~frankh/postscript/eBeW00.pdf>.

@FIPA (1997). "{FIPA} 97 Specification Part 5: Personal Assistant.", <ftp://ftp.fipa.org/Specs/FIPA97/f7a15pdf.zip>.

FISCHER, S., R. KLINKENBERG, I. MIERSWA y O. RITTHOFF (2002). "Yale: Yet Another Learning Environment - Tutorial". CI-136/02, Collaborative Research Center 531, University of Dortmund.

FISHER, D. H. (1987). "Knowledge Acquisition Via Incremental Conceptual Clustering." 139-172.

FLACHSBART, B., W. E. BOND, D. C. S. CLAIR y J. HOLLAND (1994). "Using the ID3 symbolic classification algorithm to reduce data density". *Proceedings of the 1994 ACM symposium on Applied computing*, ACM Press: 292--296.

FORBUS, K. D. (1985). "Qualitative process theory". *Qualitative reasoning about physical systems*. D. BOBROW, MIT press.

FORGY, C. L. (1989). "Rete: A Fast Algorithm for the Many Pattern/ Many Object Pattern Match Problem." *Artificial Intelligence*(19): 17-37.

FRANK, M. R. y J. D. FOLEY (1994). "A pure reasoning engine for programming by demonstration". *Proceedings of the 7th annual ACM symposium on User interface software and technology*, Conference: 95--101 Location.

- FRANKEL, J. A. y D. ROMER** (1996). "Trade and Growth: An Empirical Investigation." NBER Working Paper 5476.
- FREEMAN, W. T., P. A. BEARDSLEY, H. KAGE, K.-I. TANAKA, K. KYUMA y C. D. WEISSMAN** (2000). "Computer vision for computer interaction." ACM SIGGRAPH Computer Graphics 33(4): 65--68.
- GABRIEL, K.** (1995). "Engineering Microscopic Machines". Scientific America. 273: 118-121.
- GANDOLFO, G.** (1998). "International Trade Theory and Policy", Springer.
- GARELLI, S.** (2000). "Competitiveness of Nations: The Fundamentals". IMD, Switzerland., The World Competitiveness Yearbook.
- GARUD, R. y P. NAYYAK** (1994). "Transformative capacity: continual structuring by interporal tecnology transfer." Strategic Management Journal 15(5): 365-385.
- GEENHUIZEN, M. V.** (1993). "Knowledge transfer: managing change under increased uncertainty". *The economics of science parks*. M. GUEDES y P. FORMICA. Amprotec. Brasil.
- GENESERETH, M. R. y R. E. FIKES** (1992). "Knowledge Interchange Format, Version 3.0 Reference Manual." (Logic-92-1).
- GENNARI, J., P. LANGLEY y D. FISHER** (1989). "Models of Incremental Concept Formation." Artificial Intelligence 40: 11-62.
- GENNARI, J., M. A. MUSEN, R. W. FERGERSON, W. E. GROSSO, M. CRUBÉZY, H. ERIKSSON, N. F. NOY y S. W. TU** (2002). "The Evolution of Protégé: An Environment for Knowledge-Based Systems Development", Stanford Medical Informatics.
- GEORGE, J. H. y P. LANGLEY** (1995). "Estimating Continuous Distributions in Bayesian Classifiers". *Uncertainty in Artificial Intelligence: Proceedings of the Eleventh Conference (UAI-1995)*. San Francisco, CA, Morgan Kaufmann Publishers: 338-345.
- GLOVER, F.** (1986). "Future paths for Integer Programming and Links to Artificial Intelligence." Computers and Operations Research(5): 533-549.
- GLOVER, F.** (1990). "Tabu Search: A Tutorial." Interfaces 4(20): 74-94.

- GORDON, B. D.** (2002). "Anytime/Anyplace Computing and the Future of Knowledge Work." *Communications of the ACM* 45(12): 67--73. <http://doi.acm.org/10.1145/585597.585617>.
- GRANT, R.** (1996a). "Prospering in dynamically-competitive environments: Organizational capability as knowledge integration." *Organizational Science* 7(4): 375-388.
- GRANT, R.** (1996b). "Toward a knowledge-based theory of the firm." *Strategic Management Journal* 17(Winter Special Issue): 109-122.
- GRAFENSTETTE, J. J., R. GOPAL, B. ROSMAITA y D. G. GUCHT** (1985). "Genetic Algorithms for the Traveling Salesman Problem". *Proc. of an Int. Conf. on Genetic Algorithms and Their Applications*. J. J. GRAFENSTETTE. Hillsdale, NJ, Lawrence Erlbaum: 160--168.
- GROSKY, W. I.** (1997). "Managing multimedia information in database systems." *Communications of the ACM* 40(12): 72--80.
- GROSOFF, B. N. y T. C. POON** (2003). "SweetDeal: representing agent contracts with exceptions using XML rules, ontologies, and process descriptions". *Proceedings of the twelfth international conference on World Wide Web*, Conference: 340--349 Location, ACM Press, <http://doi.acm.org/10.1145/775152.775200>.
- GROUP, W. C. H. W.** (2003). "*XHTML* 2.0", The World Wide Web Consortium.
- GRUBER, T. R.** (1995). "Toward principles for the design of ontologies used for knowledge sharing." *Int. J. Hum. Comput. Stud.* 43(5/6): 907-928.
- GRUNINGER, M. y J. LEE** (2002). "Ontology applications and design: Introduction." *Communications of the ACM* 45(2): 39 - 41.
- HALL, M. A.** (1999). "Correlation-based feature selection for machine learning". Department of Computer Science. Hamilton, New Zealand, University of Waikato.
- HAN, J. y M. KAMBER** (2001). "*Data Mining: Concepts and Techniques*", Morgan Kaufman.

HAUGENEDER, H., STEINER y F. G. MCCABE (1994). "*{IMAGINE}: {A} framework for building multi-agent systems*". *Proceedings of the 1994 International Working Conference on Cooperating Knowledge Based Systems (CKBS-94)*. S. M. DEEN. UK, DAKE Centre, University of Keele: 31--64.

HEDLUND, G. (1994). "A model of knowledge management and the N-form corporation." *Strategic Management Journal* 15(73-90).

HENDLER, J., T. BERNERS-LEE y E. MILLER (2002). "*Integrating Applications on the Semantic Web*." Vol 122(10): 676-680.<http://www.w3.org/2002/07/swint>.

HERTZ, J., A. KROGH y R. G. PALMER (1991). "*Introduction to the theory of neural computation*", Addison-Wesley.

HERTZ, J. A. (1989). "A Gauge Theory in Computational Vision: A Model for Outline Extraction." *pscrip* 39: 161--167.

HINNEBURG, A. y D. A. KEIM (1998). "*An Efficient Approach to Clustering in Large Multimedia Databases with Noise*". *Proceedings of the 4th International Conference on Knowledge Discovery and Data Mining, {(KDD98)}*: 58--65.

HOLSAPPLE, C. W. y K. D. JOSHI (2002). "A collaborative approach to ontology design." *Communications of the ACM* 45(2): 42--47.<http://doi.acm.org/10.1145/503124.503147>.

HOPFIELD, J. J. (1982). "Neural Networks and Physical Systems with Emergent Collective Computational Abilities". *Proceedings of the National Academy of Sciences*, Conference: Location.

HOUSTIS, E. N., A. C. CATLIN, J. R. RICE, V. S. VERYKIOS, N. RAMAKRISHNAN y C. E. HOUSTIS (2000). "*PYTHIA-II: a knowledge/database system for managing performance data and recommending scientific software*." *ACM Transactions on Mathematical Software (TOMS)* 26(2): 227--253.

HUDLICKA, E. y K. CORKER (1988). "*Integrating causal reasoning at different levels of abstraction*". *Proceedings of the first international conference on Industrial and engineering applications of artificial intelligence and expert systems*, ACM Press: 157--163.

INOKUCHI, A., T. WASHIO y H. MOTODA (2003). "*Complete Mining of Frequent Patterns from Graphs: Mining Graph Data*." *Machine Learning* 50(3): 321--354.

INSTITUTE"), A. N. S. (1992). "American national standard for information systems: database language --- SQL: ANSI {X3}.135-1992". pub-ANSI:adr, pub-ANSI.

IST-EU (2003). "Lightweight Extensible Agent Platform", LEAP Consortium. 2003.<http://leap.crm-paris.com/index.html>.

JAMIL, H. M. (2002). "Bottom-Up Association Rule Mining in Relational Databases." *Journal of Intelligent Information Systems* 19(2): 191--206.

JIAWEI, H. y F. YONGJIAN (1994). "Dynamic Generation and Refinement of Concept Hierarchies for Knowledge Discovery in Databases". *{KDD} Workshop*: 157-168.

JIN, C., D. WEI, S. H. LOW, G. BUHRMASTER, J. BUNN, D. H. CHOE, R. L. A. COTTRELL, J. C. DOYLE, W. FENG, O. MARTIN, H. NEWMAN, F. PAGANINI, S. RAVOT y S. SINGH (2003). "FAST TCP: From Theory to Experiments." Presentado a la IEEE Communications Magazine, Internet Technology Series.<http://netlab.caltech.edu/pub/papers/fast-030401.pdf>.

JONES, R. W. (1988). "Heckscher-Ohlin trade theory", The Macmillan Press Limited.

KARNOUSKOS, S. y A. VASILAKOS (2002). "Active electronic mail". *Proceedings of the 2002 ACM Madrid, Spainsymposium on Applied computing*, Conference: 801--806 Location, ACM Press, <http://doi.acm.org/10.1145/508791.508947>.

KARYPIS, G., E.-H. S. HAN y V. KUMAR (1999). "Chameleon: Hierarchical Clustering Using Dynamic Modeling." *Computer* 32(8): 68--75.

KAUFMAN, L. y P. J. ROUSSEEUW (1990). "Finding Groups in Data: An Introduction to Cluster Analysis", John Wiley & Sons.

KING, R. y R. LEVINE (1993). "Finance and Growth: Schumpeter Might be Right." *Schumpeter Might be Right*. *Quarterly Journal of Economics*, 108: 717-738.

KIRA, K. y L. RENDELL (1992). "The feature selection problem: Traditional methods and a new algorithm". *The Proceedings of the Tenth National Conference on Artificial Intelligence*, Conference: 129-134 Location, AAAI Press.

- KIRKPATRICK, S., C. D. GELATT y M. P. VECCHI** (1983). "Optimization by Simulated Annealing." *Science*, Number 4598, 13 May 1983 220, 4598: 671--680.
- KNACK, S. y P. KEEFER** (1995). "Institutions and Economic Performance". New York, Association of Economics and Politics.
- KOHAVI, R. y G. H. JOHN** (1997). "Wrappers for feature subset selection." *Artificial Intelligence Journal*, Special Issue on Relevance(97(1-2)): 273-324.
- KOHONEN, T.** (1995). "Self-organizing maps". Berlin, Springer.
- KOLODNER, J.** (1983). "Maintaining Organization in a Dynamic Long-Term Memory". *Cognitive Science*. E. CHARNIAK, D. NORMAN y E. SMITH, Norwood, NJ. Ablex Publishing Company. 7: 243-280.
- KOUBARAKIS, M.** (1993). "Representation and querying in temporal databases: the power of temporal constraints". *Proceedings. Ninth International Conference on Data Engineering*, Conference: 327-334 Location.
- KRUGMAN, P.** (2000). "Ricardo's Difficult Idea.", <http://web.mit.edu/krugman/www/ricardo.htm>.
- KUIPERS, B.** (1977). "Modeling Spatial Knowledge". *{IJCAI}*: 292-298.
- KUIPERS, B.** (1986). "Qualitative Simulation." *Artificial Intelligence*, 1986. Reprinted in *Qualitative Reasoning about Physical Systems*, ed. Daniel Weld and J. De Kleer, Morgan Kaufmann, 1990, p.236-260 26: 289--338.
- KUIPERS, B.** (1994). "Qualitative Reasoning, Modeling and Simulation with incomplete knowledge", MIT Press, Cambridge, MA, Artificial Intelligence Series.
- KULKARNI, A. D. y V. K. MUNIGANTI** (1996). "Fuzzy neural network models for clustering". *Proceedings of the 1996 ACM symposium on Applied Computing*, Conference: 523--528 Location, ACM Press.
- KURUTACH, W.** (1995). "Modelling fuzzy interval-based temporal information: a temporal database perspective". *Proceedings of 1995 IEEE International Conference on Fuzzy Systems, 1995. International Joint Conference of the Fourth IEEE International Conference on Fuzzy Systems and The Second International Fuzzy Engineering Symposium*, Conference: 741-748 Location.
- LACHER, M. Y. y M. KOCH** (2000). "An agentbased knowledge management framework". *Proc. AAAI Spring Symposium 2000*, Conference: 145-147 Location.

- LANCASTER, K. J.** (1979). "Variety, Equity and Efficiency". New York, Columbia University Press.
- LAWTON, T. C.** (1999). "European Industrial Policy and Competitiveness, Concepts and Instruments." Macmillan Press Ltd.
- LEE, W. W., C. CHIU y B. KUIPERS** (1987). "Developments Towards Constructing Qualitative Simulation". *Qualitative Reasoning Workshop Abstracts*, Qualitative Reasoning Group, University of Illinois at Urbana-Champaign.
- LEFF, L. L.** (2001). "Automated reasoning with legal XML documents". *Proceedings of the 8th international conference on Artificial intelligence and law*, Conference: 215--216 Location, ACM Press, <http://doi.acm.org/10.1145/383535.383560>.
- LENT, B., A. N. SWAMI y J. WIDOM** (1997). "Clustering Association Rules". *ICDE*: 220-231.
- LEONARD, D. y S. SENSIPER** (1998). "The role of tacit knowledge in group innovation." *California Management Review* 40(3): 112-131.
- LEONTIEF, W. W.** (1953). "Domestic Production and Foreign Trade: The American Position Reexamined", New York, Oxford University Press, Chap. 5.
- LI, B. y D. SHASHA** (1998). "Free parallel data mining". *Proceedings of the 1998 ACM SIGMOD international conference on Management of data*, ACM Press: 541-543.
- LINDER, S. B.** (1961). "An Essay on Trade and Transformation". New York, Wiley.
- LIU, B., W. HSU y Y. MA** (1998). "Integrating Classification and Association Rule Mining". *Knowledge Discovery and Data Mining*: 80-86.
- LLOYD-REASON, L. y S. WALL** (2000). "Dimensions of Competitiveness, Theory and Policies", Edward Elgar Publishing Limited.
- LU, H., L. FENG y J. HAN** (2000). "Beyond intratransaction association analysis: mining multidimensional intertransaction association rules." *ACM Transactions on Information Systems (TOIS)* 18(4): 423--454.
- LUCK, M., P. MCBURNEY y C. PRIST** (2003). "Agent Technology: Enabling Next Generation Computing", AgentLink.

LYLES, M. A. y H. K. STEEMSMA (1997). "An examination of social exchange and knowledge based theories in explaining IJV performance and survival in a transitional economy", University of Cambridge.

M.PORTER (1990). "The Competitive Advantage of Nations." The Free Press.

M.PORTER y R. CHRISTENSEN (1999). "Microeconomic competitiveness: findings from the 1999 Executive Survey.The Global Competitiveness Report 1999." Oxford University, World Economic Forum.

MAES, P. (1994). "{Agents that Reduce Work and Information Overload}." CACM 37(7): 31--40.

MAGUIRE, L. P., B. ROCHE, T. M. MCGINNITY y L. J. MCDAID (1998). "Predicting a Chaotic Time Series using Fuzzy Neural network." Information Sciences 112(1-4): 125-136.

MANDHANI, B., S. JOSHI y K. KUMMAMURU (2003). "A matrix density based algorithm to hierarchically co-cluster documents and words". *Proceedings of the twelfth international conference on World Wide Web*, ACM Press: 511--518.

MANGINA, E. (2002). "Review of Software Products for Multi-Agents Systems", AgentLink.

MARCUS, S. y V. S. SUBRAHMANIAN (1996). "Foundations of multimedia database systems." Journal of the ACM (JACM) 43(3): 474--523.

MATTHEWS, J. (1999). "Associative Neural Networks". <http://www.generation5.org/content/1999/associate.asp>.

MCCARTHY, J. (1980). "Circumscription---A Form of Non-Monotonic Reasoning." Artificial Intelligence 13: 27--39.

MCCULLOCH, W. S. y W. PITTS (1943). "A Logical Calculus of the Ideas Immanent in Nervous Activity". *Bull. Math. Biophysics*, Conference: 115-33 Location.

MCDERMOTT, D., M. Y. BURSTEIN y D. SMITH (2001). "Overcoming Ontology Mismatches in Transactions with Self-Describing Agents". *swws01-booktitle*. swws01-address: 285--302.

- MCKAY, D., T. FININ y A. O'HARE** (1990). "The Intelligent Database Interface: Integrating {AI} and Database Systems". *Proceedings of the 8th National Conference on Artificial Intelligence (AAAI-90)*. T. DIETTERICH y W. SWARTOUT. Boston, MA, USA, AAAIP: 677--684.
- MEECH, J. F., K. BAKER, E. LAW y R. LISCANO** (2000). "A multi-agent system for personal messaging". *Proceedings of the fourth international conference on Autonomous agents*, Conference: 144--145 Location, ACM Press, <http://doi.acm.org/10.1145/336595.337093>.
- MEO, R., G. PSAILA y S. CERI** (1998). "An Extension to SQL for Mining Association Rules." *Data Mining and Knowledge Discovery* 2(2): 195--224.
- MINSKY, M.** (1974). "A Framework for Representing Knowledge", MIT-AI Laboratory.
- MINSKY, M.** (1975). "A Framework for Representing Knowledge". *The Psychology of Computer Vision*, New York: McGraw-Hill: 211-277.
- MINSKY, M.** (1980). "A Framework for Representing Knowledge". *Frame Conceptions and Text Understanding*. D. METZING. Berlin, de Gruyter: 1-25.
- MOON, H. C., A. M. RUGMAN y A. VERBEKE** (1995). "The generalized double diamond approach to international competitiveness", En Alan M. Rugman, editor, *Research in Global Strategic Management : A Research Annual*.
- MYLOPOULOS, J. y H. J. LEVESQUE** (1984). "An Overview of Knowledge Representation". *On Conceptual Modelling: Perspectives from Artificial Intelligence, Databases, and Programming Languages*. M. L. BRODIE, J. MYLOPOULOS y J. W. SCHMIDT. New York, Springer: 3-17.
- NELSON, R. R. y S. G. WINTER** (1982). "An evolutionary theory of economic change". Cambridge, MA, Inglaterra.
- NEWELL, A. y H. A. SIMON** (1961). "GPS, A Program that Simulates Human Thought". *Lernende Automaten*. H. BILLING. München, Oldenbourg: 109-124.
- NEWELL, A. y H. A. SIMON** (1972). "Human Problem Solving". Englewood Cliffs, NJ, Prentice-Hall.
- NONAKA, I. y H. TAKEUCHI** (1995). "The knowledge-creating company". New York, Estados Unidos.

- NONAKA, I., R. TOYAMA y N. KONNON** (2001). "SEC1: ba and leadership: unified model of dynamic knowledge creation". *Managing industrial knowledge: Creation transfer and utilization*. I. NONAKA y D. TEECE. Londres, Inglaterra, SAGE Publications Ltd.: 13-43.
- NORMAN, A. D.** (2000). "El Ordenador Invisible", Ediciones Paidós Ibérica, S.A.
- ORPONEN, P.** (2000). "An Overview of the Computational Power of Recurrent Neural Networks". *Proceedings of the 9th Finnish {AI} Conference {STeP} 200. Symposium on Theory, Finnish {AI} Society, Vaasa, Finland, 2000*: 89-96.
- PAIS, J. y C. PINTO-FERREIRA** (1998). "Search Strategies for Reasoning about Spatial Ontologies."
- PAPADIMITRIOU, C. H., D. SUCIU y V. VIANU** (1996). "Topological Queries in Spatial Databases." 81--92.
- PASTOR, J., D. MCKAY y T. FININ** (1992). "View-Concepts: Knowledge-Based Access to Databases". *International Conference on Information and Knowledge Management, Baltimore*.
- PATEL-SCHNEIDER, P. F. y D. FENSEL** (2002). "Layering the Semantic Web: Problems and Directions". *iswc02-booktitle*. @ISWC02-EDITOR, SV. 2348: 16--29.
- PATEL-SCHNEIDER, P. F. y J. SIMÉEON** (2002). "Building the Semantic Web on {XML}". *iswc02-booktitle*. @ISWC02-EDITOR, SV. 2348: 147--161.
- PELL, P., E. B. GAMBLE, E. GAT, R. KEESING, J. KURIEN, W. MILLAR, P. P. NAYAK, C. PLAUNT y B. C. WILLIAMS** (1998). "A Hybrid Procedural/Deductive Executive for Autonomous Spacecraft". *aa98*. @AA98EDITOR. New York, acmp: 369--376.
- PERKOWITZ, M. y O. ETZIONI** (1994). "Database Learning for Software Agents". *Proceedings of the 12th National Conference on Artificial Intelligence*. Menlo Park, CA, USA, AAAIP: 1485--1485.
- PITOURA, E. y B. BHARGAVA** (1995). "A framework for providing consistent and recoverable agent-based access to heterogeneous mobile databases." *SIGMOD Record (ACM Special Interest Group on Management of Data)* 24(3): 44--49.
- PIZZUTI, C. y D. TALIA** (2003). "P-AutoClass: Scalable Parallel Clustering for Mining Large Data Sets." *IEEE Transactions on Knowledge and Data Engineering* 15(3): 629--641.

- POLANYI, M.** (1958). "Personal knowledge". Chicago, IL, Estados Unidos, University Chicago Press.
- PORTER, M.** (1990). "¿Dónde Radica la Ventaja Competitiva de las Naciones?" Harvard Deusto Business Review Especial 100.
- PORTER, M., J. SACHS y A. WARNER** (2000). "Executive Summary: Current Competitiveness and Growth Competitiveness", Oxford University Press.
- PURVIS, R. L., V. SAMBAMURTHY y R. W. ZMUD** (2000). "The development of knowledge embeddedness in CASE technologies within organizations." IEEE Transactions on Engineering Management 47(2): 245-257.
- RAMOS, R.** (2002). "La Competitividad Internacional de Canarias".
- REGAN, P. M. y B. M. SLATOR** (2002). "Case-based tutoring in virtual education environments". *Proceedings of the 4th international conference on Collaborative virtual environments*, ACM Press: 2--9.
- REIMER, U. y A. MARGELISCH** (1995). "A hybrid knowledge representation approach to reusability of legal knowledge bases". *Proceedings of the fifth international conference on Artificial intelligence and law*, Conference: 246--255 Location, ACM Press.
- REITERER, H., G. MUBLER, T. M. MANN y S. HANDSCHUH** (2000). "INSYDER — an information assistant for business intelligence". *Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval*, ACM Press: 112--119.
- RILEY, G.** (2003). "CLIPS: A Tool for Building Expert Systems". 2003.<http://www.ghg.net/clips/CLIPS.html>.
- RITTHOFF, O., R. KLINKENBERG, S. FISCHER, I. MIERSWA y S. FELSKE** (2001). "YALE: Yet Another Machine Learning Environment". Germany, Department of Computer Science, University of Dortmund: 84-92.
- ROBERT, J. A.** (1998). "A CBR Architecture for Project Knowledge Management." Lecture Notes in Computer Science 1488: 414--424.

- RODHAIN, F.** (1999). "Tacit to explicit: transforming knowledge through cognitive mapping—an experiment". *Proceedings of the 1999 ACM SIGCPR conference on Computer personnel research*, ACM Press: 51--56.
- ROSS, B. H.** (1989). "Some Psychological Results on Case-Based Reasoning". *Case-Based Reasoning Workshop*, Pensacola Beach. Morgan Kaufmann: 144-147.
- ROYCE, W.** (1970). "Management the Development of Large Software Systems". *Proc. Westcon*, Conference: 328-339 Location, IEE CS Press.
- RUGGLES, R.** (1998). "The state of the notion: knowledge management in practice." *California Management Review* 40(3): 80-89.
- RUGMAN, A. M. y J. D'CRUZ** (1991). "Fast Forward: Improving Canada's International Competitiveness". Toronto, Kodak Canada.
- RUMELHART, D. E., G. E. HINTON y R. J. WILLIAMS** (1987). "Learning Internal Representations by Error Propagation". *Parallel Distributed Processing: Volume 1: Foundations*. D. E. RUMELHART, J. L. MCCLELLAND y OTHERS. Cambridge, MIT Press: 318-362.
- RUMELHART, D. E. y D. ZIPSER** (1985). "Feature Discovery by Competitive Learning." *cogsci* 9: 75--112.
- RUST, R. T. y P. K. KANNAN** (2003). "E-service: a new paradigm for business in the electronic environment." *Communications of the ACM* 46(6): 36--42. <http://doi.acm.org/10.1145/777313.777336>.
- RUST, R. T. y K. N. LEMON** (2001). "E-service and the consumer." *International J. Electronic Commerce* 5(3): 85-102.
- RYLE, G.** (1949). "The Concept of Mind", Huchinson's University Library, Londres.
- SACHS, J. y A. WARNER** (1995). "Economic Reform and the Process of Global Integration".
- SACHS, J. y A. WARNER** (2000). "Globalization and International Competitiveness: Some Broad Lessons of the Past Decade". Oxford University Press, World Economic Forum.

- SANDER, J., M. ESTER, H.-P. KRIEDEL y X. XU** (1998). "Density-Based Clustering in Spatial Databases: The Algorithm GDBSCAN and Its Applications." *Data Mining and Knowledge Discovery* 2(2): 169--194.
- SANDHOLM, T. y V. LESSER** (1995). "Issues in Automated Negotiation and Electronic Commerce: Extending the Contract Net Framework". *First International Conference on Multi-Agent Systems (ICMAS-95)*, Conference: 328-335 Location.
- SANTOS-PEREIRA, C. M. y A. M. PIRES** (2002). "Detection of outliers in multivariate data: a method based on clustering and robust estimators". *Proceedings in Computational Statistics*: 22.
- SARAWAGI, S., S. THOMAS y R. AGRAWAL** (2000). "Integrating Association Rule Mining with Relational Database Systems: Alternatives and Implications." *Data Mining and Knowledge Discovery* 4(2-3): 89--125.
- SATYANARAYANAN, M., B. NOBLE, P. Y. KUMAR y M. PRICE** (1994). "Application-Aware Adaptation for Mobile Computing. Technical Report", Carnegie Mellon University, School of Computer Science: 7.
- SAWY, O. A. E., I. ERIKSSON, S. A. CARLSSON y A. RAVEN** (1997). "Shared knowledge creation spaces around the new product development process", University of Southern California.
- SCHAPIRE, R. E.** (1999). "A brief introduction to boosting". *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence*, Conference: Location, Morgan Kaufmann.
- SCHERRER, B. y F. CHARPILLET** (2002). "Coevolutive planning in markov decision processes". *International Joint Conference on Autonomous Agent and Multiagent Systems*, Conference: 843--844 Location, ACM Press.
- SCHÖN, D. A.** (1983). "The reflective practitioner". Cambridge, MA, Estados Unidos, Harvard University Press.
- SEDBROOK, T. A.** (1998). "A collaborative fuzzy expert system for the Web." *ACM SIGMIS Database* 29(3): 19--30. <http://doi.acm.org/10.1145/313310.313331>.
- SHEIKHOESLAMI, G., S. CHATTERJEE y A. ZHANG** (1998). "WaveCluster: A Multi-Resolution Clustering Approach for Very Large Spatial Databases". *Proc. 24th Int. Conf. Very Large Data Bases, {VLDB}*: 428--439.

- SHINTANI, T. y M. KITSUREGAWA** (1998). "Parallel mining algorithms for generalized association rules with classification hierarchy". *Proceedings of the 1998 ACM SIGMOD international conference on Management of data*, ACM Press: 25--36.
- SIEGELMANN, H. T.** (1998). "Neural Networks and Analog Computation: Beyond the Turing Limit, Boston". *Neural Networks and Analog Computation: Beyond the Turing Limit*, Boston: Birkhauser.
- SIEGELMANN, H. T. y E. D. SONTAG** (1999). "Turing Computability with Neural Networks." *Applied Mathematics Letters* 4: 77-80.
- SIMA, J. y P. ORPONEN** (2001). "Computing with Continuous-Time Liapunov Systems". *33rd {ACM} {STOC}*, ACM.
- SIMS, K.** (1994). "Evolving virtual creatures". *Proceedings of the 21st annual conference on Computer graphics and interactive techniques*, ACM Press: 15--22.
- SLADE, S., R. ARUNKUNDRAM, M. FISH y R. MADHAVAN** (1994). "Qualitative Business Decision Making". *AAAI workshop on AI in Business*,
- SMITH, A.** (1937 (1776)). "An inquiry into the nature and causes of the wealth of nations". New York, P. F. Collier & son Corporation.
- SMITH, B.** (1982). "Reflection and Semantics in a Procedural Language". MIT laboratory for Computer Science.
- SMITH, M. K., D. MCGUINNESS, R. VOLZ y C. WELTY** (2003). "OWL Web Ontology Language (OWL) Guide Version 1.0." The World Wide Web Consortium.
- SONG, S. y S. LEE** (2002). "A Strategy of Dynamic Reasoning in Knowledge-Based System with Fuzzy Production Rules." *Journal of Intelligent Information Systems* 19(3): 303--318.<http://dx.doi.org/10.1023/A:1020193822476>.
- SPARR, T. M.** (1982). "A language for a scientific and engineering database system". *Proceedings of the nineteenth design automation conference*: 865--871.
- SPINGH, M. P.** (1999). "Know-how". *Foundations of Rational Agency*. W. M. Y. R. A., Kluwer Academic Publishers: 105 -132.
- SRIMANI, P. K., W.-C. LEE, S. KUMAR y S. GUPTA** (2002). "Guest Editorial: Special Section on Data Management Systems and Mobile Computing." *IEEE Transactions on Computers* 50(10): 1121-1123.

- STEELE, G. L.** (1990). "Common Lisp the Language, 2nd edition", Digital Press.
- STEPANKOVA, O.** (1992). "An Introduction to Qualitative Reasoning". *Advanced Topics in Artificial Intelligence: Proc. of the International Summer School, Conference: 404--??* Location, Springer.
- STERN, A., A. FRANK y B. RESNER** (1998). "Virtual Petz: A Hybrid Approach to Creating Autonomous, Lifelike Dogz and Catz". *Second Int. Conference on Autonomous Agents*, Conference: 334--335 Location, ACM Press.
- STUCKENSCHMIDT, H.** (2002). "Ontology-Based Information Sharing in Weakly Structured Environments". Faculty of Sciences. Division of Computer Science. Amsterdam, Vrije Universiteit.
- SURE, Y., S. STAAB y R. STUDER** (2002). "Methodology for development and employment of ontology based knowledge management applications." *ACM SIGMOD Record* 31(4): 18--23.
- TAMBE, M., W. L. JOHNSON, R. M. JONES, F. KOSS, J. E. LAIRD, P. S. ROSENBLOOM y K. SCHWAMB** (1995). "Intelligent Agents for Interactive Simulation Environments". *AI Magazine*. 16: 15--39.
- TAMMA, V. A. M. y T. BENCH-CAPON** (2001). "An enriched knowledge model for formal ontological analysis". *Proceedings of the international conference on Formal Ontology in Information Systems*, Conference: 81--92 Location, ACM Press.
- TARDO, J. y L. VALENTE** (1996). "Mobile agent security and telescript". *COMPCON Spring '96 - 41st IEEE International Computer Conference*, Conference: 58-64 Location.
- TEECE, D.** (1998). "Capturing value from knowledge assets: the new economy, markets for know-how, and intangible assets." *California Management Review* 40(3): 55-79.
- TERNA, P.** (2003). "Decision Making and Enterprise Simulation with jES and Swarm". *Seventh Annual Swarm Users/Researchers Conference*, Department of Computer Science & Engineering and the Center for the Study of Biocomplexity at the University of Notre Dame, and the Swarm Development Group (SDG).
- TOH, M. H. y K. Y. TAN** (1998). "Competitiveness of the Singapore Economy: A Strategy Perspective", Singapore University Press.

TOPICMAPS.ORG (2001). "XML Topic Maps (XTM) 1.0: TopicMaps.Org Specification", TopicsMaps.Org. 2003.<http://www.topicmaps.org/xtm/index.html>.

TSCHICHOLD-GÜRMAN, N. (1994). "Fuzzy RuleNet: an artificial neural network model for fuzzy classification". *Proceedings of the 1994 ACM symposium on Applied computing*, Conference: 145--149 Location, ACM Press, <http://doi.acm.org/10.1145/326619.326686>.

TURING, A. M. (1950). "Computing machinery and intelligence", Prentice-Hall.

VALIANT, L. (1988). "Functionality in Neural Nets". *Learning and Knowledge Acquisition*, Proc. AAAI: 629-634.

VERNON, R. (1966). "International Investment and International Trade in the Product Cycle." *Quarterly Journal of Economics* 80: 190-207.

W3C (1999). "HTML 4.01 Specification", W3 Consortium. 2003.<http://www.w3.org/TR/html4/>.

W3C (2003). "OWL Web Ontology Language: Overview W3C Working Draft 31 March 2003", W3C Consortium.<http://www.w3.org/TR/owl-features/#s1.2>.

WANG, F. y E. MCKENZIE (1999). "A multi-agent based evolutionary artificial neural network for general navigation in unknown environments". *Proceedings of the third annual conference on Autonomous Agents*, Conference: 154--159 Location, ACM Press.

WANG, H., J. MYLOPOULOS y S. LIAO (2002). "Intelligent agents and financial risk monitoring systems." *Communications of the ACM* 45(3): 83--88.<http://doi.acm.org/10.1145/504729.504733>.

WANG, W., J. YANG y R. R. MUNTZ (1997). "STING: A Statistical Information Grid Approach to Spatial Data Mining". *Twenty-Third International Conference on Very Large Data Bases*. M. JARKE, M. J. CAREY, K. R. DITTRICH, F. H. LOCHOVSKY, P. LOUCOPOULOS y M. A. JEUSFELD. Athens, Greece, Morgan Kaufmann: 186--195.

WATSON, I. (2002). "Applying Knowledge Management: Techniques for Building Organisational Memories." *Lecture Notes in Computer Science* 2416: 6--??

WEISER, M. (1991). "The Computer for the Twenty-First Century". *Scientific American*: 94-10.

- WEISER, M. y J. S. BROWN** (1997). "The Coming Age of Calm Technology". *Beyond Calculation: The Next Fifty Years of Computing*. P. J. DENNINGS y R. M. METCALFE, Springer-Verlag: 75-90.
- WEIZENBAUM, J.** (1976). "Computer Power and Human Reason". San Francisco, W.H. Freeman & Company.
- WHITE, T., A. BIESZCZAD y B. PAGUREK** (1998). "Distributed Fault Location in Networks Using Mobile Agents". *iata:98:booktitle*. @IATA:98:EDITOR, SV. 1437.
- WILKINS, D. E.** (1988). "Casual Reasoning in Planning." *Computational Intelligence* 4: 373--380.
- WILSON, R. y F. KEIL** (1999). "MIT Encyclopedia of Cognitive Science", Cambridge, MA: MIT Press.
- WINTER, S.** (1987). "Knowledge and competence as strategic assets". *The competitive challenge*. D. TEECE. Carbondale, Estados Unidos, Southern Illinois University Press: 161-204.
- WORLD-BANK** (1999). "Low income countries", World Bank. 2003.<http://193.129.255.248/glossary/GlossaryKtoN.htm>.
- WU, L. G. y H. LU** (2000). "F3MCNN: a fuzzy minimum mean maximum clustering neural network". *Proceedings of the 2000 ACM symposium on Applied computing*, Conference: 10--14 Location, ACM Press, <http://doi.acm.org/10.1145/335603.335615>.
- XIAOWEI, X., M. ESTER, H. P. KRIEGEL y J. SANDER** (1998). "A distribution-based clustering algorithm for mining in large spatial databases". *Data Engineering, 1998. Proceedings., 14th International Conference on*, Conference: 324-331 Location.
- YIU-KAI, N. y A. SEETHARAMAN** (1997). "A query-based horizontal fragmentation approach for disjunctive deductive databases". *Eighth International Workshop on Database and Expert Systems Applications, 1997. Proceedings.*, Conference: 604-609 Location.
- YUEN, H. S., S. HO y J. ZELEZNIKOW** (1992). "Adding Qualitative Reasoning to an Organizational Database for Management Decision Support." *Lecture Notes in Computer Science* 611: 79--??

ZADEH, L. A. (1965). "Fuzzy Sets." *Information Control* 18: 338-353.

ZAKI, M. J., S. PARTHASARATHY y W. LI (1997). "A localized algorithm for parallel association mining". *Proceedings of the ninth annual ACM symposium on Parallel algorithms and architectures*, ACM Press: 321--330.

ZHANG, T., R. RAMAKRISHNAN y M. LIVNY (1996). "BIRCH: an efficient data clustering method for very large databases." 103--114.

ZYSMAN, J. y L. TYSON (1983). "American Industry in International Competition, Government Policies and Corporate Strategies". Ithaca, N.Y., London:, Cornell University Press.

ANEXOS

Anexo I. Relación de países utilizados en el estudio

Id País	Nombre en español	Nombre en inglés
1	ARGENTINA	ARGENTINA
2	AUSTRALIA	AUSTRALIA
3	AUSTRIA	AUSTRIA
4	BÉLGICA	BELGIUM
5	BOLIVIA	BOLIVIA
6	BRASIL	BRAZIL
7	BULGARIA	BULGARIA
8	CANADÁ	CANADA
9	ISLAS CANARIAS	CANARY ISLANDS
10	CHILE	CHILE
11	CHINA	CHINA
12	COLOMBIA	COLOMBIA
13	COSTA RICA	COSTA RICA
14	REPÚBLICA CHECA	CZECH REPUBLIC
15	DINAMARCA	DENMARK
16	ECUADOR	ECUADOR
17	EGIPTO	EGYPT
18	EL SALVADOR	EL SALVADOR
62	ESTONIA	ESTONIA
19	FINLANDIA	FINLAND
20	FRANCIA	FRANCE
21	ALEMANIA	GERMANY
22	GRECIA	GREECE
23	HONG KONG	HONG KONG
24	HUNGRÍA	HUNGARY
25	ISLANDIA	ICELAND
26	INDIA	INDIA
27	INDONESIA	INDONESIA
28	IRLANDA	IRELAND
29	ISRAEL	ISRAEL
30	ITALIA	ITALY
31	JAPÓN	JAPAN
32	JORDANIA	JORDAN
33	COREA	KOREA
34	LUXEMBURGO	LUXEMBOURG
35	MALASIA	MALAYSIA
36	MAURIZIO	MAURITIUS

Anexo I. Relación de países utilizados en el estudio

Id País	Nombre en español	Nombre en inglés
37	MÉXICO	MEXICO
38	HOLANDA	NETHERLANDS
39	NUEVA ZELANDA	NEW ZEALAND
40	NORUEGA	NORWAY
41	PERÚ	PERU
42	FILIPINAS	PHILIPPINES
43	POLONIA	POLAND
44	PORTUGAL	PORTUGAL
61	RUSIA	RUSSIA
45	RUSIA	RUSSIAN FEDERATION
46	SINGAPUR	SINGAPORE
47	REPÚBLICA ESLOVACA	SLOVAK REPUBLIC
63	ESLOVENIA	SLOVENIA
48	SUDÁFRICA	SOUTH AFRICA
49	ESPAÑA	SPAIN
50	SUECIA	SWEDEN
51	SUIZA	SWITZERLAND
52	TAIWÁN	TAIWAN
53	TAILANDIA	THAILAND
54	TURQUÍA	TURKEY
55	UCRANIA	UKRAINE
56	REINO UNIDO	UNITED KINGDOM
57	ESTADOS UNIDOS	USA
58	VENEZUELA	VENEZUELA
59	VIETNÁN	VIETNAM
60	ZIMBABWE	ZIMBABWE

Anexo II. Relación de variables utilizadas en el estudio

Id	Nombre de variable	Factor	Fecha
1	Producto Interior Bruto (PIB). Billones \$USA	Actividad económica	2000
2	PIB-PPP. Estimaciones de billones \$USA a paridad de poder de compra	Actividad económica	2000
3	Gastos de consumo final privado. Billones \$USA.	Actividad económica	2000
4	Gastos de consumo final del gobierno. Billones \$USA.	Actividad económica	2000
5	Inversión Interior Bruta. Billones \$USA.	Actividad económica	2000
6	Ahorro interior bruto. Billones \$USA	Actividad económica	2000
7	Sectores económicos: Agricultura. Relación de los sectores económicos, porcentaje del PIB.	Actividad económica	2000
8	Sectores económicos: Industria. Relación de los sectores económicos, porcentaje del PIB.	Actividad económica	2000
9	Sectores económicos: Servicios. Relación de los sectores económicos, porcentaje del PIB.	Actividad económica	2000
10	PIB real - Crecimiento. Cambio porcentual, calculado a la moneda local sobre una base de precios constantes.	Actividad económica	2000
11	PIB real per capita - Crecimiento. Cambio porcentual, calculado a la moneda local sobre una base de precios constantes.	Actividad económica	2000
12	Gasto de consumo final privado - Crecimiento real. Cambio porcentual, calculado a la moneda local sobre una base de precios constantes.	Actividad económica	2000
13	Inversión interior bruta - Crecimiento real. Cambio porcentual, calculado a la moneda local sobre una base de precios constantes.	Actividad económica	2000
14	Ahorro interior bruto - Crecimiento real. Cambio porcentual, calculado a la moneda local sobre una base de precios constantes.	Actividad económica	1999
15	Gastos de consumo final del gobierno - Crecimiento real. Cambio porcentual calculado a la moneda local sobre una base de precios constantes.	Actividad económica	2000

Anexo II. Relación de variables utilizadas en el estudio

Id	Nombre de variable	Factor	Fecha
16	Sector de la agricultura - Crecimiento real. Cambio porcentual, calculado a la moneda local sobre una base de precios constantes.	Actividad económica	2000
17	Producción industrial - Crecimiento real. Cambio porcentual, calculado a la moneda local sobre una base de precios constantes.	Actividad económica	2000
18	Sector de servicios - Crecimiento real. Cambio porcentual calculado a la moneda local sobre una base de precios constantes.	Actividad económica	2000
19	PIB per capita. \$USA per capita	Actividad económica	2000
20	PIB per capita (PPP). Estimaciones \$USA per capita a paridad de poder de compra.	Actividad económica	2000
21	Gasto de consumo final privado per capita. \$USA per capita.	Actividad económica	2000
22	Gasto de consumo final del gobierno per capita. \$USA per capita.	Actividad económica	2000
23	Inversión interior bruta per capita. \$USA per capita.	Actividad económica	2000
24	Ahorro interior bruto per capita. \$USA per capita.	Actividad económica	2000
25	Previsiones: Crecimiento real del PIB. Cambio porcentual, calculado a la moneda local sobre una base de precios constan	Actividad económica	2001
26	Previsiones: Gasto de consumo privado - Crecimiento real. Cambio porcentual, calculado a la moneda local sobre una base precios constantes.	Actividad económica	2001
27	Previsiones: Inversión interior bruta - Crecimiento real. Cambio porcentual, calculado a la moneda local sobre una base precios constantes.	Actividad económica	2001
28	Previsiones: Inflación. Cambio porcentual	Actividad económica	2001
29	Previsiones: Desempleo. Porcentaje del total de la fuerza laboral	Actividad económica	2001
30	Previsiones: Balance de la cuenta corriente. Porcentaje del PIB/PNP	Actividad económica	2001
31	Balance de la cuenta corriente. Billones \$USA	Actividad económica	2000
32	Balance de la cuenta corriente. Porcentaje	Actividad económica	2000

Anexo II. Relación de variables utilizadas en el estudio

Id	Nombre de variable	Factor	Fecha
	del PIB		
33	Balanza comercial. Billones \$USA	Actividad económica	2000
34	Balanza comercial. Porcentaje del PIB	Actividad económica	2000
35	Balanza de los servicios comerciales. Billones \$USA	Actividad económica	1999
36	Balanza de los servicios comerciales. Porcentaje del PIB	Actividad económica	1999
37	Exportaciones de bienes. Billones \$USA	Actividad económica	2000
38	Exportaciones de bienes. Porcentaje del PIB	Actividad económica	2000
39	Exportaciones de bienes- Crecimiento real. Cambio porcentual, calculado a la moneda local sobre una base de precios constantes.	Actividad económica	2000
40	Exportaciones de los servicios comerciales. Billones \$USA	Actividad económica	1999
41	Exportaciones de los servicios comerciales. Porcentaje del PIB	Actividad económica	1999
42	Exportaciones de los servicios comerciales - Crecimiento real. Cambio porcentual, calculado a la moneda local sobre una	Actividad económica	1999
43	Desglose de exportaciones por sectores económicos: Agricultura. Porcentaje del total de exportaciones	Actividad económica	1999
44	Desglose de exportaciones por sectores económicos: Industria. Relación de los sectores económicos, porcentaje del PIB.	Actividad económica	1999
45	Desglose de exportaciones por sectores económicos: Servicios. Relación de los sectores económicos, porcentaje del PIB.	Actividad económica	1999
46	Importaciones de bienes y servicios comerciales. Billones \$USA	Actividad económica	1999
47	Importaciones de bienes y servicios comerciales. Porcentaje del PIB	Actividad económica	1999
48	Importaciones de los bienes y servicios comerciales - Crecimiento real. Cambio porcentual calculado a la moneda local so	Actividad económica	1999
49	Desglose de importaciones por sector económico: Agricultura. Porcentaje del total de importaciones.	Actividad económica	1999

Anexo II. Relación de variables utilizadas en el estudio

Id	Nombre de variable	Factor	Fecha
50	Desglose de importaciones por sector económico: Industria. Porcentaje del total de importaciones.	Actividad económica	1999
51	Desglose de importaciones por sector económico: Servicios. Porcentaje del total de importaciones.	Actividad económica	1999
52	Ratio de comercio y PIB. (Exportaciones + Importaciones) / (2 x PIB)	Actividad económica	1999
53	Índice de términos comerciales. Valor unitario de las exportaciones sobre el valor unitario de las importaciones (1995=1	Actividad económica	1999
54	Recepciones por turismo. Recepciones por turismo extranjero como porcentaje del PIB.	Actividad económica	1999
55	Flujos de inversión directa en el extranjero. Billones \$USA.	Actividad económica	1999
56	Stocks de inversión directa en el extranjero. Billones \$USA.	Actividad económica	1999
57	Stocks de inversión directa en el extranjero - Crecimiento real. Cambio porcentual, calculado a la moneda local sobre una	Actividad económica	1999
58	Flujos de inversión directa extranjera. Billones \$USA	Actividad económica	1999
59	Stocks de inversión directa extranjera. Billones \$USA	Actividad económica	1999
60	Stocks de inversión directa extranjera - Crecimiento real. Cambio porcentual, calculado a la moneda local sobre una base	Actividad económica	1999
61	Cartera de activos de inversión. Billones \$USA.	Actividad económica	1999
62	Cartera de pasivos de inversión. Billones \$USA.	Actividad económica	1999
63	Empleo. Empleo total en millones.	Actividad económica	2000
64	Empleo. Porcentaje de la población.	Actividad económica	2000
65	Empleo - Crecimiento. Cambio porcentual	Actividad económica	2000
66	Empleo por sector económico: Agricultura. Porcentaje del total de empleo	Actividad económica	2000
67	Empleo por sector económico: Industria. Porcentaje del total de empleo	Actividad económica	2000

Anexo II. Relación de variables utilizadas en el estudio

Id	Nombre de variable	Factor	Fecha
68	Empleo por sector económico: Servicios. Porcentaje del total de empleo	Actividad económica	2000
69	Empleo privado. Total empleo privado en millones	Actividad económica	1998
70	Empleo del Gobierno. Empleo del Gobierno total en millones.	Actividad económica	1998
71	Tasa de paro. Porcentaje de la fuerza laboral	Actividad económica	2000
72	Desempleo de jóvenes. Desempleo de la población menor de 24 años como porcentaje del total de desempleo.	Actividad económica	2000
73	Inflación. Tasa anual media.	Actividad económica	2000
74	Índice del coste de vida. Índice del coste de la cesta de la compra en las principales áreas/zonas, excluyendo la vivienda	Actividad económica	2000
75	Renta de un apartamento. Renta mensual de un apartamento de 3 dormitorios en las principales ciudades, \$USA	Actividad económica	2000
76	Renta de una oficina. Coste de ocupación total (\$USA/m ² por año)	Actividad económica	2000
77	Déficit/Superávit presupuestario del Gobierno. Billones \$USA	Eficiencia del Gobierno	1999
78	Déficit/Superávit presupuestario del Gobierno. Porcentaje del PIB	Eficiencia del Gobierno	1999
79	Deuda interior del Gobierno. Billones de \$USA	Eficiencia del Gobierno	1999
80	Deuda interior del Gobierno. Porcentaje del PIB	Eficiencia del Gobierno	1999
81	Deuda extranjera del Gobierno. Billones \$USA	Eficiencia del Gobierno	1999
82	Deuda extranjera del Gobierno. Porcentaje del PIB	Eficiencia del Gobierno	1999
83	Deuda interior del Gobierno - Crecimiento real. Cambio porcentual, calculado a la moneda local sobre una base de precios	Eficiencia del Gobierno	1999
84	Pago de intereses. Porcentaje de los ingresos actuales	Eficiencia del Gobierno	1999
85	Gestión del Presupuesto Público. En los últimos años, la gestión del presupuesto	Eficiencia del Gobierno	2001

Anexo II. Relación de variables utilizadas en el estudio

Id	Nombre de variable	Factor	Fecha
	público ha mejorado		
86	Reservas Totales (Total Reserves 1999). Incluyendo el oro y las reservas oficiales (con oro a SDR 35 por onza).	Eficiencia del Gobierno	2000
87	Gasto general del Gobierno. Porcentaje del PIB	Eficiencia del Gobierno	1999
88	Recaudación fiscal total. Porcentaje del PIB	Eficiencia del Gobierno	1999
89	Tasa del Impuesto sobre la renta personal efectiva. Porcentaje del PIB per capita.	Eficiencia del Gobierno	2000
90	Impuestos recaudados sobre la renta personal. Sobre los beneficios, ingresos y ganancias del capital, como porcentaje del PIB	Eficiencia del Gobierno	1999
91	Tasa de Contribución a la Seguridad Social de los Empleados. Contribución obligatoria como % del PIB per capita.	Eficiencia del Gobierno	2000
92	Contribución a la Seguridad Social de los empleados recaudada. Contribución obligatoria como % del PIB.	Eficiencia del Gobierno	1999
93	Impuestos personales reales. Los impuestos personales reales (después de deducciones) no disuaden a las personas de trabajar.	Eficiencia del Gobierno	2001
94	Tasa Impositiva media sobre los beneficios empresariales. Porcentaje de los beneficios antes de impuestos.	Eficiencia del Gobierno	1999
95	Impuestos Empresariales Recaudados. Sobre los beneficios, ingresos y ganancias del capital, como porcentaje del PIB.	Eficiencia del Gobierno	1999
96	Tasa de Contribución a la Seguridad Social de la Empresa. Contribución obligatoria como % del PIB per capita.	Eficiencia del Gobierno	2000
97	Contribución a la Seguridad Social de la empresa recaudada.	Eficiencia del Gobierno	1999
98	Sistema fiscal. El sistema fiscal desanima la inversión y las actividades empresariales.	Eficiencia del Gobierno	2001
99	Impuestos a la propiedad y al capital recaudados. Porcentaje del PIB.	Eficiencia del Gobierno	1999

Anexo II. Relación de variables utilizadas en el estudio

Id	Nombre de variable	Factor	Fecha
100	Beneficios de los impuestos indirectos recaudados. Impuestos en bienes y servicios, como porcentaje del PIB.	Eficiencia del Gobierno	1999
101	Fraude fiscal. El fraude fiscal es muy bajo.	Eficiencia del Gobierno	2001
102	Tipo de interés a corto plazo. Descuento real promedio/tasa bancaria, Enero - Diciembre	Eficiencia del Gobierno	2000
103	Coste del Capital. El coste del capital no dificulta el desarrollo empresarial.	Eficiencia del Gobierno	2001
104	Margen de intermediación. Tasa de préstamo menos tasa de depósito.	Eficiencia del Gobierno	2000
105	Grating (valor) del crédito. Grating sobre una escala de 0 a 100 valorada por el ranking del Institutional Investor Magazine.	Eficiencia del Gobierno	2000
106	Política del Banco Central. La política del banco central tiene un impacto positivo en el desarrollo económico de su país.	Eficiencia del Gobierno	2001
107	Tipo de cambio. El tipo de cambio favorece la competitividad de las empresas locales.	Eficiencia del Gobierno	2001
108	Estabilidad del tipo de cambio. Paridad de cambio de la moneda nacional a SDR, 2000/1998.	Eficiencia del Gobierno	2000
109	Consenso sobre la dirección política. El consenso entre los partidos políticos que forman la coalición gobernante es muy alto.	Eficiencia del Gobierno	2001
110	Actividad legislativa del Parlamento. La actividad legislativa del Parlamento satisface las necesidades competitivas de la economía.	Eficiencia del Gobierno	2001
111	Políticas económicas del gobierno. El gobierno adapta sus políticas a los cambios del entorno económico.	Eficiencia del Gobierno	2001
112	Decisiones Gubernamentales. Las decisiones del Gobierno son ejecutadas de forma efectiva.	Eficiencia del Gobierno	2001
113	Sistema Político. El gobierno adapta sus políticas a los cambios en el entorno económico.	Eficiencia del Gobierno	2001

Anexo II. Relación de variables utilizadas en el estudio

Id	Nombre de variable	Factor	Fecha
114	Transparencia. Las intenciones políticas del Gobierno se comunican claramente.	Eficiencia del Gobierno	2001
115	Funcionarios públicos. Los funcionarios públicos actúan al margen de las presiones políticas.	Eficiencia del Gobierno	2001
116	Burocracia. Las regulaciones administrativas (burocracia) son mínimas, no dificultan el desarrollo empresarial.	Eficiencia del Gobierno	2001
117	Soborno y corrupción. El soborno y la corrupción en la esfera pública se dan con poca frecuencia.	Eficiencia del Gobierno	2001
118	Administración aduanera. La administración aduanera no dificulta el tránsito fluido de bienes.	Eficiencia del Gobierno	2001
119	Justicia. La justicia se administra de forma adecuada.	Eficiencia del Gobierno	2001
120	Seguridad personal y propiedad privada. La eficacia del sistema de seguridad es muy alta.	Eficiencia del Gobierno	2001
121	Riesgo de inestabilidad política. El riesgo de inestabilidad política en los próximos años es muy bajo.	Eficiencia del Gobierno	2001
122	Cohesión Social. La cohesión social es una prioridad para el gobierno.	Eficiencia del Gobierno	2001
123	Crímenes importantes. Número de asesinatos, crímenes violentos o robos armados por 100.000 habitantes.	Eficiencia del Gobierno	1998
124	Participación en bloques comerciales regionales. La participación de su país en bloques comerciales regionales ofrece acceso suficiente a los mercados extranjeros.	Eficiencia del Gobierno	2001
125	Proteccionismo. Las leyes y medidas proteccionistas no impiden que los productos y servicios extranjeros sean importados.	Eficiencia del Gobierno	2001
126	Contratos del sector público. Los contratos del sector público son suficientemente abiertos a postores extranjeros.	Eficiencia del Gobierno	2001
127	Créditos y seguros a la exportación.	Eficiencia del	2001

Anexo II. Relación de variables utilizadas en el estudio

Id	Nombre de variable	Factor	Fecha
	Cuando las empresas locales deciden exportar, el apoyo activo que reciben de las instituciones públicas es muy alto.	Gobierno	
128	Economía sumergida. La economía sumergida no impide el desarrollo económico en su país.	Eficiencia del Gobierno	2001
129	Subvenciones gubernamentales. A las empresas públicas y privadas, como porcentaje del PIB.	Eficiencia del Gobierno	1998
130	Marco legal. El sistema legal no va en detrimento de la competitividad de su país	Eficiencia del Gobierno	2001
131	Leyes a la competencia. Las leyes a la competencia impiden la competencia desleal en su país.	Eficiencia del Gobierno	2001
132	Obligaciones sobre productos y servicios. La legislación sobre las obligaciones a cumplir en productos y servicios no re	Eficiencia del Gobierno	2001
133	Controles de precios. Los controles de precios del gobierno no afectan al precio de los productos en la mayoría de las industrias.	Eficiencia del Gobierno	2001
134	Regulaciones laborales. Las regulaciones laborales (prácticas de contratación y despido, salarios mínimos,...) son suficientes.	Eficiencia del Gobierno	2001
135	Legislación sobre desempleo. La legislación sobre desempleo ofrece un incentivo para buscar trabajo.	Eficiencia del Gobierno	2001
136	Leyes a la inmigración. Las leyes a la inmigración no impiden a su empresa de emplear mano de obra extranjera.	Eficiencia del Gobierno	2001
137	Regulaciones legales de las instituciones financieras. Las regulaciones legales de las instituciones financieras son adecuadas para la estabilidad financiera.	Eficiencia del Gobierno	2001
138	Derechos y responsabilidades de los accionistas. Los derechos y responsabilidades de los accionistas están bien definidos.	Eficiencia del Gobierno	2001
139	Confidencialidad de las transacciones	Eficiencia del	2001

Anexo II. Relación de variables utilizadas en el estudio

Id	Nombre de variable	Factor	Fecha
	financieras. La confidencialidad de las transacciones financieras está garantizada.	Gobierno	
140	Riesgos en el extranjero. Los riesgos a asumir de inversiones en el extranjero pueden ser negociados con los socios externos.	Eficiencia del Gobierno	2001
141	Inversores extranjeros. Los inversores extranjeros son libres de adquirir el control de compañías domésticas.	Eficiencia del Gobierno	2001
142	Empresas foráneas y locales. Las empresas foráneas y locales son tratadas de igual forma.	Eficiencia del Gobierno	2001
143	Instituciones Financieras Extranjeras. Las instituciones financieras extranjeras tienen acceso a los mercados locales.	Eficiencia del Gobierno	2001
144	Acceso a los mercados de capital locales. El acceso a los mercados de capitales locales no está restringido para las empresas extranjeras.	Eficiencia del Gobierno	2001
145	Acceso a los mercados de capitales extranjeros. El acceso a los mercados de capitales extranjeros no está restringido para las empresas locales.	Eficiencia del Gobierno	2001
146	Incentivos a la inversión. Los incentivos a la inversión son atractivos para los inversores extranjeros.	Eficiencia del Gobierno	2001
147	Proyectos de protección a la inversión. Los proyectos de protección a la inversión (en contra del nacionalismo, o de la expropiación) están disponibles para los países extranjeros.	Eficiencia del Gobierno	2001
148	Gasto público total en educación. Porcentaje del PIB.	Eficiencia del Gobierno	1999
149	Ratio alumno-profesor (Educación Primaria). Número de alumnos por profesor.	Eficiencia del Gobierno	1998
150	Ratio alumno-profesor (Educación Secundaria). Número de alumnos por profesor.	Eficiencia del Gobierno	1998
151	Matriculación en escuela secundaria.	Eficiencia del	1997

Anexo II. Relación de variables utilizadas en el estudio

Id	Nombre de variable	Factor	Fecha
	Porcentaje de grupo de edad relevante recibiendo educación a tiempo completo.	Gobierno	
152	Logros en educación superior. Porcentaje de la población que ha alcanzado al menos el nivel de educación superior.	Eficiencia del Gobierno	1998
153	Logros educativos: Matemáticas. TIMSS logro medio de alumnos, en nivel octavo.	Eficiencia del Gobierno	1998
154	Logros educativos: Ciencias. TIMSS logro medio de alumnos, en nivel octavo.	Eficiencia del Gobierno	1998
155	Sistema Educativo. El sistema educativo satisface las necesidades de una economía competitiva.	Eficiencia del Gobierno	2001
156	Educación universitaria. La educación universitaria satisface las necesidades de una economía competitiva.	Eficiencia del Gobierno	2001
157	Analfabetismo. Tasa de analfabetismo de adulto (mayor de 15 años) como porcentaje de la población.	Eficiencia del Gobierno	1998
158	Cultura económica. La cultura económica en la población local es generalmente muy alta.	Eficiencia del Gobierno	2001
159	Educación en finanzas. La formación en finanzas es en general suficiente.	Eficiencia del Gobierno	2001
160	Ingenieros cualificados. Los ingenieros cualificados necesarios para un desarrollo económico competitivo están disponibles.	Eficiencia del Gobierno	2001
161	Cooperación Universidad-Empresa. La transferencia tecnológica entre las universidades y las empresas es muy alta.	Eficiencia del Gobierno	2001
162	Productividad general (PPP). Estimaciones PIB(PPP) por persona empleada, en \$USA.	Eficiencia Empresarial	2000
163	Productividad general. PIB por persona empleada, \$USA	Eficiencia Empresarial	2000
164	Productividad general - Crecimiento real. Estimaciones: Cambio porcentual del PIB real por persona empleada.	Eficiencia Empresarial	2000
165	Productividad laboral (PPP). Estimaciones PIB (PPP) por persona empleada por hora en \$USA.	Eficiencia Empresarial	2000

Anexo II. Relación de variables utilizadas en el estudio

Id	Nombre de variable	Factor	Fecha
166	Productividad laboral. PIB por persona empleada por hora, \$USA	Eficiencia Empresarial	2000
167	Productividad agrícola (PPP). Estimaciones: PIB relacionado (PPP) por persona empleada en agricultura, \$USA	Eficiencia Empresarial	2000
168	Productividad agrícola. PIB relacionado (PPP) por persona empleada en agricultura, \$USA.	Eficiencia Empresarial	2000
169	Productividad en la industria (PPP). Estimaciones Relacionadas PIB (PPP) por persona empleada en industria.	Eficiencia Empresarial	2000
170	Productividad en la industria. PIB relacionado por persona empleada en la industria, \$USA	Eficiencia Empresarial	2000
171	Productividad en el sector servicios (PPP). Estimaciones Relacionadas PIB (PPP) por persona empleada en servicios.	Eficiencia Empresarial	2000
172	Productividad en el sector servicios. PIB relacionado por persona empleada en servicios, \$USA.	Eficiencia Empresarial	2000
173	Niveles de compensación. Estimaciones Compensación horaria total para los trabajadores en manufactura (salario más beneficios)	Eficiencia Empresarial	2000
174	Costes unitarios laborales en el sector industrial. Cambio porcentual.	Eficiencia Empresarial	2000
175	Remuneración en las profesiones del sector servicios: Administrativo bancario. Renta anual bruta incluyendo suplementos.	Eficiencia Empresarial	2000
176	Remuneración en las profesiones del sector servicios: Director de Departamento. Renta anual bruta incluyendo suplementos.	Eficiencia Empresarial	2000
177	Remuneración en las profesiones del sector servicios: Profesor de enseñanza primaria. Renta anual bruta incluyendo suplementos.	Eficiencia Empresarial	2000
178	Remuneración en las profesiones del sector servicios: Secretario. Renta anual bruta incluyendo suplementos tales como bonos, \$USA	Eficiencia Empresarial	2000
179	Remuneración de dirección: Director	Eficiencia	2000

Anexo II. Relación de variables utilizadas en el estudio

Id	Nombre de variable	Factor	Fecha
	Económico. Salario base total más bonus e incentivos a largo plazo, \$USA	Empresarial	
180	Remuneración de dirección: Ingeniero. Salario base total más bonus e incentivos a largo plazo, \$USA	Eficiencia Empresarial	2000
181	Remuneración de dirección: Director Industrial. Salario base total más bonus e incentivos a largo plazo, \$USA	Eficiencia Empresarial	2000
182	Remuneración de dirección: Director de Recursos Humanos. Salario base total más bonus e incentivos a largo plazo, \$USA	Eficiencia Empresarial	2000
183	Horas laborales. Media de número de horas laborales por año.	Eficiencia Empresarial	2000
184	Relaciones laborales. Las relaciones laborales son generalmente de colaboración.	Eficiencia Empresarial	2001
185	Motivación de los trabajadores. El grado de identificación de los empleados con los objetivos de las empresas es muy alto.	Eficiencia Empresarial	2001
186	Disputas industriales. Días laborales perdidos por 1000 habitantes por año.	Eficiencia Empresarial	1999
187	Formación de los empleados. Para las empresas locales, la formación de los empleados es una alta prioridad.	Eficiencia Empresarial	2001
188	Fuerza laboral. Empleado y desempleado registrado.	Eficiencia Empresarial	2000
189	Fuerza laboral. Porcentaje de la población	Eficiencia Empresarial	2000
190	Fuerza laboral - Crecimiento. Cambio porcentual.	Eficiencia Empresarial	2000
191	Fuerza laboral femenina. Porcentaje del total de la fuerza laboral.	Eficiencia Empresarial	2000
192	Trabajadores cualificados. Los trabajadores cualificados (es decir, relevantes para la economía) están disponibles en el mercado laboral local.	Eficiencia Empresarial	2001
193	Disponibilidad de personas cualificadas en finanzas. Los buenos profesionales especializados en finanzas necesarios para un desarrollo económico competitivo están	Eficiencia Empresarial	2001

Anexo II. Relación de variables utilizadas en el estudio

Id	Nombre de variable	Factor	Fecha
	disponibles en el mercado laboral local.		
194	Fuga de cerebros. Las personas bien formadas por lo general no emigran fuera del país.	Eficiencia Empresarial	2001
195	Experiencia internacional. La experiencia de los directivos locales en la gestión de negocios internacionales y puestos.	Eficiencia Empresarial	2001
196	Disponibilidad de altos directivos. Los altos directivos competentes necesarios para un desarrollo económico están disponibles en el mercado laboral local.	Eficiencia Empresarial	2001
197	Nivel de competencia. Los directivos locales en relación a los foráneos tienen el mismo nivel de competencia.	Eficiencia Empresarial	2001
198	Recursos del sector bancario. Porcentaje del PIB	Eficiencia Empresarial	2000
199	Créditos. En el último año, los créditos han llegado a ser más fáciles de obtener.	Eficiencia Empresarial	2001
200	Número de tarjetas de crédito emitidas. Per cápita	Eficiencia Empresarial	1999
201	Transacciones de las tarjetas de crédito. \$USA per cápita	Eficiencia Empresarial	1999
202	Capital-riesgo. La financiación mediante capital riesgo está muy desarrollada.	Eficiencia Empresarial	2001
203	La bolsa. La bolsa (incluyendo los mercados secundarios) ofrece financiación adecuada a las empresas.	Eficiencia Empresarial	2001
204	Capitalización de la bolsa. Billones \$USA	Eficiencia Empresarial	1999
205	Valor comercializado en las bolsas. \$USA per cápita	Eficiencia Empresarial	1999
206	Empresas locales en cotización. Número de empresas locales con cotización .	Eficiencia Empresarial	1999
207	Índice de la bolsa. Cambio porcentual sobre el índice.	Eficiencia Empresarial	2000
208	Transparencia de las instituciones financieras. Las instituciones financieras proporcionan información adecuada sobre sus actividades.	Eficiencia Empresarial	2001

Anexo II. Relación de variables utilizadas en el estudio

Id	Nombre de variable	Factor	Fecha
209	Uso fraudulento de información comercial. El uso fraudulento de información comercial no es común en la bolsa	Eficiencia Empresarial	2001
210	Capacidad de auto-financiación. empresas generan suficiente dinero para financiar sus inversiones.	Las Eficiencia Empresarial	2001
211	Factoring. Porcentaje de exportaciones comercializadas.	Eficiencia Empresarial	2000
212	Adaptabilidad. Las empresas se adaptan rápidamente a los cambios en el ciclo económico.	Eficiencia Empresarial	2001
213	Prácticas éticas. En la gestión empresarial de las empresas locales las prácticas y valores éticos son tomados en consideración.	Eficiencia Empresarial	2001
214	Credibilidad Corporativa. empresarios/directivos de las empresas gozan de la confianza de la sociedad.	Los Eficiencia Empresarial	2001
215	Consejos de Administración. profesionalidad de los consejos de administración de las empresas locales es muy alta.	La Eficiencia Empresarial	2001
216	Valor del accionista. El patrimonio de las empresas (entendido como el valor del accionista) se gestiona de forma eficiente.	Eficiencia Empresarial	2001
217	Satisfacción del consumidor. importancia que conceden las empresas locales a la satisfacción del cliente es muy alta.	La Eficiencia Empresarial	2001
218	Espíritu emprendedor. Generalmente los directivos tienen un espíritu empresarial innovador.	los Eficiencia Empresarial	2001
219	Creación de empresas. La situación económica actual favorece el desarrollo de nuevos negocios.	Eficiencia Empresarial	2001
220	Cultura de marketing. En las empresas locales, una cultura de marketing dinámica sí existe.	Eficiencia Empresarial	2001
221	Responsabilidad social. Los líderes empresariales no descuidan su	Eficiencia Empresarial	2001

Anexo II. Relación de variables utilizadas en el estudio

Id	Nombre de variable	Factor	Fecha
	responsabilidad hacia la sociedad.		
222	Salud, seguridad y medio ambiente. En las empresas locales, la salud laboral, la seguridad en el trabajo y el respeto al medio ambiente se gestionan adecuadamente.	Eficiencia Empresarial	2001
223	Globalización. Para las empresas locales la globalización es considerada una oportunidad importante.	Eficiencia Empresarial	2001
224	Actitudes hacia la globalización. Las actitudes hacia la globalización son generalmente positivas en su país	Eficiencia Empresarial	2001
225	Relocalización de la producción. El traslado de las actividades productivas fuera del país no supone una amenaza para su futuro económico.	Eficiencia Empresarial	2001
226	Relocalización de los centros de I+D. La relocalización de los centros de I+D no supone una amenaza para el futuro de su economía.	Eficiencia Empresarial	2001
227	Imagen de su país en el extranjero. La imagen de su país en el extranjero es buena: apoya el desarrollo de los negocios.	Eficiencia Empresarial	2001
228	Población - Tamaño del mercado. Estimaciones en millones	Infraestructura	2000
229	Población menor de 15 años. Porcentaje del total de la población.	Infraestructura	1999
230	Población mayor de 65 años. Porcentaje del total de la población.	Infraestructura	1999
231	Ratio de dependencia. Población menor de 15 y mayor de 64 años, dividido por la población activa (15 a 64 años).	Infraestructura	1999
232	Desarrollo y mantenimiento de las infraestructuras. El desarrollo y mantenimiento de las infraestructuras es adecuada.	Infraestructura	2001
233	Carreteras. Densidad de la red.	Infraestructura	1998
234	Infraestructura de distribución. La infraestructura de distribución de bienes y servicios es generalmente eficiente.	Infraestructura	2001

Anexo II. Relación de variables utilizadas en el estudio

Id	Nombre de variable	Factor	Fecha
235	Ferrocarriles. Densidad de la red.	Infraestructura	1999
236	Transporte aéreo. Número de pasajeros transportados por las principales compañías.	Infraestructura	1999
237	Transporte marítimo. La infraestructura de transporte marítimo satisface las necesidades empresariales.	Infraestructura	2001
238	Área cultivable. Metros cuadrados per capita	Infraestructura	1998
239	Urbanización. Las ciudades no agotan los recursos nacionales.	Infraestructura	2001
240	Intensidad energética. Cantidad de energía comercial consumida por cada dólar de PIB en kilojoules	Infraestructura	1996
241	PIB y consumo de energía. Crecimiento real del PIB menos crecimiento del consumo de energía.	Infraestructura	1997
242	Producción de energía indígena total. Porcentaje de los requerimientos totales en toneladas de aceite equivalente.	Infraestructura	1998
243	Importaciones de energía versus exportaciones de mercancías. Importaciones de energía como porcentaje de las exportaciones	Infraestructura	1998
244	Auto-suficiencia en materias primas no energéticas. Balanza comercial per capita en \$USA	Infraestructura	1998
245	Costes de electricidad para los clientes industriales. \$USA per kwh	Infraestructura	2000
246	Inversión en telecomunicaciones. Porcentaje del PIB	Infraestructura	1998
247	Líneas telefónicas. Número de líneas por 1000 habitantes.	Infraestructura	2000
248	Subscriptores de teléfonos móviles. Número de subscriptores por 1000 habitantes	Infraestructura	2000
249	Costes de una llamada internacional. \$USA por 3 minutos en horas punta a USA (para USA a Europa)	Infraestructura	2001

Anexo II. Relación de variables utilizadas en el estudio

Id	Nombre de variable	Factor	Fecha
250	Ordenadores en uso. Cuota de ordenadores mundiales en uso / Fuente: Computer Industry Almanac	Infraestructura	2000
251	Ordenadores per cápita. Número de ordenadores por 1000 personas / Fuente: Computer Industry Almanac	Infraestructura	2000
252	Potencia de los ordenadores. Cuota mundial de MIPS (millones de instrucciones por segundo) / Fuente: Computer Industry Amanac	Infraestructura	1998
253	Potencia de los ordenadores per capita. MIPS (millones de instrucciones por segundo) por 1000 personas / Fuente: Compute	Infraestructura	1998
254	Usuarios de Internet. Número de usuarios de Internet por 1000 personas / Fuente: Computer Industry Almanac	Infraestructura	2000
255	Nueva tecnología de información. Las nuevas tecnologías de información satisfacen las necesidades empresariales.	Infraestructura	2001
256	Comercio electrónico. El comercio electrónico está muy desarrollado.	Infraestructura	2001
257	Personas cualificadas en tecnología de información. Las personas cualificadas en tecnología de información están disponibles.	Infraestructura	2001
258	Cooperación tecnológica. La cooperación tecnológica entre las empresas es muy alta.	Infraestructura	2001
259	Gasto total en I+D. Millones \$USA.	Infraestructura	1999
260	Gasto total en I+D per capita. \$USA per capita.	Infraestructura	1999
261	Gasto total en I+D. Porcentaje del PIB	Infraestructura	1999
262	Gasto empresarial en I+D. Millones \$USA.	Infraestructura	1999
263	Gasto empresarial en I+D per capita. \$USA per capita a tasas de cambio y precios actuales.	Infraestructura	1999
264	Trabajadores en I+D totales en la nación. Trabajo equivalente a tiempo completo (FTE).	Infraestructura	1999

Anexo II. Relación de variables utilizadas en el estudio

Id	Nombre de variable	Factor	Fecha
265	Trabajadores en I+D totales en la nación per capita. Trabajo equivalente a tiempo completo (FTE) por 1000 personas.	Infraestructura	1999
266	Trabajadores en I+D totales en el mundo empresarial. Trabajo equivalente a tiempo completo.	Infraestructura	1999
267	Trabajadores en I+D totales en el mundo empresarial per capita. Trabajo equivalente a tiempo completo (FTE) por 1000 per capita.	Infraestructura	1999
268	Investigación básica. La investigación básica que se realiza para garantizar el desarrollo económico y tecnológico es su	Infraestructura	2001
269	Desarrollo y aplicación de tecnología. El desarrollo y la aplicación de la tecnología está apoyado por el entorno legal.	Infraestructura	2001
270	Fondos para desarrollo tecnológico. La escasez de suficientes recursos financieros no dificulta el desarrollo tecnológico	Infraestructura	2001
271	Premios Nóbeles. Número otorgados en física, química, fisiología o medicina y economía desde 1950.	Infraestructura	2000
272	Premios Nóbeles per capita. Número otorgados en física, química, fisiología o medicina y economía desde 1950 por millón	Infraestructura	2000
273	Ciencia y Educación. La ciencia y la educación se enseñan de forma adecuada en los colegios.	Infraestructura	2001
274	Ciencia, tecnología y juventud. La ciencia y la tecnología son temas que interesan mucho a la juventud	Infraestructura	2001
275	Patentes otorgadas a residentes. Número anual medio de patentes otorgadas a residentes.	Infraestructura	1998
276	Cambio en las patentes otorgadas a residentes. Cambio porcentual.	Infraestructura	1998
277	Obtención de patentes en el extranjero. Número de patentes aseguradas en el extranjero por residentes del país.	Infraestructura	1998

Anexo II. Relación de variables utilizadas en el estudio

Id	Nombre de variable	Factor	Fecha
278	Patente y protección de los derechos de autor. Las patentes y la protección de los derechos de autor es aplicada en su p	Infraestructura	2001
279	Número de patentes en vigor. Por 100.000 habitantes.	Infraestructura	1998
280	Gasto sanitario total. Porcentaje del PIB	Infraestructura	1998
281	Gasto público en salud. Porcentaje del gasto total sanitario.	Infraestructura	1998
282	Esperanza de vida al nacer. Estimaciones medias	Infraestructura	2000
283	Asistencia médica: Por médico. Número de habitantes por médico y enfermera	Infraestructura	1999
284	Asistencia médica: Por enfermera. Número de habitantes por médico y enfermera.	Infraestructura	1999
285	Infraestructura sanitaria. El sistema sanitario satisface las necesidades de la población.	Infraestructura	2001
286	Población urbana. Porcentaje del total de la población	Infraestructura	1999
287	Distribución de la renta - 20% más pobre. Porcentaje de rentas de familias dirigidas hacia el 20% de la población más pobre.	Infraestructura	1998
288	Distribución de la renta - 20% más rico. Porcentaje de rentas de familias dirigidas hacia el 20% de la población más rica.	Infraestructura	1998
289	Índice de desarrollo humano. Combina indicadores económicos - sociales - educativos / Fuente: Human Development Report	Infraestructura	2000
290	Abuso de alcohol y de drogas. El abuso del alcohol y de las drogas en los centros de trabajo plantean un problema poco importante.	Infraestructura	2001
291	Calidad de vida. La calidad de vida en su país es alta.	Infraestructura	2001
292	Proporción de papel y cartón reciclable. Porcentaje del consumo aparente.	Infraestructura	1997
293	Plantas de tratamiento de agua residual. Porcentaje de la población servida por	Infraestructura	1998

Anexo II. Relación de variables utilizadas en el estudio

Id	Nombre de variable	Factor	Fecha
	tratamiento de aguas.		
294	Emisiones de dióxido de carbono. CO2 emisiones industriales en toneladas métricas por un millón de \$USA de PIB	Infraestructura	1997
295	Desarrollo sostenible. El desarrollo sostenible es considerado una prioridad en su país.	Infraestructura	2001
296	Problemas de polución e infraestructuras. La infraestructura no está afectada por serios problemas de polución.	Infraestructura	2001
297	Leyes medioambientales. Las leyes existentes para proteger el medioambiente no dificultan el desarrollo de los negocios.	Infraestructura	2001
298	Cultura nacional. La cultura nacional está abierta a la influencia extranjera	Infraestructura	2001
299	Flexibilidad y adaptabilidad. La predisposición de la población activa para adaptarse a los nuevos desafíos es muy alta.	Infraestructura	2001
300	Igualdad de oportunidades. La discriminación social (raza, sexo,..) es un problema poco importante.	Infraestructura	2001
301	Acoso y violencia. El acoso y la violencia en los centros de trabajo plantean problemas poco importantes.	Infraestructura	2001
302	Valores de la sociedad. Los valores que mejoran la competitividad como el trabajo duro, la innovación) están presentes en la sociedad.	Infraestructura	2001