Universidad de Granada
Departamento de Lenguajes y Sistemas Informáticos

TESIS DOCTORAL

UNA APROXIMACIÓN AL ANÁLISIS DE LA INNOVACIÓN Y LOS PROCESOS DE GESTIÓN MEDIOAMBIENTAL:

GESTIÓN DE PATENTES BASADA EN ONTOLOGÍAS Y ESTUDIO DE WEBS CORPORATIVAS

Presentada por:

María del Campo Bermúdez Edo

Dirigida por:

Dr. Manuel Noguera García

Dra. Nuria Esther Hurtado Torres

Dr. José Luis Garrido Bullejos.

GRANADA, 2013

Editor: Editorial de la Universidad de Granada
Autor: María del Campo Bermúdez Edo
D.L.: GR 89-2014
ISBN: 978-84-9028-666-1

ÍNDICE DE CONTENIDOS

ÍNDICE DE FIGURAS

ÍNDICE DE TABLAS

PREFACIO

PREFACIO

Delimitación del Tema Objeto de Estudio

Indicadores medioambientales, económicos y sociales revelan que el modelo actual de progreso es insostenible debido a la fragilidad de los ecosistemas y la intensidad de la presión económica para explotarlos a través de sistemas poco respetuosos con el entorno (Stern, 2007). En este contexto, el papel de las empresas es clave (Fundación-Entorno, 2011).

Los distintos agentes relacionados con el funcionamiento de las empresas (*stakeholders*) vienen reclamando que éstas desarrollen innovaciones avanzadas que puedan ayudar a paliar los problemas medioambientales (p.e. Kassinis & Vafeas, 2006). Hasta el momento, las innovaciones medioambientales en productos y el desarrollo de nuevas oportunidades de comercialización, ha propiciado un interés importante entre la población y los investigadores. Es necesario destacar que las innovaciones en procesos de gestión, han recibido una menor atención. Las mismas son especialmente relevantes por su carácter transversal, reduciendo los impactos medioambientales generados como consecuencia del funcionamiento cotidiano de la empresa, sea cual sea su ámbito sectorial, tamaño u orientación comercial. La expansión del uso de las tecnologías de la información y las comunicaciones (TIC) puede contribuir a reducir las consecuencias medioambientales (Fundación-Entorno, 2011; Webb et al., 2008).

La tesis que se presenta aborda el uso de las TIC en el desarrollo de innovaciones medioambientales relacionadas con procesos de negocio en las empresas. En este sentido, la forma más eficaz de proteger una innovación es a través de patentes. Los documentos de patentes están almacenados en grandes bases de datos bajo diferentes modelos en cuanto a la estructura de la información, lo cual hace difícil el procesamiento automático y eficiente de la información que contienen. En este contexto, las ontologías, entendidas como una "*specification of a conceptualization*" (Gruber, 1993a), han demostrado ser una tecnología adecuada para capturar aspectos relevantes de un dominio y dotar de una semántica bien definida a la información, permitiendo compartir dicha información entre diferentes fuentes con distintas estructuras de datos. Las ontologías permiten la representación del conocimiento, posibilitando la identificación de contexto y dependencia de la información de una manera más sencilla que las bases de datos (Giereth et al., 2006). Establecer las relaciones y vínculos entre la información contenida en los documentos de patentes, es necesario para múltiples propósitos, tales como el desarrollo de estudios de internacionalización de patentes, la cooperación entre empresas, etc. Adicionalmente, el poder del razonamiento en las ontologías permite llevar a cabo la integración de diferentes dominios, cómo pueden ser el dominio de las patentes con dominios económicos o de gestión (Seol et al., 2011). Esta integración puede resultar de utilidad para analizar la relación entre las patentes y la gestión económica de la empresa.

Este trabajo pretende idear métodos y técnicas, así como su aplicación mediante su implementación mediante herramientas, que permitan validar la hipótesis sobre la posibilidad de mejorar la gestión medioambiental. El presente documento plantea un puente explícito entre los trabajos de investigación relacionados con la gestión de empresas (*management*) y aquellos relacionados con los aspectos teóricos y prácticos del ámbito de las TIC.

Objetivos de la Tesis

El objetivo principal de este trabajo se concreta en dos objetivos generales diferenciados a la par que interconectados. El primero estaría especialmente relacionado con aspectos de gestión y el segundo tiene un vínculo principal con las tecnologías de la información y las comunicaciones. A continuación, se describen los contenidos de cada categoría.

El primer objetivo general está globalmente relacionado con analizar aspectos específicos de la gestión de las innovaciones medioambientales en las empresas. En relación al mismo, la investigación desarrollada se centran en:

- Estudiar la utilización de las Webs corporativas para ofrecer información sobre la innovación medioambiental de la empresa y comprobar cómo afecta a la confianza de sus grupos de interés o *stakeholders*. En particular se pretende analizar las dimensiones de la confiabilidad de las empresas a partir de la información que las mismas ofrecen a sus potenciales empleados a través de sus Webs corporativas.

- Analizar cómo influye en el desempeño financiero de la empresa, la protección y explotación a nivel internacional de la innovación medioambiental. Específicamente, la tesis analiza la protección de innovaciones medioambientales, a través de patentes, en un ámbito internacional.

Alcanzados estos objetivos, identificamos vías a partir de las cuales es factible profundizar en aplicaciones tecnológicas basadas en ontologías que ayuden a gestionar y analizar información relevante para las empresas en distintos ámbitos de gestión y, particularmente, en lo que respecta a las innovaciones medioambientales. Así pues, el segundo objetivo general de la presente tesis doctoral está relacionado con idear y utilizar herramientas y procedimientos avanzados que, desde el campo de las TIC, permitan la recuperación y clasificación de la información sobre innovación en procesos de gestión medioambiental. El mismo se alcanzará dotando de semántica a la información sobre patentes, lo cual permite no sólo la recopilación de información de distintas fuentes de datos, sino también la inferencia de relaciones no explícitas en dichas fuentes. Este objetivo general se concreta en los siguientes objetivos específicos:

- Estudiar cómo la representación de la información basada en ontologías puede ser aplicada a la información contenida en las bases de datos de patentes facilitando el análisis sobre innovación medioambiental. Específicamente, se pretende estudiar cómo la representación ontológica de la información y, en particular, la representación jerárquica de los códigos tecnológicos de patentes, permite homogeneizar la información contenida en distintas bases de datos e inferir información no explícita. Este objetivo requerirá también la propuesta de una solución técnica para el poblado automático de ontologías con los datos contenidos en diferentes repositorios de información. En particular, se estudia en detalle la población de códigos jerárquicos aprovechando toda la información que dichos códigos contienen y que no se representa explícitamente en los repositorios tradicionales de patentes.

- Proponer un método concreto que generalice la creación de nuevas relaciones entre conceptos de una ontología de patentes, que permita la inferencia automática de información. En particular, se pretende estudiar cómo implementar en una ontología OWL los principales indicadores utilizados en el análisis de patentes, como por ejemplo, indicadores de la internacionalización de dichas patentes, calidad de las patentes, transferencia de conocimiento entre sector público y privado, etc.

Estructura de la Tesis

La tesis se estructura como sigue:

- El Capítulo 1 introduce y pone en contexto el primer gran objetivo general de la tesis, planteando los antecedentes y la evolución de la gestión medioambiental en las empresas en el ámbito internacional, y señalando el interés de estudiar la comunicación y protección de la innovación medioambiental.

- El Capítulo 2 introduce el segundo gran objetivo general de esta tesis, describiendo la aplicación de las tecnologías en la gestión de la información en general y de las ontologías en particular.

Los objetivos específicos de este trabajo de investigación se desarrollan entre los Capítulos 3 y 6 (ambos inclusive). Cada uno de ellos tiene una estructura similar que incluye: una revisión sobre el estado del arte, y una propuesta para alcanzar el objetivo específico de cada capítulo en forma de hipótesis, método o modelo. Todas estas propuestas serán validadas a través estudios empíricos con datos primarios o secundarios, o bien mediante casos de estudio de razonamiento sobre ontologías. Finalmente, cada capítulo proporciona las conclusiones particulares de sus respectivas propuestas. A continuación, planteamos un somero repaso de las características específicas de cada uno de los capítulos incluidos en este grupo.

- El Capítulo 3 estudia cómo las TIC, y en particular la Web corporativa, ayudan a comunicar la información medioambiental de la empresa. En particular, se estudia la confianza que genera la información sobre la gestión medioambiental de la empresa ofrecida por la Web corporativa, analizando las tres dimensiones de confiabilidad: integridad, benevolencia y capacidad de la empresa.

- El Capítulo 4 estudia cómo influye en el rendimiento de la empresa la internacionalización de la innovación medioambiental, medida a través de patentes. En particular, se analiza la diversidad geográfica de las fuentes de conocimiento de dichas innovaciones medioambientales patentadas y la amplitud geográfica su explotación.

- El Capítulo 5 plantea el desarrollo de una ontología específica que facilita el procesamiento de la información obtenida de las bases de datos de patentes. Para ello, el capítulo estudia los códigos jerárquicos tecnológicos de las patentes, define una ontología y presenta una solución técnica para poblar automáticamente la ontología propuesta con los datos procedentes de diversas bases de datos de patentes.

- El Capítulo 6 propone una metodología para crear relaciones entre conceptos de forma sistemática permitiendo inferir nueva información que no podría derivarse directamente en las bases de datos de las que procede la información. De forma más concreta, este capítulo infiere una clasificación de internacionalización de las patentes basándose en las definiciones de patrones de internacionalización de las patentes descritas por la Organización para la Cooperación y el Desarrollo Económico (OECD, 2009).

- El Capítulo 7 presenta las principales conclusiones globales de la tesis y una recopilación de las contribuciones específicas alcanzadas en los capítulos previos. Además, el capítulo recoge implicaciones de los trabajos desarrollados y propuestas para futuras investigaciones.

Referencias del Capítulo

Fundación-Entorno. (2011). *I+D+i y Cambio Climatico. Cómo aprovechar la capacidad tecnológica de la empresa española*. Fundación Entorno. Disponible en http://www.fundacionentorno.org/xtras/difusion/idi/FE_Publicacion_I+D+i.pdf.

Giereth, M., St, A., Rotard, M., & Ertl, T. (2006). Application of Semantic Technologies for Representing Patent Metadata. *1st International Workshop on Applications of Semantic Technologies AST06*, *94*, 297–305.

Gruber, T. R. (1993a). A Translational Approach to Portable Ontology Specifications. *Knowledge Acquisition*, *5*(2), 199–220.

Kassinis, G., & Vafeas, N. (2006). Stakeholder Pressures and Environmental Performance. *Academy of Management Journal*, *49*(1), 145–159.

OECD. (2009). OECD Patent Statistics Manual. Disponible en http://www.oecd.org/science/innovationinsciencetechnologyandindustry/oecdpatentstatisticsmanual.htm.

Seol, H., Lee, S., & Kim, C. (2011). Identifying New Business Areas using Patent Information: A DEA and Text Mining Approach. *Expert Systems with Applications*, *38*(4), 2933–2941.Stern, N. (2007). *The Economics of Climate Change: The Stern Review*. Cambridge University Press.

Webb M. et al., (2008). *SMART 2020 : Enabling the Low Carbon Economy in the Information Age*. The Climate Change Group. Disponible en http://www.smart2020.org/_assets/files/02_Smart2020Report.pdf

CAPÍTULO 1

GESTIÓN MEDIOAMBIENTAL DE LA EMPRESA: *STAKEHOLDERS* E INNOVACIÓN

GESTIÓN MEDIOAMBIENTAL DE LA EMPRESA: *STAKEHOLDERS* E INNOVACIÓN

1.1 Introducción

El objetivo principal de la gestión medioambiental de las empresas consiste en la integración de la protección del medioambiente en los procesos y estrategias internas de la empresa (Hart, 1995). La literatura existente en este campo destaca la importancia de diferenciar entre estrategias medioambientales proactivas y reactivas (Aragón-Correa & Sharma, 2003; Hart, 1995; Sharma & Vredenburg, 1998). Las proactivas se definen como aquellos planteamientos medioambientales de la organización que se anticipan a las regulaciones y a las tendencias sociales, desarrollando innovaciones en productos y procesos capaces de reducir el impacto medioambiental de los procesos de negocio. Las estrategias reactivas, por el contrario, se orientan básicamente a la reparación de daños medioambientales ya ocurridos y, por lo general, son consecuencia de una obligación externa (normalmente impuesta de forma legal).

Este capítulo pretende revisar los desarrollos principales relacionados con la gestión medioambiental de las empresas, en general, y profundizar en el estudio de la gestión medioambiental en un ámbito internacional. Esta temática constituye una de las bases de la tesis que se presenta y, por tanto, resulta oportuno hacer una delimitación y revisión de algunos de los conceptos en ese ámbito que servirán de punto de referencia en capítulos posteriores. El resto del capítulo se divide en cuatro partes principales que van desde los planteamientos más genéricos a los más específicamente relacionados con la investigación desarrollada. En el epígrafe siguiente, repasaremos la evolución del marco institucional relacionado con la gestión medioambiental. A continuación se hace una revisión de los principales desarrollos de investigación en el campo de la gestión medioambiental. Finalmente, se incluye un apartado amplío donde se comienza revisando la relación entre innovación medioambiental y gestión internacional, para posteriormente plantear los aspectos generales que tienen que ver con la comunicación de los logros medioambientales de la empresa y la protección de los mismos. Estos dos aspectos serán desarrollados en los siguientes capítulos.

1.2 Marco Institucional en Torno a la Gestión Medioambiental

La preocupación generalizada por el medio ambiente en el conjunto de la sociedad surgió en los años 60 y 70, con frecuencia ligada a los movimientos por los derechos civiles de aquellos años. Los primeros ámbitos de preocupación social estuvieron vinculados a la denuncia de problemas puntuales de contaminación local. La preocupación social por este tema aumenta en los 80, década en la que simultáneamente se empiezan a identificar problemas ambientales de corte global en el planeta (Morrison & Dunlap, 1986).

Esta creciente preocupación a escala mundial por los temas medioambientales, tiene como consecuencia a nivel institucional, el desarrollo en 1972 de la 1ª Conferencia de las Naciones Unidas sobre MedioAmbiente y Desarrollo, también conocida como Cumbre de la Tierra. Posteriormente, en 1987 las Naciones Unidas comienzan a hablar del concepto "*desarrollo sostenible*", es decir, conseguir un progreso que satisfaga las necesidades presentes sin comprometer las de futuras generaciones (WCED, 1987). Este concepto será el punto de referencia para orientar los planteamientos medioambientales de las décadas siguientes.

La consolidación de los temas medioambientales en el plano institucional se encuentra ligada al desarrollo de segunda edición de la Cumbre de la Tierra en 1992 en

Río de Janeiro. Esta Cumbre de Río supuso el establecimiento de la Agenda 21 para el desarrollo sostenible por parte de gobiernos, empresas y ciudadanos. La Agenda 21 estableció un plan de acción para preservar el medioambiente a nivel global que incluía áreas como biodiversidad, ciencia, pobreza, tecnología, o educación, entre otros.

El comienzo de siglo también ha traído cambios importantes en los marcos de referencia internacional que pretenden impulsar los avances medioambientales en el planeta. La Cumbre de la Tierra de Johannesburgo del año 2002, la entrada en vigor del Protocolo de Kyoto en el año 2005 y la IV Cumbre de la Tierra de Río de Janeiro de 2012, han constituido tres de los hechos más destacados de este siglo.

El Protocolo de Kyoto ha buscado un acuerdo a nivel mundial por parte de los gobiernos para la reducción de la emisión de gases a la atmósfera. Este protocolo recogía el compromiso de cada uno de los países firmantes de reducir sus emisiones contaminantes según su grado de desarrollo económico. A la finalización del periodo de referencia, el año 2012, se preveía la puesta en marcha de un nuevo mecanismo que continuara con el seguimiento de ese esfuerzo, sin embargo, hasta el momento no se ha desarrollado ningún planteamiento.

En 2012 se desarrolladó, en Río de Janeiro, la IV Cumbre de la Tierra, donde se volvió a destacar la importancia de apoyar las innovaciones medioambientales en un marco internacional, conocer los factores que las impulsan e identificar las herramientas tecnológicas que pueden facilitar su estudio y desarrollo. Además, en noviembre de 2012 en Doha se celebró la décimo octava sesión de la Conferencia de los países participantes en la UNFCCC (United Nations Framework Convention on Climate Change) y la sesión 8º de los países firmantes del Protocolo de Kyoto. No obstante, no está aún claro si los acuerdos adoptados serán una base lo suficientemente precisa que permita la continuidad de los ambiciosos planteamientos de Kyoto.

1.3 Evolución de la Literatura sobre la Gestión Medioambiental de la Empresa

La literatura de gestión de sostenibilidad medioambiental puede considerarse divulgativa y sin una base teórica fundada, hasta mediados de los 90 (Newton & Harte, 1997). La gestión medioambiental se empieza a estudiar de forma más sistemática a raíz del monográfico sobre organizaciones ecológicamente sostenibles de la revista *Academy of Management Review*, en el año 1995. Dentro de este monográfico, destacan el trabajo de Hart (1995) y el desarrollado por Jennings y Zandbergen (1995). Ambos sentarán las fundamentaciones teóricas, en las que se han apoyado gran parte de la literatura posterior. Por una parte, Hart (1995) aborda el estudio de la gestión medioambiental bajo la perspectiva de recursos y capacidades. Entre sus principales conclusiones destacan cómo las empresas que incorporen estrategias de prevención de la polución y desarrollo sostenible, podrán obtener ventajas competitivas. Por otra parte, Jennings y Zandbergen (1995) establecen las bases para aplicar planteamientos institucionales a la sostenibilidad de las organizaciones.

A partir de la publicación de este monográfico se experimenta una gran expansión de las investigaciones sobre los planteamientos medioambientales en la empresa. De este modo destacan estudios medioambientales cuyo marco teórico es la Teoría Institucional (Hoffman, 1999; Westley & Vredenburg, 1997), la Teoría de los Agentes del Entorno o *stakeholders* (Fineman & Clarke, 1996; Henriques & Sadorsky, 1996), o la Perspectiva de Recursos y Capacidades (Aragón-Correa, 1998; Sharma & Vredenburg, 1998), entre otras.

Paralelamente, desde mediados de los 90, los trabajos de investigación empiezan a analizar la gestión medioambiental como base de las ventajas competitivas de la empresa. Los pioneros en esta vertiente de pensamiento Porter y Van Der Linde (1995a) resaltaron que unos planteamientos medioambientales exigentes y con un diseño apropiado, podían mejorar la innovación de la empresa y permitir a ésta reducir los costes y obtener ventajas competitivas en países con regulaciones medioambientales menos estrictas. Posteriormente, otros trabajos de investigación se han centrado en el análisis de la repercusión de la innovación medioambiental en el resultado financiero de las empresas (Judge & Douglas, 1998; Klassen & McLaughlin, 1996; Russo & Fouts, 1997).

A finales de los 90, empiezan a aparecer los primeros trabajos que relacionan la estrategia medioambiental de las empresas y sus planteamientos internacionales. Algunas de estas investigaciones centran su atención en estudiar la influencia de las legislaciones nacionales e internacionales (Rugman & Verbeke, 1998a, 1998b). Posteriormente en este capítulo profundizaremos en los aspectos que relacionan los planteamientos de gestión medioambiental e internacionalización.

En el año 2000, la publicación de un segundo monográfico de la revista *Academy of Management Journal*, supuso otro impulso muy importante a la investigación sobre gestión medioambiental. Este monográfico relaciona la gestión medioambiental con aspectos específicos del funcionamiento estratégico de la organización, tales como el liderazgo (Andersson & Bateman, 2000; Egri & Herman, 2000), el desarrollo de mejores prácticas en dirección medioambiental internacional (Christmann, 2000), o los planteamientos y características de los directivos (Cordano & Frieze, 2000; Flannery & May, 2000; Ramus & Steger, 2000; Sharma, 2000). Por otra parte, se sigue también avanzando en el desarrollo de planteamientos institucionales que permiten entender mejor la relación del entorno con los planteamientos medioambientales de la empresa (Bansal & Roth, 2000; King & Lenox, 2000).

En esta última década son numerosos los trabajos que han destacado cómo una gestión medioambiental bien diseñada permite a las empresas mejorar su productividad (Garcés Ayerbe & Galve Górriz, 2001), o su rentabilidad financiera (Aragón-Correa et al., 2005; Bansal & Hunter, 2003; Burgos Jiménez, et al., 2001). Margolis y Walsh (2003) ofrecen una completa revisión de la literatura en la que se describen los planteamientos más tradicionales y los más actuales. Podemos señalar que la literatura ha pasado de considerar la gestión medioambiental de las empresas como un asunto de responsabilidad social corporativa, a considerarla como un ámbito de gestión estratégica y que afecta a diversas funciones de la empresa (Aragón-Correa & Sharma, 2003; Buysse & Verbeke, 2003; Sharma & Henriques, 2005).

En ese marco de referencia, actualmente, una línea importante de investigación se ha centrado en analizar la gestión medioambiental de las empresas que operan en un contexto internacional. Es dentro de esta línea donde se enmarca la presente tesis doctoral. El interés de nuestra investigación partió de examinar cómo las empresas pueden facilitar el control y acceso a la información referente a sus innovaciones medioambientales, a nivel internacional, a través de las tecnologías de la información y las comunicaciones (TIC). Adicionalmente, otro aspecto relevante de la innovación medioambiental en el ámbito internacional, está relacionado con la protección y explotación de dichas innovaciones a nivel mundial, mediante el uso de patentes. Finalmente, con esos desarrollos previos sobre la gestión medioambiental de las empresas que operan en un contexto internacional, encontramos vías a partir de las cuales era factible profundizar en aplicaciones tecnológicas basadas en ontologías que ayudarán a gestionar y analizar esos procesos. Estas tecnologías proporcionan mecanismos para representar, compartir, intercambiar, reutilizar y comunicar información y conocimiento referente a la gestión de la innovación medioambiental en la empresa. En el siguiente apartado de la tesis, plantearemos precisamente una revisión de la literatura sobre la relación entre la innovación medioambiental y la internacionalización de las empresas que supone la base de los desarrollos comentados sobre la gestión medioambiental de las empresas que operan en un contexto internacional. Inmediatamente, a continuación, abordaremos de forma específica la comunicación de esas innovaciones a través de la adopción de tecnologías de la información y las comunicaciones y la protección de las mismas a través de patentes.

1.4. Innovación en Gestión Medioambiental en el Ámbito Internacional: Comunicación y Protección

1.4.1. Innovación Medioambiental en el Ámbito Internacional.

La creciente globalización y la eliminación de las fronteras comerciales entre países han llevado a la necesidad de analizar cómo desarrollar los planteamientos de gestión de las empresas en un contexto internacional (Dowell et al., 2000). La internacionalización permitirá a las empresas aprovechar oportunidades de negocio de nuevos mercados, adquirir nuevos conocimientos y para reforzar globalmente su ventaja competitiva frente a la competencia (Sapienza et al., 2006).

En los últimos años, la relación entre la internacionalización de las empresas y su gestión medioambiental está despertando un creciente interés. Por un lado, una gestión medioambiental proactiva puede contribuir a un mayor éxito en la internacionalización de empresas (Martin-Tapia et al., 2008). Por otro lado, la internacionalización que experimentan las organizaciones puede favorecer la generación de prácticas medioambientales novedosas y avanzadas (Christmann & Taylor, 2001).

La literatura que analiza la gestión medioambiental e internacionalización ha prestado especial atención al estudio del comportamiento medioambiental de la empresa en las empresas multinacionales (Kostova, 1999). Estas empresas deben decidir si transferir sus prácticas medioambientales a lo largo de su red interna organizativa o mantener planteamientos diferentes en los distintos países (Rugman & Verbeke, 1998a, 1998b). Numerosos trabajos se centran en analizar el efecto de las legislaciones medioambientales de los distintos países en la estrategia medioambiental de dichas organizaciones multinacionales (Christmann & Taylor, 2001; Porter & Linde, 1995a; Rappaport & Flaherty, 1992). Considerando que las legislaciones medioambientales son diferentes en cada país, las empresas deben elegir qué modelo de gestión medioambiental quieren seguir, es decir, un modelo más tradicional de adaptar su gestión medioambiental a las exigencias mínimas del país de acogida (Leonard, 1988; Vernon, 1992) o, por el contrario, un modelo más actual en el que la empresa estandarice sus prácticas de gestión medioambiental (Christmann & Taylor, 2006; Darnall, 2006; Darnall et al., 2008). Este segundo modelo se fundamenta en la búsqueda, por parte de la empresa, de la ventaja competitiva a través de su reputación (Christmann, 2000), la legitimación internacional (Bansal, 2005; Kostova et al., 2008), o mediante su eficiencia interna (Sharma & Vredenburg, 1998; Shrivastava, 1995). Sin embargo, la estricta adaptación al contexto del país suele asociarse con una forma menos costosa de alcanzar los niveles medioambientales mínimos de referencia.

Es preciso destacar que, junto a las multinacionales, existen otro tipo de organizaciones que están experimentando un proceso de expansión internacional muy intenso. Se trata de empresas exportadoras que, aunque no cuentan con unidades organizativas ubicadas en países extranjeros, actúan en diferentes mercados fundamentalmente a través de sus exportaciones. El proceso de internacionalización permite a estas empresas adquirir conocimiento y aprendizaje valioso que pueden constituir una fuente de ventaja competitiva para las mismas (Hitt et al., 1997). En el ámbito medioambiental, las empresas adquieren nuevos conocimientos, derivados del desarrollo de actividades en entornos con distintos perfiles institucionales. Por tanto, la actuación de las empresas en diversas regiones con perfil medioambiental diferente, denominada diversificación internacional medioambiental, puede facilitar la adquisición de conocimiento medioambiental innovador y valioso (Bansal, 2005). Así, mientras las empresas multinacionales centran sus preocupaciones en el desarrollo de planteamientos globales o específicos para las instalaciones de cada país, las empresas exportadoras centran su interés en conocer las implicaciones de la gestión medioambiental en la relación con agentes específicos de cada mercado internacional (Aguilera-Caracuel et al., 2011).

En este mismo ámbito de investigación, otros trabajos han analizado cómo determinados recursos y capacidades de la empresa puede reforzar el vínculo existente entre los planteamientos medioambientales proactivos y sus desarrollos internacionales. Así, el hecho de que la empresa pueda generar exceso de recursos (*slack resources*) en su gestión internacional puede contribuir positivamente a que exista una inversión en prácticas medioambientales avanzadas (Bansal, 2005; Henriques & Sadorsky, 1996), pudiéndose transferir y aplicar las referidas prácticas en diferentes mercados con independencia de las peculiaridades institucionales de los diferentes entornos. Además, la orientación al aprendizaje en los mercados internacionales (Yeoh, 2004), así como, el grado de aprendizaje organizativo de las propias empresas (Cohen & Levinthal, 1990; Huber, 1991) pueden favorecer en gran medida la generación, integración y aplicación de prácticas medioambientales novedosas al resto de mercados. La introducción de planteamientos medioambientales novedosos buscará la conversión de las potenciales amenazas derivadas de las nuevas tendencias medioambientales en oportunidades susceptibles de ser aprovechadas por la organización. Por tanto, la introducción en el mercado de nuevos productos y procesos más avanzados en términos medioambientales puede ayudar a alcanzar una auténtica ventaja competitiva (Aragón-Correa, 1998).

Por tanto, tal y como afirmó Grant (1996), sostener una ventaja competitiva en entornos tan dinámicos como los mercados internacionales actuales, requiere una capacidad de innovación continua. En este sentido, las innovaciones medioambientales responden a las demandas externas, permitiendo a las empresas obtener legitimidad y asegurar la supervivencia, y, al mismo tiempo, favorecen el desarrollo de recursos y capacidades que son únicos, difícil de imitar, y de valor reconocido para mejorar sus resultados (Markman et al., 2004).

La literatura sobre capacidades medioambientales ha identificado la capacidad de innovar con aspectos tales como: la mejora continua para reducir y eliminar residuos (Hart, 1995); la innovación continua en procesos, servicios y sistemas (Sharma & Vredenburg, 1998); la innovación medioambiental para la prevención de la contaminación (Russo & Fouts, 1997); la innovación en procesos e implementación de la tecnología medioambiental (Christmann, 2000); y el desarrollo y reordenación de la cartera de tecnología medioambiental (Klassen & Whybark, 1999). En definitiva, podemos señalar que la innovación en temas medioambientales es esencial para crear las capacidades que permiten a la empresa aprovechar las oportunidades de mercado para el desarrollo sostenible (Bansal, 2005).

La innovación medioambiental en un contexto internacional plantea algunos retos relevantes para los investigadores. Uno de esos retos consiste en ser capaces de innovar en los procesos que permiten gestionar la información de los desarrollos medioambientales adoptados por las empresas y facilitar el acceso y la comunicación de los mismos a los agentes externos e internos. La posibilidad de hacerlo a través del uso de las tecnologías de la información y las comunicaciones e Internet es especialmente apropiada en el contexto global actual. Otro reto relevante está relacionado con la protección de esas innovaciones medioambientales en un marco internacional. En este sentido, las patentes son uno de los medios más efectivos para proteger innovaciones en la empresa. En los siguientes apartados introducciones a nivel teórico cada una de las temáticas planteadas.

1.4.2. Integración de las Tecnologías de la Información y las Comunicaciones en la Comunicación de la Innovación Medioambiental en un Ámbito Global

En el entorno actual competitivo, cobra una especial relevancia el uso de tecnologías de la información y las comunicaciones a través de las cuales las empresas pueden crear, guardar, gestionar, diseminar y comunicar información sobre sus desarrollos medioambientales. Las TIC pueden ayudar a las empresas a mejorar su habilidad para identificar y aprovechar oportunidades y responder a amenazas externas (Lee et al., 2008). Un adecuado uso de la TIC, en conjunción con otros recursos y capacidades de la empresa, mejorará la gestión de los procesos de negocio de una empresa, y como consecuencia el resultado financiero de la misma (Melville et al., 2004).

Son numerosos los estudios, en el área de sistemas de información, que han analizado el desarrollo, a nivel de empresa, de las TIC, junto con el análisis de cómo dichas tecnologías interactúan con los individuos, grupos, organizaciones y/o los mercados (Sidorova et al., 2008). Cada uno de estos factores corresponde a un nivel de análisis diferente. La Figura 1.1 muestra cómo las TIC están implicadas en todos los niveles de la empresa.

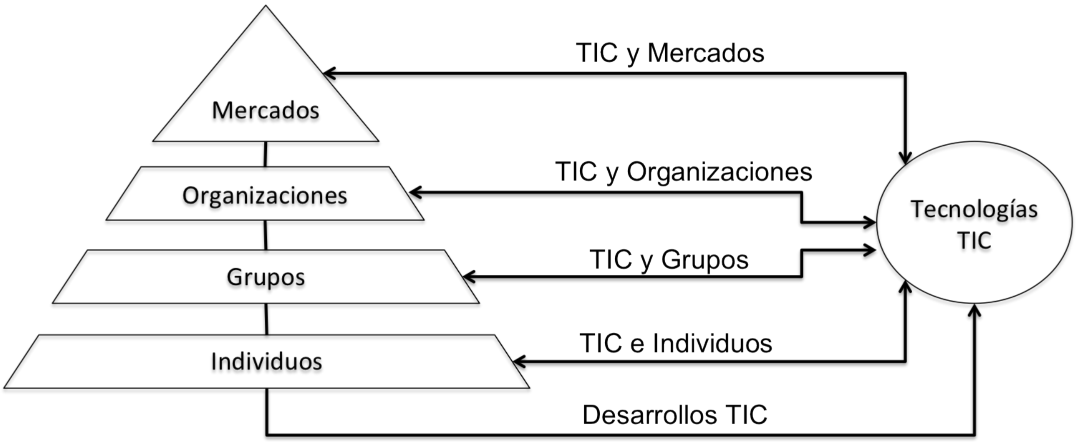

Figura 1.1 Modelo de identificación de investigación TIC (Sidorova et al., 2008)

En las empresas es interesante conocer no sólo las tecnologías TIC sino también, como dichas tecnologías afectan a los diferentes niveles de la pirámide. Benbasat y Zmud (2003) distinguen cuatro áreas de interés que se ven afectadas por la inclusión de las TIC en las empresas (ver Figuras de la 1.2 a la 1.5):

- El uso que se hace de dichas tecnologías
- El impacto que la utilización de dichas tecnologías tiene en cada nivel
- Las prácticas de gestión, apoyo metodológico y operativo del uso y la adopción de dichas tecnologías
- Las capacidades de gestión y las prácticas, metodologías y técnicas para la planificación, diseño, construcción e implementación de dichas tecnologías

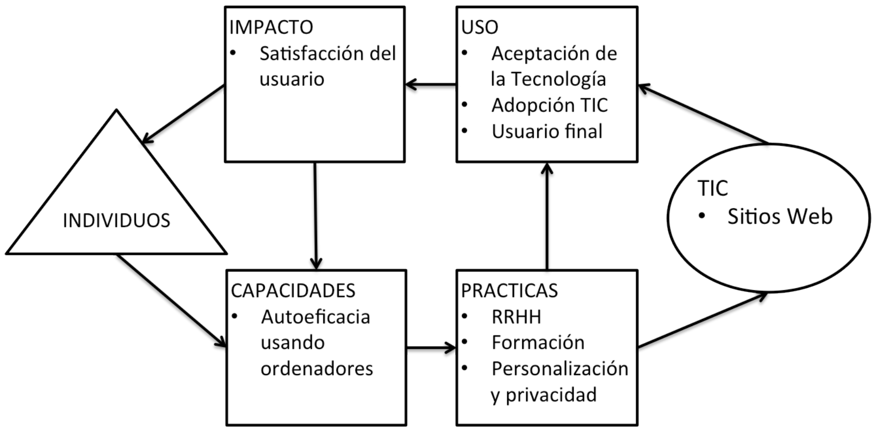

Figura 1.2 TIC e individuos (Sidorova et al., 2008)

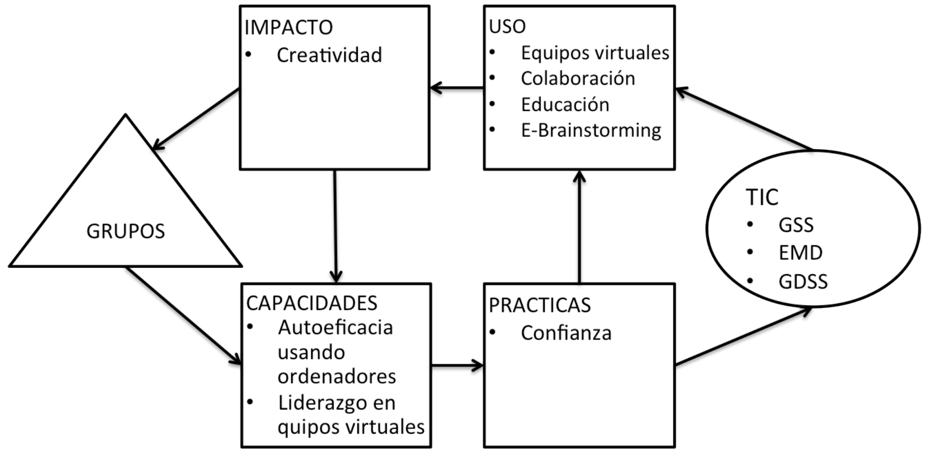

Figura 1.3 TIC y grupos (Sidorova et al., 2008)

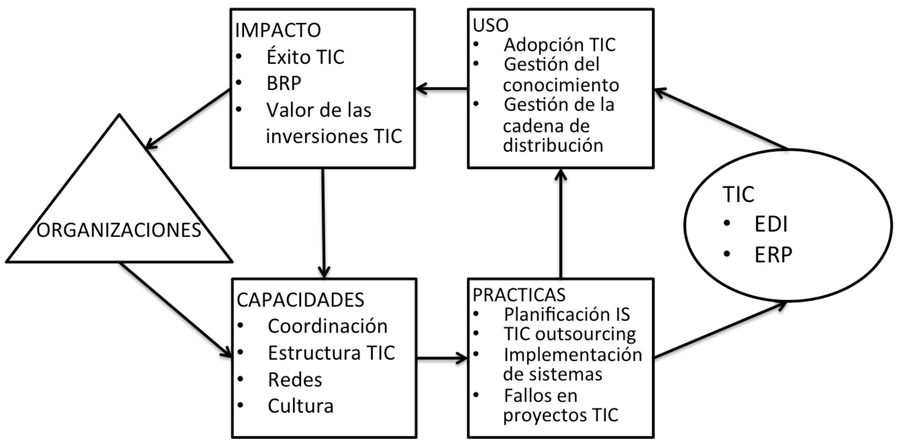

Figura 1.4 TIC y organizaciones (Sidorova et al., 2008)

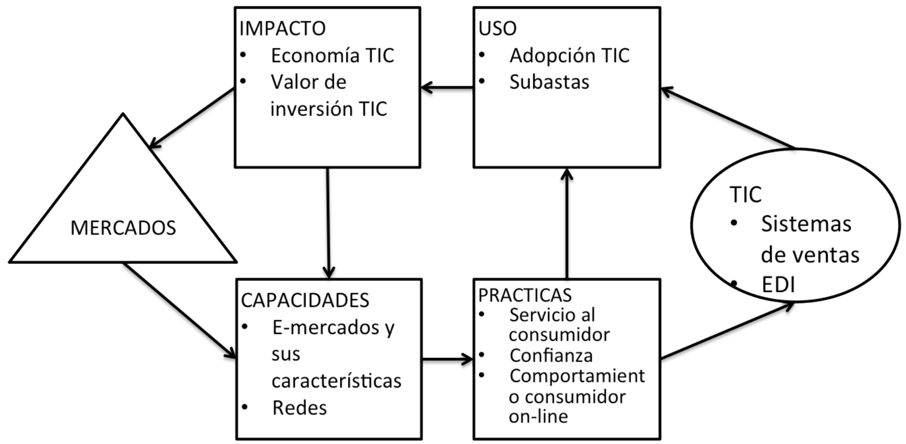

Figura 1.5 TIC y mercados (Sidorova et al., 2008)

Las investigaciones desarrolladas en el nivel de TIC y los individuos se centran en los aspectos psicológicos de las interacciones persona-ordenador (Figura 1.2). Entre los temas analizados destacan, el estudio de los modelos de aceptación de la tecnología a nivel individual y autoeficacia en el uso de ordenadores, entre otros. En el siguiente nivel, las TIC y su interacción con los grupos, la investigación ha destacado los sistemas usados para apoyar el trabajo en grupo (*groupware*) (Garrido, 2003; Garrido et al. 2003), incluyendo sistemas de apoyo a la decisión, así como, la influencia de tales sistemas en las dinámicas de grupo o en las relaciones de confianza en comunidades virtuales (Figura 1.3).

En lo que respecta a las TIC y las organizaciones, los trabajos se han centrado en el uso de TIC, analizando su papel estratégico, el impacto de la inversión en TIC en el resultado financiero y el efecto de las TIC en los procesos de negocio. Las investigaciones en este nivel también se han enfocado en las prácticas de gestión de TIC, tales como planificación y *outsourcing* (Figura 1.4). Los estudios desarrollados a nivel de TIC y los mercados versan sobre como el uso de TIC afecta a las relaciones con otras organizaciones y con otros agentes externos a la empresa. Por ejemplo, las investigaciones se centran en analizar un tipo específico de sistema de información y cómo dicho sistema impacta en potenciales consumidores, vendedores etc. (Figura 1.5).

Además de los cuatro niveles de investigación descritos con anterioridad, dentro del área de conocimiento de sistemas de información, son importantes los trabajos que versan sobre el desarrollo de dichas tecnologías de la información y las comunicaciones. Estas investigaciones, de carácter socio-técnico, se centran en el diseño y las funcionalidades de diferentes tipos de sistemas CSCW (*Computer Supported Cooperative Work*) (Garrido et al., 2003), sistemas de apoyo a la decisión, sistemas expertos, etc.

La relación de las tecnologías de la información y las comunicaciones con la innovación, en general, y con la innovación medioambiental, en particular, ha sido objeto de estudio en los distintos niveles de análisis descritos anteriormente. En este sentido, la integración de las TIC en los procesos de gestión de la innovación medioambiental pueden ser un pilar clave en el fomento y utilización de las mismas en los distintos ámbitos o procesos de la empresa (ej. garantizando un acceso fiable a la información relativa a dichas innovaciones, tanto interna como externamente). Las TIC se convierten así en un instrumento que facilita la comunicación a los agentes externos e internos de los desarrollos medioambientales adoptados por las empresas.

En el contexto actual, una de las tecnologías que más influencia tiene en la visibilidad en Internet de una organización y de sus innovaciones ha sido la "*World Wide Web*", convirtiéndose en un espacio virtual de comunicación y colaboración. Especialmente, las Web corporativas definen la presencia de la empresa en Internet, y se convierten en un instrumento para transmitir los valores fundamentales de la misma, proporcionar la imagen que la empresa desea que se conozca a nivel global y facilitar la colaboración entre agentes internos y externos a la organización.

Es en ese escenario, interactúan lo social (dónde y cómo se producen las comunicaciones y relaciones) y lo tecnológico (nuevas herramientas, sistemas, plataformas, aplicaciones y servicios) provocando cambios de lo uno sobre lo otro.

Analizar las implicaciones de los contenidos publicados por una empresa en su Web corporativa ha ido ganando en importancia en los últimos años como consecuencia de la generalización de su utilización (y, por tanto, un mayor acceso de cualquier persona a esos contenidos). La compra por Internet es seguramente el proceso que ha despertado más atención sobre el efecto generado por las Webs (Jones & Leonard, 2008; Kuan & Bock, 2007; Lee et al., 2009). Sin embargo, otros procesos, tales como, los relacionados con el reclutamiento de personal online o la colaboración e interacción a través de redes sociales propiciadas por la Web 2.0, han llevado también a subrayar el papel de la Web en otros procesos de gestión de las empresas, llegándose a acuñar el término "*Empresa 2.0*". Por ello, desarrollamos en esta tesis propuestas para analizar la Web corporativa como plataforma para la difusión de información medioambiental relevante por parte de la empresa y el uso de tecnologías propias de la Web semántica, como son las ontologías, para representar el conocimiento y la innovación medioambiental, definiendo formalmente los conceptos de los diferentes dominios y sus relaciones, con capacidad para realizar deducciones con este conocimiento.

1.4.3. La Protección de la Innovación Medioambiental a Través de Patentes

1.4.3.1. Las Patentes: Aspectos Generales

Una patente es un derecho legal que autoriza a su poseedor a excluir a terceros de actos de fabricación, uso, comercialización, oferta para la venta, venta, o importación para estos fines del producto, o proceso, objeto de la patente (TRIPS –"*Trade Related aspects of Intellectual Property Rights-Agreement*", art. 28). Este derecho es limitado en el tiempo, en la mayoría de los países es 20 años, desde el día de presentación de la patente en la oficina de patentes. La concesión de la patente está condicionada al cumplimiento de unos requisitos que son evaluados por la oficina de patentes. En este sentido, destacar que la propuesta debe ser novedosa, inventiva y de aplicación industrial (Somaya, 2012).

El objetivo de las patentes es la estimulación de actividades inventivas e innovadoras, previniendo la copia inmediata de dichas actividades. A cambio el inventor debe hacer público su invento, en lugar de mantenerlo en secreto, para permitir a terceros generar mayor conocimiento a partir de dicha patente y, de este modo, fomentar el desarrollo de nuevas innovaciones que vayan más allá de la propia patente. Dado que la inversión inicial en una patente es grande, tanto en lo relativo a los desarrollos de innovación, como en términos administrativos (Park, 2010), las patentes tratan de otorgar protección a las rentas que se pueden generar de las mismas. No obstante, la protección de las patentes es imperfecta en algunas ocasiones, por las dificultades derivadas del hecho de que la patente pueda cubrir completamente todas las opciones que impedirían la imitación de la innovación que trata de proteger (Lemley & Shapiro, 2005).

En los últimos años se ha observado un crecimiento exponencial en el número de innovaciones que son patentadas por las empresas en todo el mundo. Por ejemplo, en el año 2008 se presentaron aproximadamente 1.91 millones de patentes en todo el mundo (Gurry et al., 2012). La estrategia de las empresas y su ámbito de actividad influye en la mayor o menor propensión de las empresas a patentar (Pisano, 2006; Somaya, 2012). Las patentes representan sólo una parte de las innovaciones de la empresa, ya que no todas las innovaciones son patentadas y, además, las empresas pueden optar por otras estrategias de protección de sus innovaciones. No obstante, es importante destacar la importancia de las innovaciones patentadas. Dicha relevancia viene reforzada por al menos dos factores. Por una parte, se trata de las innovaciones que han pasado el escrutinio de una oficina de patentes en la evaluación de su novedad y su capacidad inventiva. Por otra parte, han superado el "*test del esfuerzo*" en cuanto a inversión y recursos comprometidos por la empresa, indicando una expectativa favorable, por parte de la empresa, en lo que respecta a su utilidad y aceptabilidad en el mercado. Por todo ello, las patentes han sido consideradas como indicadores adecuados para valorar la innovación de las empresas, en términos de resultados, (*outputs*) (OECD, 2009).

En los años 60, los trabajos desarrollados por Schmookler y Scherner son pioneros en la consideración de las patentes como medida de la innovación (OECD, 2009). Schmookler (1966) utilizó el número de patentes para estudiar los cambios tecnológicos en ciertos sectores y afirmó que las patentes eran una medida adecuada de la innovación, ya que facilitaban información sobre el producto final de la innovación y, también, sobre el proceso de innovación. Por su parte, Scherner (1965) fue uno de los primeros en proponer la existencia de una correlación positiva entre el número de patentes y los beneficios de la empresa.

En el año 1973 la National Science Foundation (NSF) del gobierno estadounidense publica su primer informe de indicadores del estado de la ciencia (*Science Indicators* 1972), donde presta una atención importante a los datos sobre patentes.

A finales de los años 70 la USTPO (United States Trade and Patent Office) convirtió la información sobre patentes a un formato digital, legible por máquinas, y en 1986, la EPO (European Patent Office) digitalizó todos sus archivos, con documentos de patentes que databan de 1920. La informatización de las bases de datos de patentes ayudó sobremanera a que, a lo largo de los años 80, se extendiese la línea de investigación que utilizaba las patentes como medida de la innovación (Acs & Audretsch, 1987; Bound et al., 1982; Hall et al., 1986; Scherer, 1982). Los trabajos desarrollados inciden en la necesidad de relacionar los esfuerzos de I+D de la empresa con el número de patentes. Bound et al. (1982) y Grilliches y Pakes (1984) observaron la existencia de una fuerte relación entre la inversión realizada en I+D y el número de patentes de una empresa, independientemente de su tamaño. Por su parte, Griliches (1990) llegó a afirmar que las patentes son una fuente única para analizar el cambio tecnológico ante la falta de otras evidencias que describan de forma tan detallada las actividades de I+D.

En 1994 la OECD (Organisation for Economic Co-operation and Development) inició la publicación del manual "*OECD Patent Statistic Manual*" (OECD, 1994), lo que facilitó el uso de las patentes como medida de la innovación. Dicho manual se actualiza periódicamente, siendo su última versión la de febrero del 2009 (OECD, 2009). Por su parte, Eurostat también se unió a las publicaciones periódicas de estadísticas de patentes, en 1994, con su informe "*Science and Technology Indicators*". Finalmente, cabe destacar, que la EPO comienza en 2006 a publicar las estadísticas de patentes a nivel mundial a través de su base de datos PATSTAT (*PATent STATistics*).

En este contexto de creciente disponibilidad de datos estadísticos sobre las patentes generadas y donde existe un importante consenso sobre la utilidad de las mismas, los estudios académicos apoyados en datos de patentes, como medida de la innovación, han ido creciendo exponencialmente en las últimas décadas [ver (Hanel, 2006; Somaya, 2012; Vries, 2012) para una revisión de la literatura de los últimos años]. Además, es necesario destacar que en las publicaciones, a nivel internacional, que versan sobre el sistema de ciencia y tecnología, se destaca la importancia de las patentes (European Commission, 2010a; Gurry et al., 2012; NSF-US, 2012; OECD, 2009).

1.4.3.2. Las Patentes Medioambientales

Las patentes medioambientales han sido también objeto de atención específica en trabajos previos (Brunnermeier & Cohen, 2003; Jaffe & Palmer, 1997; Lanjouw &

Mody, 1996). Frente a los trabajos pioneros sobre innovaciones medioambientales de las empresas, cuyos datos procedían mayoritariamente de la propia percepción de los responsables de las empresas, el análisis de la innovación medioambiental a través de las patentes garantiza centrar la atención en las innovaciones más radicales, es decir, que presentan la importancia suficiente como para ser patentadas.

Las oficinas de patentes, siendo conscientes de la importancia de las innovaciones medioambientales, recientemente han desarrollado categorías específicas para clasificar a las patentes de innovaciones medioambientales. Así, en 2009, la Oficina Estadounidense de Patentes ha definido la categoría "*tecnologías verdes*", comprometiéndose formalmente a dar prioridad, en su trámite, a las solicitudes de esta categoría (USPTO, 2009). Por su parte, la Oficina Europea de Patentes (EPO), en colaboración con el United Nations Environment Programme (UNEP) y el International Centre for Trade and Sustainable Development (ICTSD), puso en marcha un proyecto conjunto para analizar el papel de las patentes en el desarrollo de tecnologías que permitan la lucha contra el cambio climático (Veefkind et al., 2012). Como consecuencia del citado proyecto, a finales del 2010 la EPO definió unos nuevos códigos para clasificar "*patentes ecológicas*" (Veefkind et al., 2012). Esta categoría de la EPO denominada "*technologies or applications for mitigation or adaptation against climate change*" se empezó a utilizar a partir de octubre de 2011 y contiene, a principios del 2013, cuatro subclases y 23 grupos, divididos a su vez en subgrupos referidos a innovaciones medioambientales. Esta clasificación está en continua evolución, por lo que cada día se incorporan nuevas subclases y subgrupos.

Por tanto, podemos señalar que la creciente importancia de los temas medioambientales y el aumento en la información medioambiental contenida en las bases de datos de patentes está favoreciendo el desarrollo de trabajos de índole económica en los que se analizan las patentes medioambientales (Johnstone et al., 2012; Peters et al., 2012). Además, los cambios descritos en los sistemas de clasificación de las patentes en las principales oficinas de patentes, incluyendo códigos específicos para las patentes medioambientales, hacen previsible que la atención investigadora sobre el tema aumente todavía más su importancia en los próximos años.

No obstante, el análisis de las patentes medioambientales y la obtención de información relevante sobre las mismas que pueda ser de utilidad a los gestores y a los investigadores en el ámbito de la gestión de empresas, se enfrenta con distintos problemas. Algunos de esos problemas son similares a los que encontramos en el marco general de las patentes, pero otros son nuevos o se acentúan en el ámbito medioambiental.

En primer lugar, aunque se están realizando muchos esfuerzos para homogeneizar las leyes sobre patentes, (por ejemplo, a través de asociaciones, como la WIPO – World Intellectual Property Organization), la distinta regulación de patentes que existe en los países y las diferentes clasificaciones nacionales de patentes, aumentan la complejidad en la gestión de una patente si la empresa está interesada en su explotación a nivel internacional. A su vez, la novedad de la regulación sobre las patentes medioambientales acentúa esa problemática frente a campos más consolidados. En segundo lugar, las bases de datos en las que se almacena la información de las patentes no tienen representadas explícitamente relaciones entre sus contenidos y conceptos, lo que dificulta una fácil descripción e identificación externa de los desarrollos específicamente realizados en cada caso. Finalmente, existen varias bases de datos con información sobre patentes y cada una de ellas representa la información de forma diferente, lo cual hace difícil el procesamiento automático por ordenadores de la información contenida en ellas y, por tanto, su análisis es más complejo al realizarse manualmente.

Por tanto, considerando que actualmente las empresas están aumentando exponencialmente el volumen anual de innovaciones medioambientales que son patentadas y que la información sobre tales patentes está almacenada en distintas bases de datos con diferencias en cuanto a la representación de la información (y con limitaciones en lo que respecta a las relaciones entre términos y conceptos que utilizan en su descripción), podemos señalar que las tecnologías de la información y las comunicaciones van a jugar un papel muy relevante en posibilitar la gestión, análisis y desarrollo de un apropiado sistema de patentes, en general, y de patentes medioambientales en particular. El siguiente capítulo de la tesis revisa algunos desarrollos de carácter tecnológico que pueden ser útiles en esta orientación. Capítulos posteriores de la tesis desarrollarán estas ideas con propuestas específicas.

Referencias del Capítulo

Acs, Z. J., & Audretsch, D. B. (1987). Innovation, Market Structure, and Firm Size. *The Review of Economics and Statistics*, *69*(4), 567–574.

Aguilera-Caracuel, J., Aragón-Correa, J. A., Hurtado-Torres, N. E., & Rugman, A. M. (2011). The Effects of Institutional Distance and Headquarters' Financial Performance on the Generation of Environmental Standards in Multinational Companies. *Journal of Business Ethics*, *105*(4), 461–474.

Andersson, L. M., & Bateman, T. S. (2000). Individual Environmental Initiative: Championing Natural Environmental Issues in U.S. Business Organizations. *Academy of Management Journal*, *43*(4), 548–570.

Aragón-Correa, J. A. (1998). Strategic Proactivity and Firm Approach to the Natural Environment. *Academy of Management Journal*, *41*(5), 556–567.

Aragón Correa, J. A., Hurtado Torres, N. E., & García Morales, V. J. (2005). Un Modelo Explicativo de las Estrategias Medioambientales Avanzadas para Pequeñas y Medianas Empresas y su Influencia en los Resultados. *Cuadernos De Economía y Dirección De La Empresa*, 25, 29-52.

Aragón-Correa, J. A., & Sharma, S. (2003). A Contingent Resource-Based View of Proactive Corporate Environmental Strategy. *Academy of Management Review*, *28*(1), 71–88.

Bansal, P. (2005). Evolving Sustainably: A Longitudinal Study of Corporate Sustainable Development. *Strategic Management Journal*, *26*(3), 197–218.

Bansal, P., & Hunter, T. (2003). Strategic Explanations for the Early Adoption of ISO 14001. *Journal of Business Ethics*, *46*(3), 289–299.

Bansal, P., & Roth, K. (2000). Why Companies Go Green: A Model of Ecological Responsiveness. *Academy of Management Journal*, *43*(4), 717–736.

Benbasat, I., and Zmud, R. W. (2003). The Identity Crisis Within the IS Discipline: Defining and Communicating the Discipline's Core Properties, *MIS Quarterly, 27*(2), 183-194.

Bound, J., Cummins, C., Griliches, Z., Hall, B. H., & Jaffe, A. (1982). *Who Does R&D and Who Patents?* University of Chicago Press.

Brunnermeier, S. B., & Cohen, M. A. (2003). Determinants of Environmental Innovation in US Manufacturing Industries. *Journal of Environmental Economics and Management*, *45*(2), 278–293.

Burgos Jiménez, J. De, Céspedes Lorente, J. J., & Alvarez Gil, M. J. (2001). Un Análisis Exploratorio de las Estrategias Medioambientales y el Contexto Organizativo de los Hoteles Españoles. *Cuadernos de Economía y Dirección de la Empresa*, *8*, 5–32.

Buysse, K., & Verbeke, A. (2003). Proactive Environmental Strategies: A Stakeholder Management Perspective. *Strategic Management Journal*, *24*(5), 453–470.

Christmann, P. (2000). Effects of "Best Practices" of Environmental Management on Cost Advantage: The Role of Complementary Assets. *Academy of Management Journal*, *43*(4), 663–680.

Christmann, P., & Taylor, G. (2001). Globalization and the Environment: Determinants of Firm Self-Regulation in China. *Journal of International Business Studies*, *32*(3), 439–458.

Christmann, P., & Taylor, G. (2006). Firm Self-regulation through International Certifiable Standards: Determinants of Symbolic Versus Substantive Implementation. *Journal of International Business Studies*, *37*(6), 863–878.

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly*, *35*(1), 128–152.

Cordano, M., & Frieze, I. H. (2000). Pollution Reduction Preferences of U.S. Environmental Managers: Applying Ajzen's Theory of Planned Behavior. *Academy of Management Journal*, *43*(4), 627–641.

Darnall, N. (2006). Why Firms Mandate ISO 14001 Certification. *Business Society*, *45*(3), 354–381.

Darnall, N., Henriques, I., & Sadorsky, P. (2008). Do Environmental Management Systems Improve Business Performance in an International Setting? *Journal of International Management*, *14*(4), 364–376.

Dowell, G., Hart, S., & Yeung, B. (2000). Do Corporate Global Environmental Standards Create or Destroy Market Value? *Management Science*, *46*(8), 1059–1074.

Egri, C. P., & Herman, S. (2000). Leadership in the North American Environmental Sector: Values, Leadership Styles, and Contexts of Environmental Leaders and Their Organizations. *Academy of Management Journal*, *43*(4), 571–604.

European Commission. (2010a). *Europe's Digital Competitiveness Report*. Disponible en http://ec.europa.eu/ceskarepublika/pdf/press/ks7ec2.pdf

Fineman, S., & Clarke, K. (1996). Green Stakeholders: Industry Interpretations and Response. *Journal of Management Studies*, *33*(6), 715–730.

Flannery, B. L., & May, D. R. (2000). Environmental Ethical Decision Making in the U.S. Metal-finishing Industry. *Academy of Management Journal*, *43*(4), 642–662.

Garcés Ayerbe, C., & Galve Górriz, C. (2001). Repercusión de las Inversiones en Protección del Medio Ambiente en la Productividad de las Empresas Españolas: Un Análisis Empírico. *Cuadernos de Economía y Dirección de la Empresa*, *8*, 33–50.

Garrido, J. L. (2003). *AMENITIES: Una Metodología para el Desarrollo de Sistemas Cooperativos Basada en Modelos de Comportamiento y Tareas*. Tesis Doctoral. Universidad de Granada.

Garrido, J.L., Gea, M., Rodríguez, M.L. (2003). Requirements Engineering in Cooperative Systems.. *Requirements Engineering for Sociotechnical Systems*. 226-244.

Grant, R. M. (1996). Prospering in Dynamically-Competitive Environments: Organizational Capability as Knowledge Integration. *Organization Science*, *7*(4), 375–387.

Griliches, Z. (1990). Patent Statistics as Economic Indicators: A Survey. *Journal of Economic Literature*, *28*(4), 1661–1707.

Gurry, F., Fink, C., Khan, M., Bergquist, K., Lamb, R., Feuvre, B. Le, & Hao Zhou. (2012). *World Intellectual Property Indicators 2011. WIPO Economics & Statistics Series*. Disponible en http://www.wipo.int/ipstats/en/wipi/

Hall, B. H., Griliches, Z., & Hausman, J. A. (1986). Patents and R&D: Is there a lag? *International Economic Review*, *27*(2), 265–283.

Hanel, P. (2006). Intellectual Property Rights Business Management Practices: A Survey of the Literature. *Technovation*, *26*(8), 895–931.

Hart, S. L. (1995). A Natural-Resource-Based View of the Firm. *Academy of Management Review*, *20*(4), 986–1014.

Henriques, I., & Sadorsky, P. (1996). The Determinants of an Environmentally Responsive Firm: An Empirical Approach. *Journal of Environmental Economics and Management*, *30*(3), 381–395.

Hitt, M. A., Hoskisson, R. E., & Kim, H. (1997). International Diversification: Effects on Innovation and Firm Performance in Product-Diversified Firms. *Academy of Management Journal*, *40*(4), 767–798.

Hoffman, A. J. (1999). Institutional Evolution and Change: Environmentalism and the U.S. Chemical Industry. *Academy of Management Journal*, *42*(4), 351–371.

Huber, G. P. (1991). Organizational Learning: The Contributing Processes and the Literatures. *Organization Science*, *2*(1), 88–115.

Jaffe, A. B., & Palmer, K. (1997). Environmental Regulation and Innovation: A Panel Data Study. *The Review of Economics and Statistics*, *79*(4), 610–619.

Jennings, P. D., & Zandbergen, P. A. (1995). Ecologically Sustainable Organizations: An Institutional Approach. *Academy of Management Review*, *20*(4), 1015–1052.

Johnstone, N., Haščič, I., Poirier, J., Hemar, M., & Michel, C. (2012). Environmental Policy Stringency and Technological Innovation: Evidence from Survey Data and Patent Counts. *Applied Economics*, *44*(17), 2157–2170.

Jones, K., & Leonard, L. (2008). Trust in Consumer-to-Consumer Electronic Commerce. *Information & Management*, *45*(2), 88–95.

Judge, W. Q., & Douglas, T. J. (1998). Performance Implications of Incorporating Natural Environmental Issues into the Strategic Planning Process: An Empirical Assessment. *Journal of Management Studies*, *35*(2), 241–262.

King, A. A., & Lenox, M. J. (2000). Industry Self-Regulation without Sanctions: The Chemical Industry's Responsible Care Program. *Academy of Management Journal*, *43*(4), 698–716.

Klassen, R D, & McLaughlin, C. P. (1996). The Impact of Environmental Management on Firm Performance. *Management Science*, *42*(8), 1199–1214.

Klassen, Robert D, & Whybark, D. C. (1999). The Impact of Environmental Technologies on Manufacturing Performance. *Academy of Management Journal*, *42*(6), 599–615.

Kostova, T. (1999). Transnational Transfer of Strategic Organizational Practices: A Contextual Perspective. *Academy of Management Review*, *24*(2), 308.

Kostova, T., Roth, K., & Dacin, M. T. (2008). Institutuonal Theory in the Study of Multinational Corporations: A critique and New Directions. *Academy of Management Review*, *33*(4), 994–1006.

Kuan, H., & Bock, G. (2007). Trust Transference in Brick and Click Retailers: An Investigation of the Before-Online-Visit Phase. *Information & Management*, *44*(2), 175–187.

Lanjouw, J. O., & Mody, A. (1996). Innovation and the International Diffusion of Environmentally Responsive Technology. *Research Policy*, *25*(4), 549–571.

Lee, D., Kauffman, R. J., & Bergen, M. E. (2009). Image Effects and Rational Inattention in Internet-Based Selling. *International Journal of Electronic Commerce*, *13*(4), 127–166.

Lee, O. K., Sambamurthy, V., Lim, K., & Wei, K. K. (2008). *IT-Enabled Organizational Agility and Sustainable Competitive Advantage*. Disponible en http://ssrn.com/abstract=1249301.

Lemley, M. A., & Shapiro, C. (2005). Probabilistic Patents. *The Journal of Economic Perspectives*, *19*(2), 75–98.

Leonard, H. (1988). *Pollution and the Struggle for the World Product: Multinational Corporations, Environment, and International Comparative Advantage*. Cambridge University Press.

Margolis, J. D., & Walsh, J. P. (2003). Misery Loves Companies: Rethinking Social Initiatives by Business. *Administrative Science Quarterly*, *48*(2), 268–305.

Markman, G., Espina, M. I., & Phan, P. H. (2004). Patents as Surrogates for Inimitable and Non-Substitutable Resources. *Journal of Management*, *30*(4), 529–544.

Martin-Tapia, I., Aragon-Correa, J. A., & Senise-Barrio, M. E. (2008). Being Green and Export Intensity of SMEs: The Moderating Influence of Perceived Uncertainty. *Ecological Economics*, *68*(1-2), 56–67.

Melville, N., Kraemer, K., & Gurbaxani, V. (2004). Review: Information Technology and Organizational Performance: An Integrative Model of IT Business Value. *MIS Quarterly*, *28*(2), 283–322.

Morrison, D. E., & Dunlap, R. E. (1986). Environmentalism and Elitism: a Conceptual and Empirical Analysis. *Environmental Management*, *10*(5), 581–589.

Newton, T., & Harte, G. (1997). Green Business: Technicist Kitsch? *Journal of Management Studies*, *34*(1), 75–98.

NSF-US. (2012). *Science and Engineering Indicators*. Disponible en http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA278733.

OECD. (1994). *The Measurement of Scientific and Technological Activities Using Patent Data as Science and Technology Indicators*. Disponible en http://www.oecd-ilibrary.org/science-and-technology/the-measurement-of-scientific-and-technological-activities-using-patent-data-as-science-and-technology-indicators_9789264065574-en.

OECD. (2009). OECD *Patent Statistics Manual*. Disponible en http://www.oecd.org/science/innovationinsciencetechnologyandindustry/oecdpatentstatisticsmanual.htm.

Pakes, A., & Griliches, Z. (1984). Patents and R & D at the Firm Level : A First Look. Disponible en *http://www.nber.org/papers/w0561.pdf?new_window=1*.

Park, W. G. (2010). On Patenting Costs. *The WIPO Journal*, *2*(1), 38-48.

Peters, M., Schneider, M., Griesshaber, T., & Hoffmann, V. H. (2012). The Impact of Technology-Push and Demand-Pull Policies on Technical Change – Does the Locus of Policies Matter? *Research Policy*, *41*(8), 1296–1308.

Pisano, G. (2006). Profiting from Innovation and the Intellectual Property Revolution. *Research Policy*, *35*(8), 1122–1130.

Porter, M. E., & Van Der Linde, C. (1995a). Toward a New Conception of the Environment-Competitiveness Relationship. *Journal of Economic Perspectives*, *9*(4), 97–118.

Ramus, C. A., & Steger, U. (2000). The Roles of Supervisory Support Behaviors and Environmental Policy in Employee “Ecoinitiatives” at Leading-Edge European Companies. *Academy of Management Journal*, *43*(4), 605–626.

Rappaport, A., & Flaherty, M. F. (1992). *Corporate Responses to Environmental Challenges: Initiatives by Multinational Management*. Quorum Books.

Rugman, A. M., & Verbeke, A. (1998a). Corporate Strategy and International Environmental Policy. *Journal of International Business Studies*, *29*(4), 819–833.

Rugman, A. M., & Verbeke, A. (1998b). Corporate Strategies and Environmental Regulations: an Organizing Framework. *Strategic Management Journal*, *19*(4), 363–375.

Russo, M. V., & Fouts, P. A. (1997). A Resource-Based Perspective on Corporate Environmental Performance and Profitability. *Academy of Management Journal*, *40*(3), 534–559.

Sapienza, H. J., Autio, E., George, G., & Zahra, S. A. (2006). A Capabilities Perspective on the Effects of Early Internationalization on Firm Survival and Growth. *Academy of Management Review*, *31*(4), 914–933.

Scherer, F. M. (1965). Corporate Inventive Output, Profits, and Growth. *Journal of Political Economy*, *73*(3), 290.

Scherer, F. M. (1982). Inter-Industry Technology Flows and Productivity Growth. *The Review of Economics and Statistics*, *64*(4), 627.

Schmookler, J. (1966). *Invention and Economic Growth*. Harvard University Press.

Sharma, S. (2000). Managerial Interpretations and Organizational Context as Predictors of Corporate Choice of Environmental Strategy. *Academy of Management Journal*, *43*(4), 681–697.

Sharma, S., & Henriques, I. (2005). Stakeholder Influences on Sustainability Practices in the Canadian Forest Products Industry. *Strategic Management Journal*, *26*(2), 159–180.

Sharma, S., & Vredenburg, H. (1998). Proactive Corporate Environmental Strategy and the Development of Competitively Valuable Organizational Capabilities. *Strategic Management Journal*, *19*(8), 729–753.

Shrivastava, P. (1995). The Role of Corporations in Achieving Ecological Sustainability. *Academy of Management Review*, *20*(4), 936–960.

Sidorova, A., Evangelopoulos, N., Valacich, J. S., & Ramakrishnan, T. (2008). Uncovering the Intellectual Core of the Information Systems Discipline. *MIS Quarterly*, *32*(3), 467–482.

Somaya, D. (2012). Patent Strategy and Management: An Integrative Review and Research Agenda. *Journal of Management*, *38*(4), 1084–1114.

USPTO. (2009). *The U.S. Commerce Department's Patent and Trademark Office (USPTO) will pilot a Program to Accelerate the Examination of Certain Green Technology Patent Applications*. Press Release USPTO. Disponible en http://www.uspto.gov/news/pr/2009/09_33.jsp

Veefkind, V., Hurtado-Albir, J., Angelucci, S., Karachalios, K., & Thumm, N. (2012). A New EPO Classification Scheme for Climate Change Mitigation Technologies. *World Patent Information*, *34*(2), 106–111.

Vernon, R. (1992). Transnational Corporations: Where are They Coming From, Where are They Headed? *Transnational Corporations*, *1*(2), 7–35.

Vries, D. (2012). *Leveraging Patents Financially. A Company Prespective*. Springer.

WCED, (1987). *Our Common Future*. Oxford University Press.

Westley, F., & Vredenburg, H. (1997). Interorganizational Collaboration and the Preservation of Global Biodiversity. *Organization Science*, *8*(4), 381–403.

Yeoh, P. L. (2004). International Learning: Antecedents and Performance Implications Among Newly Internationalizing Companies in an Exporting Context. *International Marketing Review*, *21*(4/5), 511–535.

CAPÍTULO 2

REPRESENTACIÓN, ALMACENAMIENTO E INTERCAMBIO DE INFORMACIÓN Y CONOCIMIENTO

REPRESENTACIÓN, ALMACENAMIENTO E INTERCAMBIO DE INFORMACIÓN Y CONOCIMIENTO

2.1 Introducción

En las últimas décadas, los sistemas de información, IS (*Information Systems*), han crecido en importancia. La mayoría de los gestores de empresas piensan que los IS son críticos para el éxito de la estrategia empresarial, y la inversión en IS representa una gran parte del presupuesto de las empresas (Chen et al., 2010). Sin embargo, hasta estos momentos no es suficiente la inversión en tecnologías de la información y las comunicaciones, en general y en IS en particular, para obtener unos resultados financieros de la empresa positivos y proporcionales a la inversión realizada (Malhotra, 2005). La eficiente utilización de dichos sistemas, e incluso la motivación de los usuarios, juega un papel muy importante en el rendimiento y beneficios a obtener de las tecnologías de la información (Bermúdez-Edo et al., 2008; Jennex et al., 2012; Malhotra & Galletta, 2003; Strassmann, 1997).

Los sistemas de gestión del conocimiento, KMS (*Knowledge Management Systems*), son un tipo especial de sistemas de información que ayudan y mejoran los procesos de conocimiento a la par que examinan la satisfacción del usuario y los beneficios percibidos por el uso de dichos sistemas (Karlinsky-Shichor & Zviran, 2012). La gestión del conocimiento es la práctica selectiva de aplicar el conocimiento de experiencias previas en la toma de decisiones para situaciones presentes y futuras mejorando la efectividad de la empresa. Es una práctica que crea sinergias entre las capacidades de los procesos de información, vistos desde la perspectiva de la informática con las capacidades innovativas de los elementos humanos y sociales de la empresa (Malhorta, 2001).

Los KMS se definen, pues, como los sistemas que son creados para facilitar la captura, almacenamiento, recuperación y reutilización del conocimiento (Alavi & Leidner, 2001). Por lo tanto, son sistemas de búsqueda, organización, filtrado y presentación de la información con el objetivo de mejorar la compresión entre las personas en un dominio específico de conocimiento (Davenport & Prusak, 1997).

Mientras que los IS se ocupan de las operaciones y gestión de las empresas, los KMS añaden valor al integrar piezas de información aparentemente inconexas, pero que son relevantes para un contexto particular, y necesarias para llevar a cabo la gestión eficiente de las empresa (Alavi & Leidner, 2001; Melville, 2010). Ello se realiza en muchas ocasiones con métodos de inteligencia artificial y sistemas colaborativos para compartir información.

Las tecnologías de bases de datos se desarrollaron en los años 60 y 70, y han emergido como la mayor tecnología subyacente en los sistemas de información. En particular, las bases de datos relacionales son las más utilizadas para el almacenamiento de información estructurada (Alexiev et al., 2005).

Hace años la información se guardaba directamente ligada a una aplicación específica, dando lugar a dificultades de conceptualización y técnicas en la utilización desde cualquier otra aplicación: redundancia, inconsistencia e imposibilidad de reutilización de datos. Con la aparición de Internet, se hizo relevante la necesidad de intercambiar datos entre aplicaciones y empresas. Con todos estos desarrollos todavía existe el problema de que los datos están almacenados en miles de formatos incompatibles y que no pueden ser sistemáticamente gestionados o integrados.

Esta incompatibilidad no está solo ligada a diferentes tecnologías de representación de datos (por ejemplo, CVS, XML, etc.), o a las múltiples tecnologías de almacenamiento de información (por ejemplo, las bases de datos relacionales pueden gestionarse con distintos Sistemas de Gestión de Bases de Datos, cómo Oracle, DB2, SQL Server, Sybase, etc. Esta incompatibilidad también reside en las diferencias semánticas, cada conjunto de datos posee su propio vocabulario y esquema.

Por lo tanto, es importante encontrar un modo eficiente de gestionar múltiples aplicaciones y fuentes de información, que permita la interoperabilidad entre aplicaciones e independencia de las tecnologías específicas subyacentes. Esto requiere un cambio del enfoque de IS evolucionando de una visión basada principalmente en aplicaciones a una visión en la cual la propia información tome un papel central (Schwartz & Schreiber, 2005), lo cual pasa por dotar de semántica común a la información.

Para conseguir tener esta visión de IS dotando de mayor relevancia a la información, es necesario tener claras las diferencias entre datos, información, conocimiento y sabiduría, lo cual se describirá con más detalle en la siguiente sección.

El resto del capítulo se divide en siete partes principales. La Sección 2.2 introduce los conceptos de datos, información, conocimiento y sabiduría. La Sección 2.3 repasa las diferentes tecnologías de representación de la información existentes. La Sección 2.4 muestra el almacenamiento de la información. La Sección 2.5 introduce la Web semántica, precursora de las ontologías. La Sección 2.6 recopila las diferentes formas de recuperación de la información. La Sección 2.7 introduce la inferencia del conocimiento. Finalmente, la Sección 2.8 reúne los contenidos de las secciones anteriores bajo el punto de vista de la gestión de información y conocimiento, y en particular para el caso de las patentes.

2.2 Datos, Información, Conocimiento y Sabiduría

Se conoce cómo la pirámide del conocimiento a la pirámide formada desde la base a la cumbre por los cuatro elementos: datos, información, conocimiento y sabiduría.

Los **datos** son un conjunto de hechos discretos sobre eventos. En el entorno de la empresa, los datos son piezas estructuradas de transacciones comerciales (Davenport et al., 1997). La **información** es un conjunto de datos procesados de manera que se le da relevancia, significado y contexto a dichos datos (véase Figura 2.1). Este procesamiento disminuye la incertidumbre de los datos, y los hace útiles para la toma de decisiones. Este procesamiento se suele realizar relacionando unos datos con otros. El **conocimiento** es información organizada, viene definida al añadir experiencia a la información, y tiene un componente dinámico, que varía con la experiencia. El último nivel del conocimiento es la **sabiduría**, que es el saber comunicar el conocimiento poniéndolo en el contexto de la audiencia. Este último escalón está, al menos de momento, fuera del alcance de los sistemas IS.

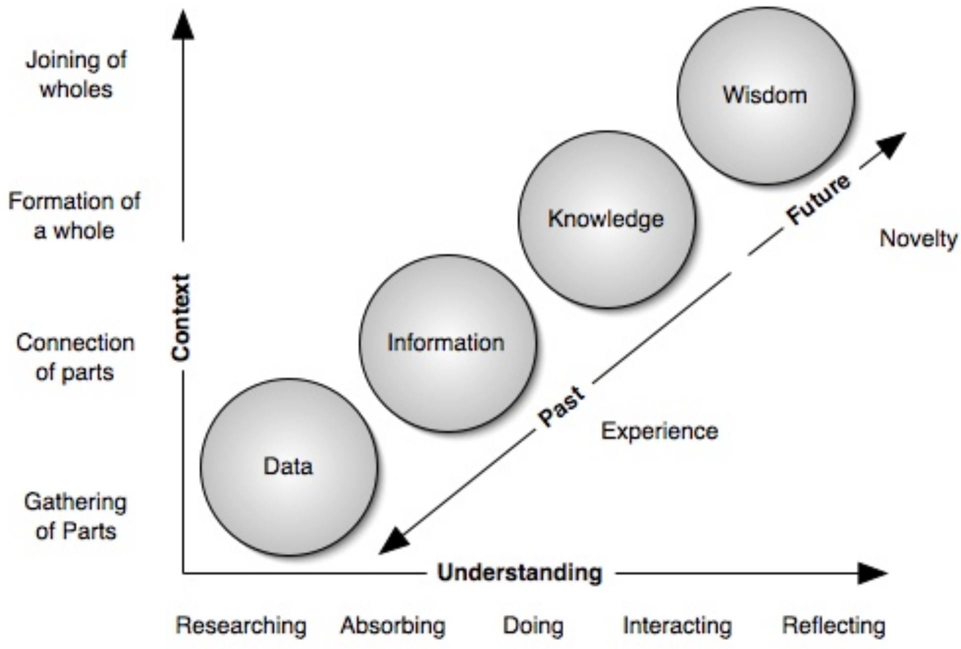

Figura 2.1 Pasos de la pirámide de conocimiento (Donald Clark)[1]

En el área de los sistemas de información, y dentro del contexto de esta tesis, las bases de datos contienen los datos, las relaciones entre esos datos (como por ejemplo, las patentes que tiene una determinada empresa) son información, y las ontologías permiten inferir conocimiento. La sabiduría depende del uso particular que se le dé a las ontologías.

A continuación se tratarán cada uno los elementos que componen un sistema de gestión del conocimiento desde el punto de vista tecnológico, centrándonos en las tecnologías de representación, almacenamiento, recuperación y gestión de la información y el conocimiento.

2.3 Tecnologías de Representación de la Información

El desarrollo de sistemas de información, requiere de la participación de diversos actores: desarrolladores, clientes, evaluadores, etc. Uno de los primeros pasos

[1] http://www.nwlink.com/~donclark/performance/understanding.html

del ciclo de vida de desarrollo de aplicaciones, tras la especificación de requisitos, es el modelado de datos. Este modelado toma especial relevancia en sistemas software complejos, o que requieren la participación de diferentes actores, ya que el modelado representa solamente los diferentes tipos de datos (no todos los datos) que se consideran en el sistema y sus relaciones mostrándolo de forma gráfica e intuitiva, lo cual permite una colaboración más eficaz entre los diferentes actores.

A principio de los 70 aparecieron los modelos de datos, concebidos como una forma de describir la estructura de una base de datos abstrayendo al usuario de detalles de implementación.

Recientemente han comenzado a utilizarse los modelos semánticos, basados en lógica descriptiva, cuya principal ventaja sobre los modelos de datos tradicionales es la aportación de una riqueza semántica que permite compartir mejor los datos entre aplicaciones. Este modelado semántico tiene su máximo exponente en las ontologías.

2.3.1 Modelos de Datos

A lo largo de la historia se han propuesto varios modelos de datos. En esta sección se describen los dos modelos más utilizados en la actualidad, el *Modelo Relacional* y el *Modelo Entidad-Relación*.

El modelo relacional (Codd, 1970) propone que un sistema de bases de datos se represente con los datos organizados en tablas bidimensionales, llamadas relaciones (Garcia-Molina, et al., 2009). El modelo relacional cuenta con unos conceptos básicos de algebra relacional, cuyas propiedades están bien definidas. Esta base matemática permite definir relaciones, propiedades y algunas operaciones sobre ellas, pero ofrece una semántica pobre acerca de cómo interpretar las relaciones del mundo real (Schmid & Swenson, 1975). En la Figura 2.2 vemos un modelo relacional, representado por tablas que están relacionadas entre ellas a través de campos comunes.

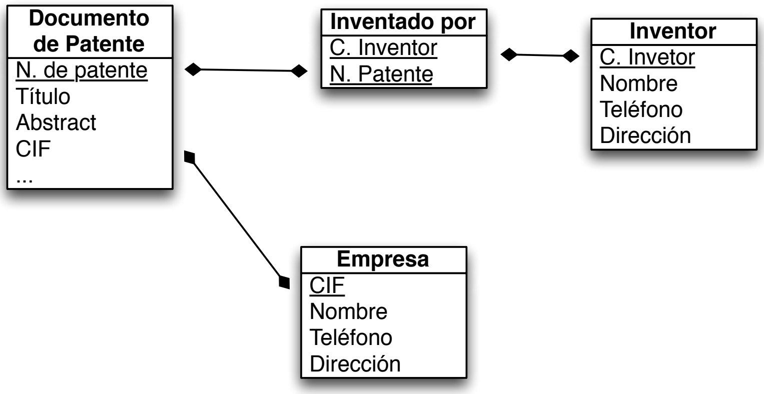

Figura 2.2 Ejemplo de modelo relacional

El modelo entidad-relación, (Chen, 1976) surge para suplir la pobre semántica del modelo relacional (Thalheim, 2000). El modelo entidad-relación combina el modelo relacional con redes semánticas, al mismo tiempo que representa grafos simples para la representación de la información contenida en las bases de datos. Este modelo consta de tres elementos: las entidades que representan objetos del mudo real, los atributos que son características que definen e identifican a las entidades y las relaciones que describen las dependencias o asociaciones entre entidades. La Figura 2.3 muestra un ejemplo del modelo entidad-relación, donde las entidades de la base de datos se representan mediante rectángulos, los atributos de dichas entidades se representan mediante óvalos y las relaciones entre las entidades mediante rombos.

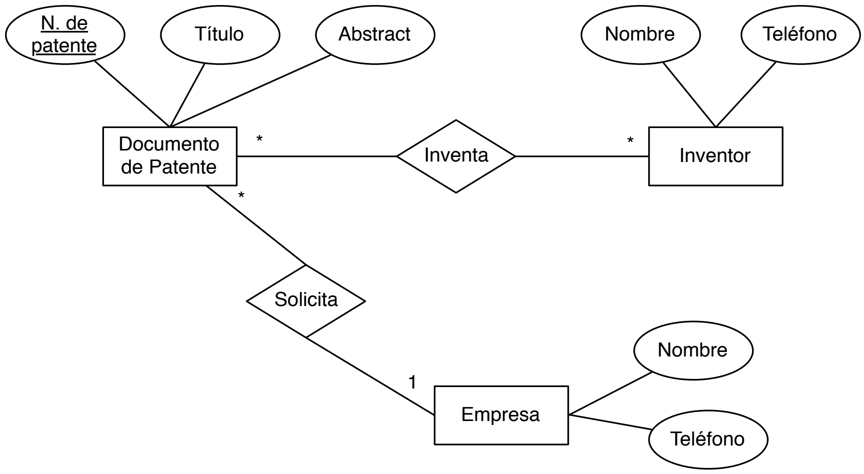

Figura 2.3 Ejemplo de modelo entidad-relación

Estos modelos de datos se limitan a estructurar los datos con una cierta semántica. Sin embargo, al ir subiendo en complejidad los sistemas de información, y al crecer la información contenida en ellos, se hizo necesario buscar abstracciones que dotaran a los modelos de una mayor capacidad expresiva. Es entonces cuando se empiezan a utilizar los modelos semánticos de representación de la información, basados en lógica descriptiva, siendo las ontologías el modelo semántico más extendido.

2.3.2 Ontologías

Las ontologías, al contrario que los modelos tradicionales de datos, dotan a la información de cierta semántica que permite una eficiente inferencia de información. La ingeniería de ontologías constituye un campo reciente de investigación y en constante evolución que cada vez se aplica a más áreas de conocimiento. Grandes empresas y organizaciones están apostando por el desarrollo de ontologías en su representación de la información. Entre otras, destacamos la Agencia Aeroespacial Americana NASA que ha desarrollado la ontología SWEET[2] (*Semantic Web for Earth and Environmental Terminology*), y las Naciones Unidas que han desarrollado la ontología *Geopolitical Ontology*[3]

A continuación recopilamos las principales definiciones de ontología, comentamos los tipos de ontologías y explicamos los pasos necesarios en el desarrollo de una ontología.

Definiciones de Ontología

El término ontología se utiliza en el área de la filosofía desde 1613 (Goclenius, 1613), para definir, la rama de la metafísica que estudia el ser en general y sus propiedades transcendentales.

En el área de la informática el término ontología se adoptó en la década de 1980, en el contexto de la Web semántica (véase la Sección 2.5), por la comunidad de Inteligencia Artificial, cuyos miembros argumentaban que podían crear nuevas ontologías, entendidas como modelos computacionales, que permitieran ciertos tipos de razonamiento automático. A partir de entonces han sido muchos los autores que han tratado de dar una definición formal al termino ontología en el ámbito de la representación del conocimiento. Entre las más citadas se encuentran:

- "*An explicit specification of a conceptualization*" (Gruber, 1993).
- "*An ontology for a body of knowledge concerning a particular task or domain describes a taxonomy of concepts for that task or domain that define the semantic interpretation of the knowledge*" (Alberts, 1993).

[2] http://sweet.jpl.nasa.gov/
[3] http://www.fao.org/countryprofiles/geoinfo/en/

- "*An ontology is a theory of what entities can exist in the mind of a knowledgeable agent*" (Wielinga & Schreiber, 1994).

- "*An ontology is an explicit, partial account of a conceptualization*" (Guarino & Giaretta, 1995).

- "*An explicit knowledge-level specification of a conceptualization, i.e. the set of distinctions that are meaningful to an agent. The conceptualization –and therefore the ontology– may be affected by the particular domain and the particular task it is intended for*" (Van Heijst et al., 1997).

- "*An explicit, partial account of the intended models of a logical language*" (Guarino, 1997).

- "*A formal specification of a shared conceptualization*" (Borst, 1997).

- "*A logical theory accounting for the intended meaning of a formal vocabulary, i.e. its ontological commitment to a particular conceptualization of the world. The intended models of a logical language using such a vocabulary are constrained by its ontological commitment. An ontology indirectly reflects this commitment (and the underlying conceptualization) by approximating these intended models"* (Guarino, 1998).

- "*Ontologies are content theories about the sorts of objects, properties of objects, and relations between objects that are possible in a specified domain of knowledge. They provide potential terms for describing our knowledge about the domain*" (Chandrasekaran et al., 1999).

- "*A formal description of entities and their properties, relationships, constraints, behaviors. It provides a common terminology that captures key distinctions and is generic across many domains*" (Grüninger et al., 2000).

- "*A formal explicit description of concepts in a domain of discourse (classes –sometimes called concepts–), properties of each concept describing various features and attributes of the concept (slots –sometimes called roles or properties–), and restrictions on slots (facets –sometimes called role restrictions–)*" (Noy & McGuinness, 2001).

Para el propósito de esta tesis adoptaremos la última de estas definiciones que define las ontologías cómo una descripción formal de conceptos, sus propiedades, relaciones, restricciones y comportamiento (Noy & McGuinness, 2001), y que a su vez está basada en la definición más citada de las existentes de Tom Gruber (1993), quien propone que una ontología es la especificación explícita de una conceptualización.

Esta formalización explícita y sistemática de los conceptos permite definir vocabularios comunes para cada dominio particular con múltiples utilidades como, por ejemplo, permitir que usuarios o agentes software puedan hacer búsquedas automáticamente por Internet en diferentes sitios Web. A veces las ontologías se comparan con taxonomías de jerarquías de clases, que ayudan a mantener una organización jerárquica de conceptos, destacando así su utilidad para administrar grandes colecciones de datos.

Sin embargo, el uso mayoritario de las ontologías es la compartición de una estructura común de información, de forma que esta pueda ser entendible por personas y máquinas. Esto permite, dentro de un dominio, la reutilización de conocimiento, hacer explicito y analizar formalmente el conocimiento, y separar el conocimiento del dominio, del conocimiento operacional (Cardoso, 2007).

Atendiendo a la definición de ontología se puede derivar que los elementos claves de las ontologías, según Gómez-Perez et al. (2004) y Noy y McGuinness (2001), consisten en:

- *Clases* o *conceptos* (*Ci*). Son las ideas a formalizar y representan los conceptos del dominio. Las clases en una ontología se suelen organizar en taxonomías, o jerarquías, a las que se les pueden aplicar las propiedades de herencia. Es decir, si *A* es una superclase de *B*, entonces cualquier instancia de *B* es necesariamente una instancia de *A*.

- *Propiedades* de dichos conceptos, también denominados *ranuras* (*slots*) o *roles*. Las propiedades describen características o atributos de las clases.

- *Restricciones sobre las propiedades*. Denominadas *facetas* (*facets*).

- *Relaciones* (*R: C1 x C2 x … x Cn-1 x Cn*). Representan asociaciones entre las clases (*Ci*). Las relaciones más habituales son binarias. Por ejemplo:

    - *SubclaseDe: Concepto1 x Concepto2* (ejemplo: *SubclaseDe: Patente x DocumentoDePatente*)

- *Funciones* (*F: C1 x C2 x … x Cn-1 -> Cn*). Son casos especiales de relaciones donde se identifican elementos mediante el cálculo de una función. Por ejemplo:

- *Madre_ de: Persona -> Mujer*

- *Precio_ coche_ usado: Modelo x Año x Kilómetros -> Precio*

- *Instancias*. Se usan para representar elementos o individuos en una ontología.

- *Axiomas*. Los axiomas formales sirven para modelar sentencias que son siempre ciertas. Normalmente se usan para representar conocimiento que no puede ser formalmente definido por los componentes descritos anteriormente. También se usan para verificar la consistencia de la propia ontología.

La Figura 2.4 muestra una representación gráfica simplificada de una ontología. Como se puede apreciar, se representan los conceptos, mediante óvalos, y las relaciones entre conceptos mediante flechas.

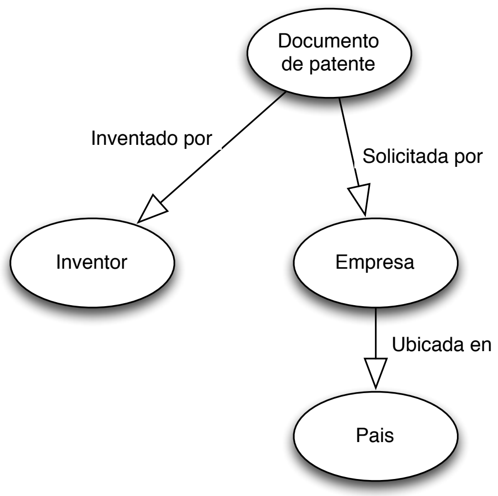

Figura 2.4 Ejemplo de modelo de ontologías

Tipos de Ontologías

Existen distintas clasificaciones de ontologías [véase (Dahlem & Hahn, 2009) para una amplia recopilación de clasificaciones]. En esta sección se introducen dos de las más reconocidas, y que ayudarán a comprender los siguientes capítulos de esta tesis.

Según el grado de generalidad. Esta clasificación diferencia las ontologías según su aplicabilidad más o menos amplia a diferentes dominios. En la Figura 2.5 se recogen los cuatro tipos de ontologías definidos por Guarino (1997). Las flechas representan relaciones de especialización.

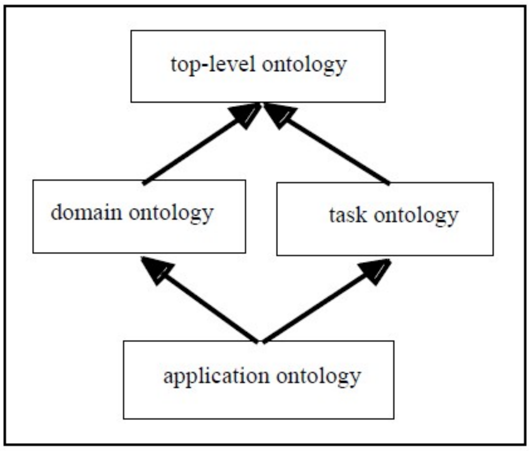

Figura 2.5 Taxonomía de ontologías según su generalidad (Guarino, 1997)

Esta clasificación es también apoyada por Van Heijst et al. (1997) distinguiendo entre:

- *Ontologías de nivel alto*. Describen conceptos generales que se utilizan en muchos dominios, como *espacio*, *tiempo*, *materia*, *objeto*, etc. Estas ontologías se suelen utilizar como base para crear otras ontologías más específicas acerca de un dominio que hagan uso de los conceptos de las ontologías de nivel alto.

- *Ontologías de dominio*. Definen un vocabulario genérico con conceptos, y relaciones entre ellos, de un dominio específico, como el dominio de las finanzas, el de las patentes, o el del medioambiente. Los conceptos y relaciones de una ontología de dominio normalmente son especializaciones de conceptos y relaciones que aparecen en ontologías de alto nivel.

- *Ontologías de tareas*. Definen el vocabulario relativo a una tarea genérica o actividad como, por ejemplo, comercializar, planificar, etc. Este tipo de ontologías también utiliza los términos introducidos en ontologías de alto nivel.

- *Ontologías de aplicación*. Describen conceptos que dependen de los dos tipos anteriores de ontologías (dominios y tareas). Con frecuencia, estos conceptos corresponden a roles desempeñados por las entidades de un dominio al realizar ciertas tareas.

Según el nivel de detalle en su especificación. Lassila y McGuinness (2001) proponen una clasificación o espectro de ontologías en función de la riqueza de su estructura interna (Figura 2.6). Siguiendo un *espectro* o *rango* lineal, las ontologías se clasifican en:

- *Vocabularios controlados.* Este tipo es el más básico y se limita a enumerar una lista finita de términos, como puede ser un catálogo.

- *Glosarios.* Como evolución al tipo anterior, la lista de términos se complementa con el significado de dichos términos, expresado por medio de lenguaje natural.

- *Tesauros.* Añaden información adicional acerca de las relaciones entre términos, como la de sinonimia.

- *Jerarquías de términos informales.* Esta categoría recoge ciertas ontologías existentes en la Web, que no cumplen estrictamente la relación de transitividad en todas sus relaciones jerárquicas entre conceptos. Es decir, una instancia de una clase puede no ser instancia de otra clase más general que englobe a la primera clase.

- *Jerarquías de subclases formales o estrictas*. Este tipo de ontologías cumplen la relación de transitividad en sus jerarquías. Las jerarquías de subclases estrictas es la primera categoría que permite explotar una de las principales ventajas de las ontologías, la herencia. A través de la herencia, se puede inferir nueva información no explícitamente representada a priori.

- *Jerarquías de subclases formales y relaciones.* Estas ontologías no sólo cumplen las relaciones de herencia, sino que además permiten definir relaciones o propiedades entre conceptos.

- *Marcos (frames).* En estas ontologías las instancias de una clase pueden contener información acerca de sus propiedades. Estas propiedades también se heredan de forma jerárquica por las instancias de las clases inferiores.

- *Restricciones de valor.* En estas ontologías se pueden definir restricciones acerca de qué valores pueden tomar las propiedades de una clase o una instancia.

- *Restricciones lógicas con carácter general*. Estas son las ontologías más expresivas. Se basan en la lógica de primer orden para definir restricciones entre clases o instancias.

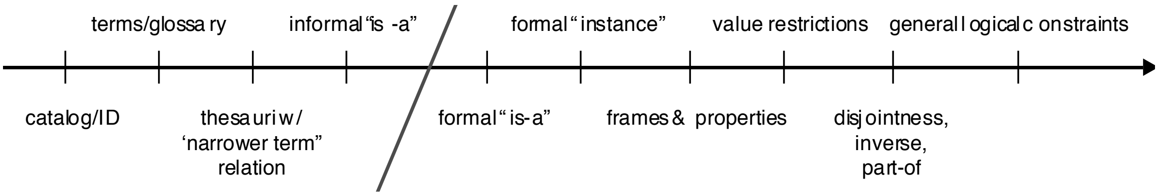

Figura 2.6 Espectro de ontologías (Lassila & McGuinness, 2001)

Desarrollo de una Ontología

El desarrollo de una ontología requiere de una metodología formal que reduzca la probabilidad de generar una base de conocimiento inconsistente, y que evite problemas estructurales cuando se integre la ontología en otros sistemas de información (Chandrasekaran et al., 1999). Se han propuesto varios métodos para la creación de ontologías (véase Dahlem & Hahn, 2009 para una recopilación de los más relevantes), pero sin duda alguna, el más extendido es el conocido cómo "*Ontology Development 101*" (Noy & McGuinness, 2001), que consta de los siguientes pasos:

1. Determinar el dominio y objetivos de la ontología.

2. Considerar reutilizar ontologías existentes, teniendo en cuenta que si existiera una ontología sobre el dominio seleccionado, se debería priorizar su uso o, en otro caso, clarificar el interés por crear una nueva ontología.

3. Enumerar los términos importantes en la ontología.

4. Definir las clases y la jerarquía de clases en un lenguaje de ontologías. Estas clases representarán los términos importantes de la ontología. Por ejemplo,

la clase *patente*, o la clase *inventor* podrían ser apropiadas para una ontología de patentes.

5. Definir las propiedades de las clases. Estas propiedades nos permiten definir las relaciones existentes entre los conceptos de la ontología. En el ejemplo de la ontología de patentes, las propiedades pueden ser: el título de la patente (*hasTitle*), el inventor (*hasInventor*), la fecha de presentación (*hasFilingDate*), etc.

6. Definir las *facetas* de las propiedades. Las *facetas* nos describen las restricciones de las propiedades. Por ejemplo: La fecha de presentación de la patente debe ser un tipo *fecha*, o la patente debe contener al menos un inventor, etc.

7. Crear las instancias (o poblar la ontología). Consiste en dar valores a los conceptos o propiedades, por ejemplo *Juan López*, *Molino de viento para generar energía con doble astas*, o *25/08/2011*.

En el Capítulo 6 de esta tesis proponemos un método que extenderá el método "*Ontology Development 101*" de Noy y McGuinness, para poder añadir axiomas, y de este modo explotar el poder de razonamiento de las ontologías.

El desarrollo de una ontología y de sus axiomas requiere del uso de editores, que son herramientas que permiten escribir texto, o programas en un lenguaje determinado. Existen varios editores específicos para la creación de ontologías, por ejemplo SWOOP

(Kalyanpur et al., 2006) o NeOn[4] (Haase, 2008). Uno de los editores de ontologías más utilizado es Protégé (Noy et al., 2000).

2.4 Almacenamiento de la Información en Bases de Datos

En los primeros años en que comenzaron a incorporarse los ordenadores a la empresa, la información se consideraba algo estático y estaba almacenada en ficheros de datos. Este sistema tradicional de ficheros acarreaba serios problemas de:

- Redundancia de datos, puesto que podían existir datos duplicados en diferentes ficheros.
- Inconsistencia entre ficheros que contenían distintas actualizaciones de los mismos datos.
- Dependencia de las aplicaciones para las cuales fueron creados. Los datos de cada fichero se podía recuperar únicamente con la aplicación con la cual fueron creados.
- Rigidez en la búsqueda de información, lo que implicaba que cada fichero se debía buscar de forma independiente.
- Problemas de seguridad y confidencialidad de los datos.

A mediados de los años 1970 se empezó a tomar conciencia de que la información era algo dinámico y de que se necesitaba un sistema de gestión de la misma más eficiente. Como consecuencia surgieron las bases de datos.

[4] http://neon-toolkit.org

Una base de datos es un conjunto de datos integrados, con redundancia controlada y con una estructura que refleja las interrelaciones existentes en el mundo real. Los datos que han de ser compartidos por diferentes usuarios y aplicaciones deben mantenerse independientes de éstas (De-Miguel et al., 1999).

El motor de una base de datos es el Sistema Gestor de Base de Datos (SGBD), una capa software formada por un conjunto de programas, procedimientos y lenguajes que permiten a los usuarios describir, recuperar y manipular los datos almacenados en la base de datos. Los SGBD imponen un modelo de datos. El más conocido es el modelo relacional, que se ha presentado en la Sección 2.3.2. Las bases de datos que siguen este modelo se llaman bases de datos relacionales.

Además de las bases de datos, hoy en día están muy extendidos otros repositorios de información, que almacenan gran cantidad de datos, destacando especialmente Internet. En sus inicios Internet contenía únicamente información sintáctica. Actualmente, para aprovechar toda la información distribuida en Internet, y para poder compartirla y procesarla automáticamente por ordenadores, ha surgido la necesidad de crear la Web semántica. En la siguiente sección abordamos dicha evolución.

2.5 Web Semántica

La World Wide Web, creada alrededor de 1989 por el inglés Tim Berners-Lee con la ayuda del belga Robert Cailliau, introdujo el hipertexto en Internet, a través del lenguaje HTML (*HiperText Markup Language*) (W3C, 2006). Esto permitió que los usuarios, utilizando programas de navegación pudieran acceder a recursos o servicios, tales como, visualizar sitios Web compuestos de páginas HTML que pueden contener texto, imágenes, vídeos u otros contenidos multimedia y navegar de unas páginas a otras a través de los hiperenlaces. HTML consiste en etiquetas que se añaden al texto, lo que permite describir el formato con el cuál se representará este texto en la pantalla. Esta Web sintáctica, compuesta por páginas HTML, se diseñó para ser inteligible por los humanos. Sin embargo, la información está representada como un texto plano, sin anotaciones semánticas, que permitan identificar el tipo de información representada (por ejemplo, si se trata de un teléfono, o del título de un libro). Por lo tanto, no se puede procesar automáticamente por ordenadores, para realizar búsquedas de información, relacionar informaciones de diferentes repositorios, etc.

A finales de los años 90, Berners-Lee (2001) propuso como solución a este problema el desarrollo de una Web Semántica, considerando a la misma como una Web consistente en datos lógicos, en la cual la información tiene un significado bien definido. En la actualidad, Berners-Lee redefine la Web Semántica como una plataforma para aplicaciones distribuidas y compartición de datos. Las ontologías son el corazón de la Web semántica, y aunque se han desarrollado sobre todo para la Web semántica, se empiezan a utilizar en otros campos.

El paso de la Web sintáctica a la Web semántica se ha acelerado por el alto volumen de información disponible a través de la Web, que hace que su procesamiento por humanos sea muy pesado, y se requiere un procesamiento automático, para aligerar el proceso de selección y análisis de la información. En este sentido la Web se puede entender como un sistema de almacenamiento de enormes cantidades de información. Por otro lado, la información contenida en Internet está distribuida en fuentes de información heterogéneas, con lo cual la compartición de información se hace especialmente ardua al tener diversos formatos e idiomas que dan información de un mismo dominio. Actualmente, se está produciendo el cambio de una Web puramente sintáctica a una Web semántica.

2.5.1 La Web Semántica como Base de Datos y Escenario de Intercambio de Información Global

La Web semántica es una extensión de la Web sintáctica que permite a las máquinas interpretar el significado de los datos mediante anotaciones semánticas basadas en descripciones ontológicas.

En la Web semántica, los contenidos además de ser entendidos por humanos, también pueden ser automáticamente procesados por máquinas, salvando las barreras de idioma, y formatos. Esta Web semántica, una vez implantada a nivel global, será de gran importancia para los usuarios de Internet, a la hora de recopilar información de diferentes repositorios, ya que podrán automáticamente procesar información.

2.5.2 Lenguajes de la Web Semántica

El World Wide Web Consortium (W3C) ha publicado una serie de recomendaciones que contienen la definición de diferentes lenguajes estándares y modelos de datos que permiten la especificación de la estructura –y su semántica-, en los documentos diseñados para la Web. Estos lenguajes de metadatos permiten la formalización de la información necesaria para crear una ontología.

Los lenguajes recomendados por la W3C conforman una pila de tecnologías (*Semantic Web Stack*, o también llamada *Semantic Web Cake*) en la que las tecnologías en los niveles superiores explotan y utilizan las capacidades de las capas inferiores (Figura 2.7).

Las tecnologías de las capas inferiores, incluida OWL, están actualmente estandarizadas, y ampliamente reconocidas para la creación de la Web semántica. Las tecnologías de las capas superiores todavía no se han estandarizado, por lo cual la pila de tecnologías podría sufrir variaciones en estas capas superiores.

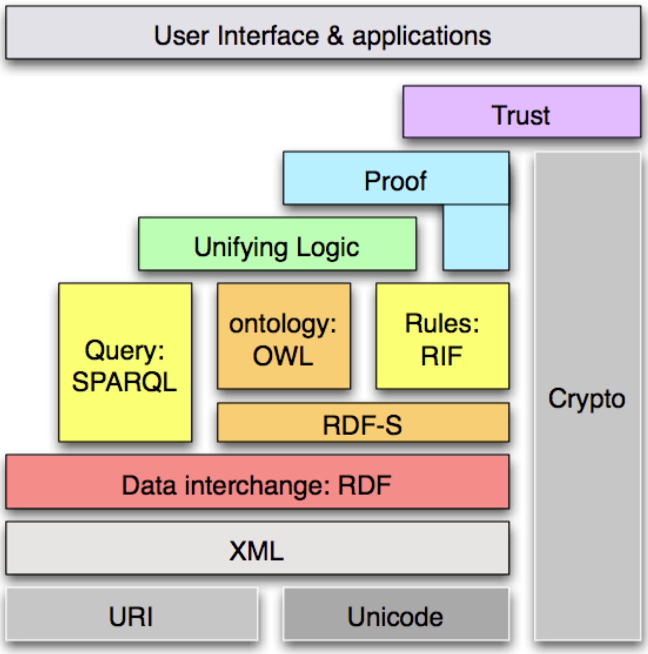

Figura 2.7 Pila de tecnologías de la Web semántica (Berners-Lee)[5]

Las dos capas inferiores representan las tecnologías de hipertexto, que son las utilizadas en la Web sintáctica, y son la base para la Web semántica. Estas tecnologías son:

- *Unicode*. Es un estándar para codificar, representar y manejar texto expresado en la mayoría de los sistemas de escritura, de manera consistente. Este estándar trata de recoger la inmensa mayoría de los caracteres utilizados en los sistemas de escritura mundiales, cómo los caracteres latinos, griegos, árabes, chinos, etc.

- *URI* (*Unifom Resource Identifier*). Es un estándar que identifica de manera única cada uno de los recursos existentes en la Web.

[5] http://www.w3.org/2009/Talks/0120-campus-party-tbl/#(44). Conferencia en el año 2009

- *XML* (*eXtensive Markup Language*) (W3C, 2006). Es un lenguaje de marcado, que permite la creación de documentos basada en datos estructurados. La Web semántica provee de significado a datos estructurados, como los datos en XML. XML se complementa con el lenguaje XML-Schema (*eXtensive Markup Language Schema*) (Fallside & Walmsley 2004), que permite definir la estructura de la información en documentos XML, para que sea conforme a un esquema determinado.

Las capas intermedias ya se consideran tecnologías semánticas para las que se han definido los siguientes estándares:

- *RDF* (*Resource Description Framework*) (Manola & Miller, 2004). Es un modelo de datos que proporciona información descriptiva simple sobre los recursos que se encuentran en la Web. Define objetos (recursos Web) y las relaciones entre ellos con una semántica simple.

- *RDF Schema* (Brickley & Guha, 2004). Proporciona el vocabulario básico para definir propiedades y clases de recursos RDF, permitiendo definir jerarquías de clases y propiedades.

- *OWL* (*Web Ontology Language*) (Smith et al., 2004). Es un lenguaje que amplía el vocabulario de *RDF Schema* para describir un mayor número de propiedades (relaciones) y clases.

- *SPARQL*. Es un lenguaje de consulta sobre RDF. Puede usarse para consultar cualquier dato basado en RDF (por tanto, también las declaraciones que contengan RDFS y OWL). Un lenguaje de consultas es necesario para recuperar información para las aplicaciones de Web Semántica.

Los niveles superiores de la pila de tecnologías Web son una propuesta de tecnologías que todavía no están estandarizadas o incluso contienen solo ideas que deberían ser implementadas.

- *RIF/SWRL*: Son lenguajes de reglas, que permiten escribir relaciones que no pueden describirse directamente en el lenguaje OWL, debido a que son relaciones que no cumplen la lógica descriptiva sobre la que se basa OWL, y los razonadores no podrían inferir conocimiento a través de estas relaciones.

- *Crypto* es la capa de criptografía. La criptografía es necesaria para dotar a la Web semántica de medidas de seguridad, al igual que en la Web sintáctica. Estas medidas asegurarán la confidencialidad de las transacciones, la autenticación de la fuente de información, el no-repudio de la información emitida, o la integridad de los datos, a través de firmas digitales.

- *Trust* es la capa de confianza. Se basa en un sistema de confianzas, dónde los usuarios reciben información de fuentes de confianza, o abaladas por autoridades de certificación, cómo las actuales autoridades de certificación de firmas digitales.

- *La interfaz de usuario* permitirá a los usuarios la interacción con aplicaciones semánticas.

A continuación nos vamos a detener un poco más en detallar las características de los lenguajes utilizados en esta tesis, así como los fundamentos en los que se sustentan.

2.5.2.1 XML y XML Schema

En 1997, XML (*eXtensive Markup Language*) (W3C, 2006) surge como un primer paso para que los contenidos de las páginas Web puedan ser procesados por computadoras. XML es uno de los formatos más extendidos actualmente para compartir información estructurada entre programas o entre personas, permitiendo a la vez el uso de hipertexto. La principal aportación de XML es la separación entre la estructura de los datos y el modo en que estos datos se presentan al usuario, que puede variar en función de la interfaz de usuario o la hoja de estilos empleada.

XML posee unas marcas o etiquetas, representando metadatos, que permiten dar información sobre los diferentes contenidos de un documento o una página Web. En la Figura 2.8 se muestra un fragmento de un documento XML con información sobre una patente. Por ejemplo *<TIEN> <p>Solid oxide fuel cell generator with mid-stack fuel feed</p> </TIEN>*, nos dice que la información contenida entre las etiquetas *<TIEN>* es el título en inglés de la patente.

```
<ROW>
   <DOCUMENT>US 8062798 B2 20111122</DOCUMENT>
   <TIEN>
     <p>Solid oxide fuel cell generator with mid-stack fuel feed</p>
   </TIEN>
   <ECLA>
     <p>H01M8/24D2</p>
     <p>H01M8/04B2C</p>
     <p>H01M8/04C2</p>
     <p>H01M8/24D4</p>
   </ECLA>
   <ECNO />
   <ICO>
     <p>T01M8:06B2</p>
     <p>T01M8:12Y</p>
     <p>Y02E60:12</p>
     <p>Y02E60:50</p>
     <p>Y02E60:52B</p>
     <p>Y02E60:52D</p>
   </ICO>
   <ISG>Y</ISG>
 </ROW>
```

Figura 2.8 Documento XML

Las principales ventajas de XML sobre su predecesor HTML son (Alexiev et al., 2005):

- *Autodescriptivo.* Al ser un formato basado en texto, las etiquetas muchas veces dan una idea del contenido del texto (por ejemplo TIEN, del inglés TItle ENglish, da idea de que el contenido es el título en inglés). Esto permite detectar errores a los humanos.

- *Globalidad.* Cualquier documento XML puede ser entendido por cualquier herramienta XML.

El mayor problema que presenta XML es la carencia de semántica. En el ejemplo, *<title> <p>Solid oxide fuel cell generator with mid-stack fuel feed</p></title>* se entendería cómo una información diferente a la recogida bajo las marcas *<TIEN>*.

XML Schema complementa a XML describiendo la estructura de los documentos XML y añadiéndoles restricciones al contenido de los elementos, como el tipo de datos que debe contener (Figura 2.9). XML Schema permite a un procesador validar un documento XML para asegurar que está libre de inconsistencias.

```
<xs:complexType name="publication-referenceType">
      <xs:sequence>
         <xs:element name="document-id">
            <xs:annotation>
               <xs:documentation>Refers to patents (and patent applications) only; see
                                                       WIPO ST.14</xs:documentation>
            </xs:annotation>
            <xs:complexType>
               <xs:sequence>
                  <xs:element name="country" type="ftxt:countryType" minOccurs="0"/>
                  <xs:element name="doc-number" type="ftxt:doc-numberType"/>
                  <xs:element name="kind" type="ftxt:kindType" minOccurs="0"/>
                  <xs:element name="date" type="ftxt:ICE-date-type" minOccurs="0"/>
               </xs:sequence>
               <xs:attribute name="document-id-type" type="xs:string"/>
            </xs:complexType>
         </xs:element>
      </xs:sequence>
      <xs:attribute name="data-format" type="xs:string"/>
   </xs:complexType>
```

Figura 2.9 Documento XML Schema

XML Schema permite (Alexiev et al., 2005):

- Proporcionar una lista de elementos y atributos en un vocabulario
- Asociar tipos (cadena de caracteres, enteros, etc.) a valores encontrados en el documento
- Para añadir restricciones sobre elementos y atributos como, por ejemplo, que dentro de un documento de patente debe aparecer un título, uno o varios inventores, etc.
- Proporcionar información inteligible por humanos y máquinas

- Dar una descripción formal de uno o más documentos

2.5.2.2 RDF y RDF Schema

A partir del lenguaje XML, distintos autores empiezan a trabajar con lenguajes que permitan relacionar datos y recursos. Uno de los lenguajes con mayor aceptación ha sido el *Resource Description Framework* (RDF) (Manola & Miller, 2004) que dota a los datos de una estructura sintáctica uniforme. Promovido por Tim Berners-Lee, el primer estándar se publicó en 1999 y fue revisado en 2004. Los modelos de datos RDF se construyen como grafos dirigidos que contienen tripletas con recursos Web, propiedades de dichos recursos y valores. Por ejemplo la frase *http://www.ejemplo/patente#inventor* tiene un *nombre* cuyo valor es *José López*, se puede descomponer en la siguiente tripleta:

- Sujeto (URI): *http://www.ejemplo/patente#inventor*
- Predicado (propiedad): *nombre*
- Objeto (valor): *José López*

Este modelo de datos consiste en nodos conectados por arcos etiquetados, donde los nodos representan los recursos Web y los arcos las propiedades de dichos recursos (Figura 2.10). RDF, en definitiva, es un modelo para definir relaciones semánticas entre distintos recursos o URIs.

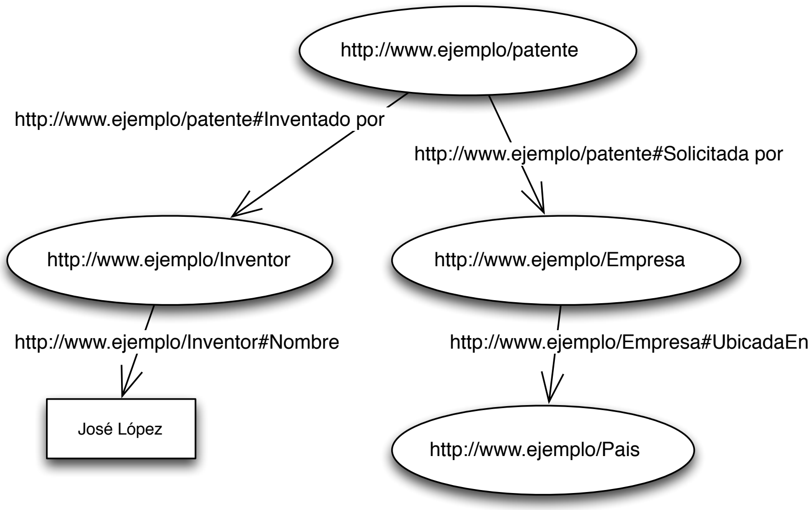

Figura 2.10 Grafo RDF

En la Figura 2.10, se muestra un grafo RDF, dónde el sujeto (desde el cual salen las flechas) consta de 4 predicados, uno de ellos es literal (cajas cuadradas), y los otros tres son objetos de tipo recurso (óvalos). RDF utiliza XML para la codificación de metadatos, y, por lo tanto, el grafo también se puede representar en notación RDF/XML, como se muestra en la Figura 2.11:

```
<rdf:Description rdf:about=" http: //www.ejemplo/patente/>
<rdf:Description rdf:about=" http: //www.ejemplo/patente#Inventor/>
<rdf:Description rdf:about=" http: //www.ejemplo/patente#InventadoPor/>
<rdf:Description rdf:about=" http: //www.ejemplo/patente#Empresa/>
<rdf:Description rdf:about=" http: //www.ejemplo/patente#PresentadaPor/>
```

Figura 2.11 Documento RDF

Una carencia importante de RDF es que no proporciona mecanismos para declarar las propiedades, ni para definir las relaciones entre estas propiedades y otros recursos (Brickley & Guha, 2004). Por ello se definió RDFS (*RDF Schema*) como una extensión a RDF, que se considera ya un lenguaje de ontologías. RDFS define clases, subclases y propiedades (Figura 2.12).

```
<rdf:Class rdf:about="http:// www.example.org/Pat.rdf#IPC">
    <rdfs:subClassOf rdf:resource="http://www.example.org/Pat.rdf#Classification"/>
</rdf:Class>
```

Figura 2.12 Documento RDF Schema

RDF y RDFS describen recursos, valores y relaciones entre ellos, pero no describen de forma estándar las relaciones entre los diferentes recursos. Adicionalmente, en cuanto a la definición de ontologías carecen de características importantes tales como la posibilidad de definir restricciones de cardinalidad, cuantificadores existenciales, propiedades inversas y transitivas, etc. Estos lenguajes carecen de una semántica formal y estándar que asegure que todos los razonadores lleguen a inferir los mismos hechos, por lo tanto, no dan soporte para el razonamiento (Hayes, 2004).

2.5.2.3 OWL y OWL 2

Para suplir algunas de las carencias de RDFS, en 1997/1999, amparado bajo el proyecto Europeo "*on-to-Knowdledge*", se definió el lenguaje OIL (*Ontology Interchange Level*) (Fensel et al., 2000). Paralelamente, en 1999 la agencia DARPA (*Defence and Advance Research Projects Agency*) inicia el programa *Agent Markup Language* que dio lugar al lenguaje DAML (Ankolenkar et al., 2001). Estos dos lenguajes están basados en RDFS y son incompatibles entre sí. En el año 2000, la NSF (*National Science Foundantion* de Estados Unidos) financia varias actividades en Japón y Australia, en un esfuerzo por llegar a una estandarización. Y en 2001 Estados Unidos y la Unión Europea aúnan esfuerzos creando el comité "US/EU *ad hoc Joint Committee on Agent Markup Language*" para la estandarización de estos dos lenguajes y crean DAML+OIL (Horrocks, 2002).

En el año 2002 se crea el grupo de trabajo WebOnt, que desarrolla el estándar OWL (Smith et al., 2004) publicado en 2004, y OWL 2 publicado en 2009, basados ambos en lógica descriptiva. OWL reemplaza al lenguaje DAML+OIL, mejorando algunos aspectos de su diseño y aplicación. OWL es un lenguaje ontológico construido sobre RDFS/XML y que fusiona las características de DAML y de OIL. En los últimos años OWL se ha convertido en el estándar de facto de los lenguajes ontológicos. OWL permite una representación de datos legible por máquinas y completa RDFS incluyendo constructores para representar clases disjuntas, propiedades transitivas, simétricas e inversas, especificar restricciones sobre el rango o la cardinalidad de las relaciones, etc. La Figura 2.13, muestra un fragmento de un documento OWL.

```
<owl:Class rdf:about="http://www.ejemplo.com/patentes.owl#Inventor">
    <rdfs:subClassOf rdf:resource="http://www.ejemplo.com/patentes.owl#Person"/>
    <rdfs:comment rdf:datatype="&xsd;string">The Inventor of a patent</rdfs:comment>
</owl:Class>
<owl:ObjectProperty rdf:about="http://www.ejemplo.com/patentes.owl#hasInventor">
    <rdfs:range rdf:resource="http://www.ejemplo.com/patentes.owl#Inventor"/>
    <rdfs:domain rdf:resource="http://www.ejemplo.com/patentes.owl#PatentDocument"/>
    <owl:inverseOf rdf:resource="http://www.ejemplo.com/patentes.owl#invented"/>
</owl:ObjectProperty>
```

Figura 2.13 Documento OWL

La especificación formal del lenguaje OWL fue influenciada por el paradigma de marcos (Minsky, 1995), la lógica descriptiva y la sintaxis de RDF/XML comentada en la Sección 2.5.2 (Horrocks et al., 2003).

El paradigma de marcos agrupa información referente a una clase en un marco, lo que permite leer y entender fácilmente las ontologías. El marco incluye el nombre de la clase, la clase general a la que pertenece la clase del marco, y lista los *slots* (ya sean propiedades con sus instancias o restricciones de los valores). Adicionalmente, cada propiedad tiene su marco, y define el rango, dominio, restricciones, transitividad, etc.).

Este paradigma de marcos es usado en gran número de sistemas de representación de conocimiento, como el editor de ontologías Protégé (Noy et al., 2000) y el modelo de conocimiento OKBL (Chaudhri et al., 1998). En el ámbito de las ontologías, donde convergen expertos de varias áreas (informática, ciencias sanitarias, gestión de negocios, etc.) es fundamental facilitar la lectura, modificación y uso de las ontologías, especialmente para usuarios no familiarizados con la lógica descriptiva.

A continuación explicaremos en qué consisten las lógicas descriptivas, los distintos tipos que existen y los sublenguajes de OWL y OWL 2 basados en los distintos tipos de lógicas descriptivas.

Lógica Descriptiva

La Lógica Descriptiva, DL (*Description Logics*), que se puede considerar como el pilar que sustenta OWL, es una familia de lenguajes lógicos de representación del conocimiento basados en clases o conceptos (Baader, 2003). Están caracterizados por el uso de varios constructores para crear clases complejas a partir de clases sencillas, poniendo énfasis en la decidibilidad de los algoritmos de razonamiento, así como, en la provisión de servicios de inferencia robustos y empíricamente tratables.

La DL formaliza la representación de un dominio, definiendo los conceptos relevantes del mismo y sus relaciones. La representación de los conceptos y sus relaciones se denomina T-box o parte terminológica y la representación de las propiedades de los individuos se denomina A-box o parte de aserciones.

Por convención, en notación abstracta se utilizan las letras A y B para los conceptos atómicos, la letra R para los roles atómicos, y las letras C y D para las descripciones de conceptos. Las principales propiedades que se pueden verificar a través de la lógica descriptiva son:

- *Satisfacibilidad de un concepto*. Un concepto C es satisfacible en una terminología T si la definición de C denota un conjunto de individuos que puede no ser vacío.

- *Subsunción de conceptos*. Un concepto D subsume a otro concepto C, expresado como $C \sqsubseteq D$, si el conjunto denotado por C es un subconjunto del denotado por D.

- *Consistencia de una base de conocimiento* (conjunto formado por una terminología y las declaraciones acerca de las instancias individuales de los conceptos que define). Una base de conocimiento es consistente si las descripciones acerca de cada individuo de la A-box no violan las descripciones y axiomas del T-box.

- *Verificación de los individuos*. Verifica si cada individuo pertenece, al menos, a un concepto C.

- *Recuperación de individuos*. Encuentra todos los individuos pertenecientes a un mismo concepto.

- *Realización de un individuo*. Encuentra todos los conceptos a los cuales pertenece un individuo.

La familia de la lógica descriptiva es muy amplia. En un extremo se encuentran las lógicas mínimas con una capacidad expresiva muy pequeña, cómo $\mathcal{AL}$ (*Attribute*

Language) básico, pero cuya simplicidad hace que la complejidad computacional para verificar las propiedades descritas sea menor y no tenga problemas de decidibilidad, es decir, que los razonadores alcancen soluciones de razonamiento, verificando las propiedades arriba expuestas, en un tiempo limitado. En el otro extremo se encuentran las lógicas complejas, con mayor capacidad expresiva, pero por ello presentan problemas graves de decidibilidad, y carecen de razonadores implementados que los soporten. Estas lógicas descriptivas se representan con lenguajes que son extensiones de $\mathcal{AL}$ básico, y pertenecen a la familia $\mathcal{AL}$.

La Figura 2.14 muestra el conjunto de constructores de $\mathcal{AL}$ básico, que combinados, permiten simular gran parte de las relaciones presentes en el mundo real. La semántica es definida mediante lógica de primer orden, usando una interpretación I, que consiste en un conjunto no vacío o dominio de interpretación Δ^I, y una función de interpretación, que asigna un set $A^I \subseteq \Delta^I$ a cada concepto atómico A, y que asigna una relación binaria $R^I \subseteq \Delta^I \times \Delta^I$ a cada propiedad atómica R. Dos conceptos C y D se dice que son equivalentes $C \equiv D$ si $C^I = D^I$ para todas la interpretaciones I.

Syntax	Semantics	Comment
A	$A^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$	atomic concept
R	$R^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$	atomic role
$\top$	$\Delta^{\mathcal{I}}$	top (most general) concept
$\bot$	$\emptyset$	bottom (most specific) concept
$\neg A$	$\Delta^{\mathcal{I}} \setminus A^{\mathcal{I}}$	atomic negation
$C \sqcap D$	$C^{\mathcal{I}} \cap D^{\mathcal{I}}$	intersection
$\forall R.C$	$\{a \in \Delta^{\mathcal{I}} \mid \forall b.(a,b) \in R^{\mathcal{I}} \Rightarrow b \in C^{\mathcal{I}}\}$	value restriction
$\exists R.\top$	$\{a \in \Delta^{\mathcal{I}} \mid \exists b.(a,b) \in R^{\mathcal{I}}\}$	limited existential quantification

Figura 2.14 Lógica *AL* básica, sintáctica y semántica (Marek Obitko)[6]

[6] http://www.obitko.com

En este tipo de lógica se puede escribir, por ejemplo, $Person \sqcap \forall hasChild.Female$, definiendo a todas las personas cuyo hijos son todos mujeres.

El $\mathcal{AL}$ básico se puede extender para aumentar su capacidad expresiva, dando lugar a diferentes lenguajes de la familia $\mathcal{AL}$. Cada lenguaje de esta familia representa una extensión al lenguaje básico $\mathcal{AL}$ y su nombre se escribe añadiendo una letra al final de la cadena $\mathcal{AL}$ indicando qué semántica se puede expresar ($\mathcal{AL}$ [$\mathcal{C}$][$\mathcal{U}$][$\mathcal{R}$][...) (véase la Figura 2.15). A continuación se enumeran algunas de las letras utilizadas para denotar distintas capacidades expresivas:

- $\mathcal{C}$, negación de conceptos compleja, (o arbitraria, extendiendo la negación sobre conceptos atómicos que permite $\mathcal{AL}$; p.e., para expresar el conjunto complementario de la unión de otros conceptos).

- $\mathcal{U}$, un concepto como resultado de la unión de otros dos (p.e., la clase $Humano$ como la unión de $Hombres$ y $Mujeres$).

- $\mathcal{R}$, axiomas complejos para definir propiedades (o relaciones, roles, slots; p.e., composición de propiedades, disyunción de propiedades, propiedades reflexivas e irreflexivas, y el complemento de una relación).

- $\mathcal{O}$, clases o conceptos, a partir de conjuntos enumerados (p.e., podría declararse la clase $Día\_de\_la\_semana$ como la formada por los elementos $lunes$, $martes$, etc.).

- $\mathcal{I}$, la propiedad inversa de otra dada (p.e., una propiedad $padreDe$ puede especificarse como la inversa de la relación $hijoDe$).

- $\mathcal{Q}$, restricciones de cardinalidad *cualificadas* sobre propiedades (o relaciones, o roles), esto es, expresiones que permiten especificar la cardinalidad de una propiedad determinada y, además, la clase a la que deben pertenecer los valores que pueda tomar dicha propiedad (p.e., una persona puede tener **dos** progenitores biológicos, sin embargo, sólo tendrá **una** madre y **un** padre).

- ($\mathcal{D}$) Usa propiedades de tipos de datos, (*datatype properties*), o valores de datos.

- $\mathcal{F}$ permite propiedades funcionales.

Construct Name	Syntax	Semantics	
atomic concept	A	$A^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$	$\mathcal{S}$
universal concept	$\top$	$\top^{\mathcal{I}} = \Delta^{\mathcal{I}}$	
atomic role	R	$R^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$	
transitive role	$R \in \mathbf{R}_+$	$R^{\mathcal{I}} = (R^{\mathcal{I}})^+$	
conjunction	$C \sqcap D$	$C^{\mathcal{I}} \cap D^{\mathcal{I}}$	
disjunction	$C \sqcup D$	$C^{\mathcal{I}} \cup D^{\mathcal{I}}$	
negation	$\neg C$	$\Delta^{\mathcal{I}} \setminus C^{\mathcal{I}}$	
exists restriction	$\exists R.C$	$\{x \mid \exists y.\langle x,y\rangle \in R^{\mathcal{I}} \text{ and } y \in C^{\mathcal{I}}\}$	
value restriction	$\forall R.C$	$\{x \mid \forall y.\langle x,y\rangle \in R^{\mathcal{I}} \text{ implies } y \in C^{\mathcal{I}}\}$	
role hierarchy	$R \sqsubseteq S$	$R^{\mathcal{I}} \subseteq S^{\mathcal{I}}$	$\mathcal{H}$
inverse role	R^-	$\{\langle x,y\rangle \mid \langle y,x\rangle \in R^{\mathcal{I}}\}$	$\mathcal{I}$
number restrictions	$\geqslant nR$ $\leqslant nR$	$\{x \mid \sharp\{y.\langle x,y\rangle \in R^{\mathcal{I}}\} \geqslant n\}$ $\{x \mid \sharp\{y.\langle x,y\rangle \in R^{\mathcal{I}}\} \leqslant n\}$	$\mathcal{N}$
qualifying number restrictions	$\geqslant nR.C$ $\leqslant nR.C$	$\{x \mid \sharp\{y.\langle x,y\rangle \in R^{\mathcal{I}} \text{ and } y \in C^{\mathcal{I}}\} \geqslant n\}$ $\{x \mid \sharp\{y.\langle x,y\rangle \in R^{\mathcal{I}} \text{ and } y \in C^{\mathcal{I}}\} \leqslant n\}$	$\mathcal{Q}$

Figura 2.15 Lógica AL, sintáctica y semántica (Horrocks et al., 2000)

De esta forma, el lenguaje $\mathcal{ALCI}$, es un lenguaje de descripción que permite expresar la negación de un concepto cualquiera (por ejemplo, $\neg\forall R.C$), y también definir una propiedad como la inversa de otra. Se suele utilizar la abreviatura $\mathcal{S}$ para hacer referencia al lenguaje $\mathcal{ALC}_{\mathcal{R}+}$, esto es, una extensión al lenguaje $\mathcal{ALC}$ para expresar roles transitivos (denotado como $_{\mathcal{R}+}$).

Sublenguajes de OWL y OWL 2

En base a la expresividad de la lógica descriptiva utilizada para expresar una ontología en OWL se distinguen tres sublenguages de OWL:

- **OWL Lite.** Soporta la lógica descriptiva $\mathcal{SHIF}(\mathcal{D})$, es decir, una extensión de $\mathcal{AAL}$ con soporte para expresar roles transitivos ($\mathcal{S}$), jerarquías de roles ($\mathcal{H}$), propiedad inversa ($\mathcal{I}$), propiedades funcionales ($\mathcal{F}$), y tipos de datos ($\mathcal{D}$). Tiene un poder expresivo pequeño, permite especificar jerarquías de clasificación y restricciones simples. Es un lenguaje útil para definir tesauros o taxonomías.

- **OWL DL**. Soporta lógica descriptiva $\mathcal{SHOIN}(\mathcal{D})$, es decir, una extensión de $\mathcal{AAL}$ con soporte para expresar roles transitivos ($\mathcal{S}$), jerarquías de roles ($\mathcal{H}$), clases a partir de conceptos enumerados ($\mathcal{O}$), propiedad inversa ($\mathcal{I}$), restricciones de cardinalidad ($\mathcal{N}$), y tipos de datos ($\mathcal{D}$). Permite la máxima potencia expresiva manteniendo la decidibilidad de los procesos de razonamiento. Es decir, garantiza que todas las conclusiones sean computables y todos los cálculos de los procesos de deducción automática que pueda llevar a cabo un razonador diseñado para este lenguaje, terminarán en un tiempo finito. OWL DL debe su nombre a las siglas *Description Logics*, debido a que soporta todos los constructores de las lógicas más complejas de esta familia.

- **OWL Full**. Permite la máxima potencia expresiva y la libertad sintáctica de RDF. Por ejemplo, en OWL Full se puede definir una clase como un conjunto de individuos, y al mismo tiempo como una propiedad. Sin embargo, carece de decidibilidad de los procesos de razonamiento, es decir, no existen razonadores que soporten OWL Full, y que puedan automáticamente validar las propiedades de la lógica descriptiva (Satisfacibilidad, subsunción, etc.).

La nueva versión del lenguaje OWL 2 que soporta lógica descriptiva $\mathcal{SROIQ}$ surgió gracias a los avances en lógica descriptiva (Grau et al., 2008). OWL 2 (Motik et al., 2009) incorpora nuevas funcionalidades y constructores con mayor potencia expresiva que permiten expresar entidades, tales como, cadenas de propiedades, propiedades asimétricas, reflexivas y disjuntas, más tipos de datos, etc.

Los perfiles o sublenguajes de OWL 2, constan de OWL 2 Full (similar a OWL 1 Full, pero con las ventajas de OWL 2 mencionadas anteriormente), y OWL DL. Dentro de OWL DL se diferencian tres sublenguajes, que son una versión reducida de OWL 2, perdiendo poder expresivo en favor de eficiencia de razonamiento (Motik et al., 2009). Los sublenguajes son independientes entre sí, y cada uno se utiliza para un fin concreto:

- **OWL 2 EL**: El acrónimo $\mathcal{EL}$++ de lógica descriptiva, expresa que la lógica subyacente permite la conjunción y las restricciones existenciales. OWL 2 EL proporciona solo cuantificación existencial (o existencia de al menos un individuo en cada clase). Este lenguaje es adecuado para aplicaciones donde se necesitan ontologías de gran tamaño, con un alto número de clases y propiedades.

- **OWL 2 QL:** El acrónimo QL hace referencia a Query Language, o lenguaje de consultas. Este lenguaje es especialmente eficiente para ontologías con un gran número de instancias, que requieren un elevado número de consultas a los datos.

- **OWL 2 RL:** El acrónimo RL hace referencia a *Rule Language*, o lenguaje de reglas. Este lenguaje es especialmente eficiente para ontologías que requieren cierto razonamiento sin perder mucho poder expresivo. El razonamiento en este tipo de ontologías se basa en reglas.

2.6 Recuperación de Información

Para poder gestionar la información de una manera adecuada, es necesario el descubrimiento de conocimiento en bases de datos KDD (*Knowledge Databases Discovery*), que se define como el proceso de extracción no trivial de patrones que sean válidos, novedosos, potencialmente útiles y entendibles a partir de datos contenidos en bases de datos (Klösgen & Zytkow, 2002; Piatetsky-Shapiro & Frawley, 1992).

El proceso de KDD consiste, tradicionalmente, en usar algoritmos de minería de datos (*data mining*), para identificar conocimiento útil que no está explícitamente representado en las bases de datos. La minería de datos permite transformar información disponible en bases de datos en conocimiento útil. Por ejemplo, a partir de los documentos de patentes, aplicando algoritmos de minería de datos que busquen texto con expresiones relacionadas con el medio ambiente, recuperaremos aquellos documentos de patentes que representen innovación medioambiental.

Hoy en día, existe un creciente interés por los sistemas de servicios Web como métodos de recuperación de información. En estos sistemas se comparte información a través de Internet, y se descubren servicios a los cuales un usuario se puede subscribir. A continuación explicaremos en qué consiste la minería de datos y los servicios Web.

2.6.1 Minería de Datos

Al ir creciendo los datos en los sistemas de información, la minería de datos recibe su nombre de un símil que compara la cantidad de datos existentes con montañas de datos, y al igual que la montaña se debe excavar, o ejecutar minería sobre ella para obtener las pepitas preciosas, la minería de datos excava entre montañas de datos para encontrar la información útil (conocimiento) en cada dominio (Fayyad et al., 1996). Este conocimiento debe ser interesante, entendiendo como tal el conocimiento novedoso y útil para el usuario (Bayardo & Agrawal, 1999).

La minería de datos es una metodología de análisis de datos tradicional enriquecida con técnicas avanzadas de descubrimiento de patrones, a priori desconocidos o no explícitos, en las bases de datos. Al contrario que los sistemas de gestión de información que se basan en indicadores de gestión o en información agregada, la minería de datos se basa en información de detalle contenida en la base de datos. Se considera una tecnología multidisciplinar, ya que hace uso de paradigmas de computo de diferentes áreas, cómo los árboles de decisión, la inducción de reglas, las redes neuronales artificiales, el aprendizaje basado en instancias, la estadística bayesiana o la programación lógica (Witten & Frank, 1999).

La minería de datos es un conjunto de tecnologías dentro del proceso de descubrimiento o extracción de conocimiento (KDD). Las etapas de las que consta un proceso KDD son:

- Pre-procesamiento: Consiste en la recogida y limpieza de datos, discretización, selección de atributos e integración de datos.

- Minería de datos: En esta etapa se seleccionan los algoritmos de minería de datos y se aplican a los datos extraídos en el pre-procesamiento.

- Post-procesamiento: En esta etapa se lleva a cabo la interpretación y evaluación de los resultados obtenidos y la utilización del conocimiento descubierto.

En la minería de datos se crean modelos (también denominados clasificadores) donde los individuos cumplen una o un conjunto de reglas con datos de entrenamiento. Posteriormente se comprueba el modelo con una función de ajuste, para que dé cabida al mayor número de individuos posible con otros datos distintos a los de entrenamiento. Finalmente se ejecuta un proceso de selección de las mejores reglas, o las que se cumplen en un mayor porcentaje de casos.

El principal problema de la minería de datos radica en que las reglas no sólo deben ser exactas y compresibles, sino que, además, deben ser interesantes, es decir deben ser útiles para el fin que se desea alcanzar. Existen dos métodos para la selección de reglas interesantes (Freitas, 2003): los métodos subjetivos y los objetivos. En los métodos subjetivos el usuario indica los atributos de las reglas para que estas sean interesantes. En los métodos objetivos se comparan las reglas según alguna métrica (Freitas, 1999), para ver qué reglas son más interesantes.

Dentro de la minería de datos, hay un área disciplinar denominada minería Web (*Web data minig*), que se utiliza en la Web semántica para crear ontologías a partir de datos dispersos por la Web. La minería Web consiste en la aplicación de técnicas de minería de datos (Klösgen & Zytkow, 2002; Michalski et al, 1998) sobre información procedente de aplicaciones Web (Pal et al., 2002).

2.6.2 Servicios Web

Un servicio Web, es un sistema software creado para la interacción entre máquinas a través de redes, mediante el intercambio de mensajes. Posee una interfaz descrita en un formato inteligible por máquinas (generalmente WSDL, *Web Service Description Language*). Por debajo, generalmente subyace la comunicación a través de mensajes SOAP (*Service Oriented Arquitecture Protocol*), típicamente usando HTML con una serialización XML, y otros estándares Web (W3C, 2004) (Figura 2.16).

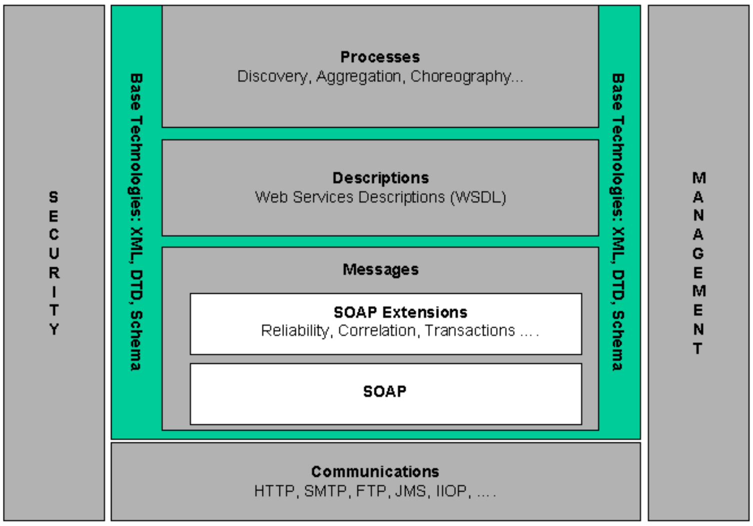

Figura 2.16 Tecnologías para los servicios Web (W3C, 2004)

Estos sistemas recuperan información diseminada por la Web a través de servicios Web. Los pasos que siguen estos sistemas para poder comunicarse entre sí son los siguientes (Figura 2.17):

1. El proveedor y el cliente se hacen accesibles mutuamente.

    a. El servicio de descubrimiento obtiene la descripción de un servicio Web (WDS, *Web Description Service*), y su descripción funcional (FD, *Functional Description*), que describe el servicio ofrecido por el proveedor.

b. El cliente le indica al servicio de descubrimiento los criterios para la selección del proveedor como, servicio ofrecido, calidad del servicio, etc.

c. El servicio de descubrimiento devuelve las descripciones de los servicios encontrados que cumplen los criterios requeridos. Si se encuentran varios servicios que cumplen los criterios, el cliente selecciona uno de ellos.

2. El cliente y el proveedor negocian la semántica (paso 2, en la Figura 3.17) que utilizarán en las comunicaciones.

3. La descripción del servicio y la semántica utilizada en las comunicaciones se guarda y utiliza correctamente por ambos agentes, el cliente y el proveedor.

4. El cliente y el proveedor intercambian mensajes SOAP.

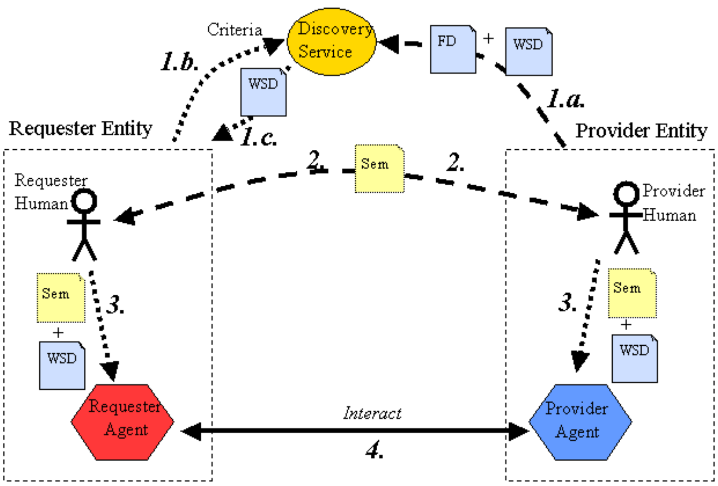

Figura 2.17 Funcionamiento de los servicios Web (W3C, 2004)

2.7 Inferencia de Conocimiento

El razonamiento, o inferencia del conocimiento, en ontologías o bases de conocimiento se conoce como la acción de desvelar hechos que no están explícitamente expresados en la base de conocimiento (Obitko, 2007). Para conseguir este razonamiento, las ontologías deben estar basadas en especificaciones formales, como la lógica descriptiva.

Los lenguajes de ontologías, como OWL, están basados en lógica descriptiva (Baader, 2003), como ya hemos comentado en la Sección 2.4.2.3. Por este motivo, los razonadores basados en OWL, como Pellet (Sirin et al., Parsia, 2007), Hermit (Shearer & Motik, 2008), o los más recientes ELK (Kazakov et al., 2012) nos permiten realizar ciertos procedimientos de verificación o razonamiento, tales como, comprobación de la consistencia, satisfacibilidad de los conceptos, equivalencia de clases y clasificación (Antoniou & Harmelen, 2009), que son las principales propiedades que se pueden verificar a través de la lógica descriptiva, comentadas en la Sección 3.4.2.3. Siendo la clasificación ontológica (o computo de las jerarquías de subsunción para clases y propiedades) el servicio de razonamiento principal proporcionado por los razonadores (Glimm et al., 2012). En los capítulos posteriores de esta tesis explotaremos las ventajas de esta clasificación a través de razonadores para analizar la internacionalización de la cartera de patentes de una empresa, o la pertenencia de una patente a una categoría de tecnología medioambiental.

Los algoritmos utilizados en los razonadores para inferir las relaciones de subsunción se pueden dividir en dos grupos; estructurales y lógicos (Obitko, 2007). Los algoritmos estructurales se basan en la comparación de la estructura sintáctica normalizada de dos conceptos. Estos algoritmos son robustos, pero tienen problemas de completitud, especialmente cuando se usan ontologías muy expresivas. Por este motivo, los algoritmos lógicos son los usados casi exclusivamente en la implementación de razonadores. Los algoritmos lógicos verifican la relación de subsunción $C \sqsubseteq D$, comprobando que $C \sqcap \neg D$ es no satisfacible, es decir, no existe ningún elemento de C

que no pertenezca a D. Hay diferentes propuestas para llevar a cabo dicha verificación, pero la más utilizada son los algoritmos de tablero (*tableau algorithms*).

Los algoritmos de tablero tratan de probar la Satisfacibilidad de un concepto C, construyendo un modelo, una interpretación I, en la cual D^I sea no vacía. Un tablero es un grafo que representa dicho modelo. En el grafo los nodos corresponden a individuos y los arcos a relaciones entre individuos. El algoritmo comienza con un individuo que satisfice un concepto (pertenece a una clase), y aplica reglas expansivas, comprobando si satisface otros conceptos hasta que no encuentra más inferencias, o encuentra una contradicción. Las reglas expansivas dependen de la lógica descriptiva empleada, y son formas equivalentes de expresar la misma relación lógica, por ejemplo, $\neg(C1 \sqcap C2) \Rightarrow \neg C1 \sqcup \neg C2$. Para ampliar la información sobre razonadores lógicos, véase (Baader, 2003). Los razonadores utilizados en esta tesis están basados en algoritmos de tableros.

2.7 Gestión de la Información de Patentes

Como ya se comentó en el capítulo anterior, una patente es un derecho legal que autoriza a su poseedor a excluir a terceros de actos de fabricación, uso, comercialización, oferta para la venta, venta, o importación para estos fines del producto, o proceso, objeto de la patente (TRIPS –*"Trade Related aspects of Intellectual Property Rights-Agreement",* art. 28)[7].

Los documentos de patentes están compuestos de una primera página con datos bibliográficos, como inventor, solicitante, título de la patente, resumen, etc. (Figura 2.18), el resto del documento se compone de la descripción de la patente, figuras, y reivindicaciones.

[7] http://www.wipo.int/treaties/en/agreement/trips.html

US007251458B2

(12) **United States Patent**
O'Connell

(10) **Patent No.: US 7,251,458 B2**
(45) **Date of Patent: Jul. 31, 2007**

(54) **SYSTEMS AND METHODS FOR RECYCLING OF CELL PHONES AT THE END OF LIFE**

(75) Inventor: **Donal O'Connell**, Irving, TX (US)

(73) Assignee: **Nokia Corporation**, Espoo (FI)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 295 days.

(21) Appl. No.: **10/444,485**

(22) Filed: **May 23, 2003**

(65) **Prior Publication Data**

US 2004/0235513 A1 Nov. 25, 2004

(51) **Int. Cl.**
H04B 1/38 (2006.01)

(52) **U.S. Cl.** **455/90.1**; 455/575.1; 455/418

(58) **Field of Classification Search** 455/575.1, 455/419, 563, 90.1, 418, 575
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,365,570	A *	11/1994	Boubelik	455/565
5,875,393	A *	2/1999	Altschul et al.	455/407
5,946,613	A *	8/1999	Hayes et al.	455/406
5,966,654	A	10/1999	Croughwell et al.	
6,233,464	B1 *	5/2001	Chmaytelli	455/556.2
6,463,263	B1 *	10/2002	Feilner et al.	455/90.1
6,584,327	B1 *	6/2003	Nilsson	455/563
2002/0065868	A1	5/2002	Moore et al.	
2002/0137476	A1 *	9/2002	Shin	455/90
2003/0087679	A1	5/2003	Naka et al.	
2003/0134629	A1	7/2003	Jacobs et al.	
2005/0009509	A1 *	1/2005	Miscopein et al.	455/418

FOREIGN PATENT DOCUMENTS

EP	1 128 692 A		8/2001
GB	2 378 854 A		2/2003
WO	WO 03/030381 A1		4/2003
WO	03/043303	*	5/2003

OTHER PUBLICATIONS

International Search Report, EPO, mailed Oct. 14, 2004.
European Search Report, mailed Sep. 9, 2004.

* cited by examiner

Primary Examiner—Lester G. Kincaid
Assistant Examiner—Fred Casca
(74) *Attorney, Agent, or Firm*—Alston & Bird LLP

(57) **ABSTRACT**

The present invention provides systems and methods for transforming cell phones at end of life into new electronic devices, such as clocks, calculators, PDAs, remote controls, etc. Importantly, the systems and methods of the present invention salvage many of the features and electronics of the cell phone that are not associated with the placement and reception of telephone calls. The electronics of the cell phone are repackaged and/or configured to hide the telecommunications options associated with the phone, leaving intact or upgrading other features of the phone, such as clock, calculator, storage device, etc. In doing so, all or most of the components of the cell phone are successfully recycled over more conventional methods that scavenge the phones for precious metals and reusable components.

26 Claims, 4 Drawing Sheets

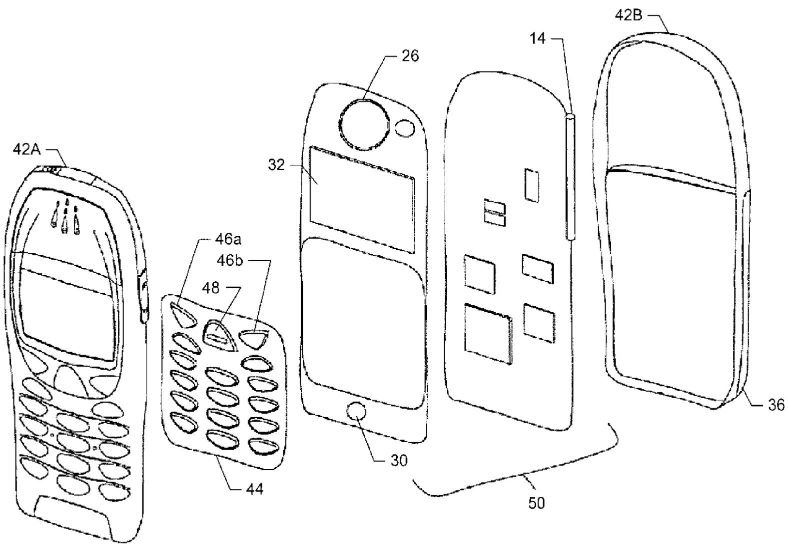

Figura 2.18 Ejemplo de un documento de patente (O'Connell, 2007)

2.7.1 Bases de Datos de Patentes

Las patentes son concedidas a nivel nacional. Cada país dispone de su oficina de patentes y sus leyes específicas sobre patentes vigentes en dicho país. Sin embargo, existe un creciente interés en homogeneizar el sistema de patentes a nivel mundial, y también existen dos oficinas que agrupan a varios países, como la Oficina Europea de Patentes (EPO), que comprende los países integrantes de la Unión Europea, y la organización de la propiedad intelectual a nivel mundial (WIPO) que actualmente aglutina a 185 países. Estas oficinas facilitan el proceso de presentación de patentes en varios países, aunque no regulan las leyes nacionales.

Cada oficina de patentes dispone de una base de datos con información sobre sus patentes, y en muchas ocasiones también recoge información sobre patentes de otros países. Las grandes oficinas de patentes, como la WIPO, EPO, USPTO (Oficina de patentes y marcas de los Estados Unidos), etc. han abierto sus bases de datos para ser consultadas online.

Por ejemplo, la EPO publica semanalmente unos 4.000 nuevos documentos de patentes en formato XML con imágenes TIFF Fax G4 y también en formato PDF. También cuenta con más de 3.3 millones de documentos de patentes, que datan desde 1978. La EPO también ofrece acceso online a la base de datos GIP (*Global Intellectual Property*), que contiene más de 80 millones de documentos de patentes a nivel mundial, y que crece a un ritmo de 2 millones de documentos anuales. La EPO ofrece dos interfaces Web de acceso a su base de datos: *Spacenet*, que es gratuito y *Patent Information Services for Expert*, que es de pago.

La versión XML de un documento de patentes en la EPO está basada en WIPO ST.36[8], que es la recomendación publicada por WIPO para crear documentos XML de patentes estándar a nivel mundial. Esta definición es ampliamente aceptada y usada mundialmente para la publicación de documentos de patentes. ST.36 representa en XML los códigos numéricos ST.9 que representan los distintos metadatos o datos bibliográficos del documento de patentes. Por ejemplo, en la Figura 3.18, los números entre paréntesis delante de los distintos campos del documento de patente representan los códigos ST.9. Sin embargo, muchas bases de datos no ofrecen los resultados XML de las búsquedas con esta codificación, sino con caracteres representando palabras, como *TIEN* para representar el título en inglés (*TItle ENglish*). ST.36 define cerca de 1000 elementos de metadatos, dónde más de la mitad son datos bibliográficos, contenidos en la primera página de un documento de patentes.

Actualmente las empresas están aumentando exponencialmente el volumen anual de innovaciones que son patentadas. La información sobre tales patentes está almacenada en distintas bases de datos, que presentan diferencias en cuanto a la representación formal de la información y carecen de semántica. Estas bases de datos tienen limitaciones en lo que respecta a las relaciones entre términos y conceptos que utilizan en su descripción, no permitiendo el descubrimiento de conocimiento implícito. En este sentido, la representación de la información acerca de patentes mediante ontologías puede ayudar a descubrir de forma más eficiente el conocimiento implícito en los documentos de patentes, y analizar mejor la cartera de patentes, para poder gestionarlas apropiadamente.

[8] http://www.wipo.int/standards/en/pdf/03-36-01.pdf

2.7.2 Acceso a Información sobre Patentes

Además del acceso online a las bases de datos, las oficinas de patentes están ofreciendo otros servicios de acceso a los datos que contienen los documentos de patentes.

A modo de ejemplo, la EPO dispone de un servicio llamado *Open Access Service* OPS, que es un servicio Web que permite el acceso a los datos de patentes a través de una interfaz estandarizada XML. Los datos pueden obtenerse de las mismas bases de datos que Spacenet, que accede a EPODOC (base de datos de información bibliográfica), *Worldwide Legal Status Database*/PRS (base de datos de estados legales por los cuales han pasado las patentes), EPOQUE (base de datos de todo el texto de las patentes), BNS (base de datos de las imágenes contenidas en los documentos de patentes). La utilización del servicio OPS por parte de una empresa, requiere la creación de un cliente Web que acceda a los servicios ofrecidos por OPS, y que vuelque los datos recuperados de las bases de datos sobre una aplicación en el lado del cliente.

Esta amplia gama de servicios hace que sea accesible nueva información sobre patentes, como por ejemplo estadísticas a nivel sectorial, estadísticas a nivel país etc. Además, posibilitan la recuperación de esta información a través de distintas aplicaciones y formatos. Incluso permiten crear aplicaciones que se conecten automáticamente a las bases de datos de patentes y vuelquen la información recuperada en una ontología o cualquier otro tipo de representación de información.

2.7.3 Ontologías de Patentes

El análisis de patentes tradicional requiere un coste significativo de tiempo y trabajo (Trappey et al., 2012), sin embargo, el uso de las ontologías ha simplificado bastante el análisis de los datos contenidos en grandes repertorios de patentes. Adicionalmente, las ontologías permiten la integración de diferentes bases de datos con estructuras de datos heterogéneas mediante una representación semántica común.

Existen dos corrientes principales de investigación en el uso de ontologías para analizar datos acerca de patentes. La primera corriente hace uso de algoritmos de inteligencia artificial para realizar minería de datos sobre los textos de las patentes. Estos algoritmos analizan los datos de los documentos de patentes y crean un modelo que se deriva de dicho análisis, es decir, el modelo se va creando en función de los algoritmos empleados y cualquier cambio en el modelo supone un cambio en las aplicaciones que hacen uso de dicha ontología.

La segunda corriente sigue un enfoque declarativo, que hace uso de técnicas basadas en declaraciones de conceptos. En esta vertiente se utilizan los metadatos de las patentes definidos en la ST.36 para crear ontologías. Los modelos declarativos son propuestos a priori y cualquier cambio posterior requiere únicamente el cambio en las ontologías, manteniéndose las herramientas software y aplicaciones estables.

En la primera corriente se encuentran las ontologías que hacen uso de la frecuencia de aparición de ciertos términos o palabras en los documentos de patentes, para analizar las tendencias tecnológicas o otra información útil (Kim et al., 2008; Liu & Hsu, 2010; Trappey et al., 2009). Dentro de esta corriente existen patentes que están implementadas en OWL (Choi et al., 2012; Seol et al., 2011; Soo et al., 2006). Estas ontologías mayoritariamente se crean para buscar nuevos campos tecnológicos basados en los términos encontrados en el texto de las patentes, pero en este enfoque no se contempla el uso de los metadatos de patentes, cómo pueden ser el inventor, el solicitante, etc., que son campos útiles para los análisis de patentes por parte de las empresas, para mejorar la gestión de la cartera de patentes.

En el segundo grupo las ontologías son modelos estructurados de conceptos que representan los metadatos de las patentes. Las ontologías de patentes más relevantes basadas en metadatos y representadas en OWL son Patexpert (Giereth et al., 2006), que es una ontología creada bajo el amparo del proyecto Europeo Patexpert (Wanner et al., 2009), y PatentOntology, creada en la universidad de Stanford (Taduri et al., 2011).

Patexpert es una ontología extensa desarrollada a partir de una ontología de alto nivel *Suggested Upper Merge Ontology* (SUMO). Patexpert está compuesta por varias ontologías más pequeñas, cómo PMO (*Patent Metadata Ontology*) que representa todos los metadatos utilizados en la mayoría de las bases de datos de patentes. Dentro del proyecto, se desarrollaron paralelamente varias aplicaciones que hacían uso de las ontologías Patexpert. Desgraciadamente las ontologías que están a disposición del público[9] no están pobladas, y se limitan a dar una estructura global para describir la información de los metadatos de patentes, sin incluir ninguna relación adicional o ningún axioma que dé lugar a inferir conocimiento a través de los razonadores.

PatentOntology es una ontología más reciente creada explícitamente para integrar dos dominios heterogéneos que contienen información sobre patentes de la USPTO: la base de datos de documentos de patentes y la base de datos de información sobre casos de juicios de patentes (Taduri et al., 2011). Esta combinación de información permite a las empresas conocer no solo las patentes existentes en un área tecnológica, sino también la validez de dichas patentes, es decir, si dichas patentes están activas, o han sido invalidadas por un tribunal de patentes. Esta ontología permite

[9] http://mklab.iti.gr/project/patexpert

ahorrar tiempo a las empresas ya que toda la información se agrupa en una única base de conocimiento.

En esta tesis se propondrá en el Capítulo 5 una ontología de patentes basada en metadatos que complementa las ontologías existentes actualmente basadas en metadatos, Patexpert y PatentOntology. Particularmente, se propondrá una metodología para hacer explicita información contenida dentro de uno de los metadatos de patentes que representa los códigos tecnológicos, y también se propondrá en el Capítulo 6 un método para inferir nueva información y conocimiento a partir de los metadatos de patentes mediante razonadores lógicos.

Referencias del Capítulo

Alavi, M., & Leidner, D. E. (2001). Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues. *MIS Quarterly*, *25*(1), 107.

Alberts L.K. (1993). *YMIR: An Ontology for Engineering Design*. PhD Tesis. University of Twente.

Alexiev, V., Breu, M., & De Bruijn, J. (2005). *Information Integration with Ontologies : Experiences from an Industrial Showcase*. Wiley.

Ankolenkar, A., Burstein, M., Hobbs, J., Lassila, O., Martin, D., McIlraith, S., … Zeng, H. (2001). DAML-S: A Semantic Markup Language For Web Services. *Proceedings of the First Semantic Web Working Symposium SWWS01 Stanford University*, 411–430.

Antoniou, G., & Harmelen, F. Van. (2009). Web Ontology Language: OWL. *Handbook on ontologies*, 91–110. Springer.

Wielinga, B.J., & Schreiber, A.T. (1994). Reusable and shareable Knowledge Bases: A European Perspective. *Knowledge Building and Knowledge Sharing*, 110–120. Ohmsha Ltd and IOS Press.

Baader, F. (2003). *The Description Logic Handbook: Theory, Implementation and Applications*. Cambridge University Press.

Bayardo, R. J., & Agrawal, R. (1999). Mining the Most Interesting Rules. *International Conference on Knowledge Discovery and Data Mining - KDD ’99*, 145–154.

Bermúdez-Edo, M., Hurtado-Torres, N., & Cordón-Pozo, E. (2008). The Influence of the Faculty Attitude on the Adoption of ICTs’ Innovations for Learning Purposes: Analysis of a Shared Web. *Times of Convergence. Technologies Across Learning Contexts* (5192), 58–61.

Berners-Lee, B. T., Hendler, J., & Lassila, O. (2001). The Semantic Web. *Scientific American*, *284*(5), 34–43.

Borst, W. N. (1997). *Construction of Engineering Ontologies for Knowledge Sharing and Reuse*. PhD Tesis Universiteit Twente.

Brickley, D., & Guha, R. V. (2004). *RDF Vocabulary Description Language 1.0: RDF Schema*. W3C Recommendation.

Cardoso, J. (2007). The Semantic Web Vision: Where Are We? *IEEE Intelligent Systems*, *22*(5), 84–88.

Chandrasekaran, B., Josephson, J. R., & Benjamins, V. R. (1999). What are Ontologies, and Why do We Need Them? *IEEE Intelligent Systems*, *14*(1), 20–26.

Chaudhri, V. K., Farquhar, A., Fikes, R., Karp, P. D., & Rice, J. P. (1998). OKBC: A Programmatic Foundation for Knowledge Base Interoperability. *National Conference on Artificial Intelligence AAAI98*, 600–607.

Chen, D. Q., Mocker, M., Preston, D. S., & Teubner, A. (2010). Information Systems Strategy: Reconceptualization, Measurement, and Implications. *MIS Quarterly*, *34*(2), 233–259.

Chen, P. P.-S. (1976). The Entity-Relationship Unified View of Data Model-Toward a. *ACM Transactions on Database Systems*, *1*(1), 9–36.

Choi, S., Kang, D., Lim, J., & Kim, K. (2012). A fact-oriented Ontological Approach to SAO-based Function Modeling of Patents for Implementing Function-based Technology Database. *Expert Systems with Applications*, *39*(10), 9129–9140.

Codd, E. F. (1970). A relational Model of Data for Large Shared Data Banks. *Communications of the ACM*, *13*(6), 377–387.

Dahlem, N., & Hahn, A. (2009). User-Friendly Ontology Creation Methodologies - A Survey. *Americas Conference on Information Systems*.

Davenport, T. H., & Prusak, L. (1997). *Information Ecology: Mastering the Information and Knowledge Environment*. Oxford University Press.

Davenport, T. H., Prusak, L., & Prusak, L. (1997). *Working Knowledge: How Organizations Manage What They Know*. Harvard Business Press.

De-Miguel, A., Piattini, M., & Marcos, E. (1999). *Diseño de Bases de Datos Relacionales*. RA-MA.

Fallside, D. C., & Walmsley, P. (2004). *XML Schema Part 0: Primer Second Edition*. W3C Recommendation.

Fayyad, U., Piatetsky-Shapiro, G., Smyth, P., & Uthurusamy, R. (1996). *Advances in Knowledge Discovery and Data Mining*. The MIT Press.

Fensel, D., Horrocks, I., Harmelen, F. Van, Decker, S., Erdmann, M., & Klein, M. (2000). OIL in a Nutshell. *Knowledge Engineering Knowledge Management Methods, Models, and Tools*, *1937*(1937), 1–16.

Freitas, A A. (2003). A survey of Evolutionary Algorithms for Data Mining and Knowledge Discovery. *Advances in Evolutionary Computation* (136), 819–845.

Freitas, A.A. (1999). On Rule Interestingness Measures. *Knowledge-Based Systems*, *12*(5-6), 309–315.

Garcia-Molina, H., Ullman, J. D., & Widom, J. (2009). *Database Systems: The Complete Book*. Prentice Hall.

Giereth, M., St, A., Rotard, M., & Ertl, T. (2006). Application of Semantic Technologies for Representing Patent Metadata. *1st International Workshop on Applications of Semantic Technologies AST06*, *94*, 297–305.

Glimm, B., Horrocks, I., Motik, B., Shearer, R., & Stoilos, G. (2012). A Novel Approach to Ontology Classification. *Web Semantics: Science, Services and Agents on the World Wide Web*.

Goclenius, R. (1613). *Lexicon Philosophicum, Quo Tanquam Clave Philosophiae Fores Aperiuntur*. Vidua Matthiae Beckeri.

Grau, B. C., Horrocks, I., Motik, B., Parsia, B., Patel-Schneider, P., & Sattler, U. (2008). OWL 2: The Next Step for OWL. *Web Semantics: Science, Services and Agents on the World Wide Web*, *6*(4), 309–322.

Gruber, T. R. (1993). Toward Principles for the Design of Ontologies Used for Knowledge Sharing. *International Journal HumanComputer Studies*, *43*(5/6), 907–928.

Grüninger, M., Atefi, K., & Fox, M. S. (2000). Ontologies to Support Process Integration in Enterprise Engineering. *Computational Mathematical Organization Theory*, *6*(4), 381–394.

Guarino, N. (1998). Formal Ontology in Information Systems. *First International Conference FOIS'98*, 3–15.

Guarino, Nicola. (1997). Understanding, Building and Using Ontologies: A Commentary to "Using Explicit Ontologies in KBS Development". *International Journal of Human and Computer Studies*, *46*, 293–310.

Guarino, Nicola, & Giaretta, P. (1995). Ontologies and Knowledge Bases: Towards a Terminological Clarification. *Towards Very Large Knowledge Bases Knowledge Building and Knowledge Sharing* (1), 25–32.

Gómez-Pérez, A., Fernández-López, M., & Corcho, O. (2004). *Ontological Engineering: with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web*. Springer.

Haase, P., Lewen, H., & Studer, R. (2008). The NeOn Ontology Engineering Toolkit. *World Wide Web*.

Hayes, P. (2004). *RDF Semantics*. W3C Recommendation.

Horrocks, I, Patel-Schneider, P. F., & Harmelen, F. van. (2003). From SHIQ and RDF to OWL: The Making of a Web Ontology Language. *Web Semantics Science Services and Agents on the World Wide Web*, *1*(1), 7–26.

Horrocks, Ian. (2002). DAML+OIL: a Description Logic for the Semantic Web. *IEEE Data Engineering Bulletin*, *25*(1), 4–9.

Horrocks, Ian, Sattler, U., & Tobies, S. (2000). Practical Reasoning for Expressive Description Logics. *Logic Journal of the IGPL*, *8*(1705), 161–180.

Jennex, M. E., Smolnik, S., & Croasdell, D. (2012). Where to Look for Knowledge Management Success. *International Conference on System Sciences,* 3969–3978.

Kalyanpur, A., Parsia, B., Sirin, E., Grau, B., & Hendler, J. (2006). Swoop: A Web Ontology Editing Browser. *Web Semantics Science Services and Agents on the World Wide Web*, *4*(2), 144–153.

Karlinsky-Shichor, Y., & Zviran., M. (2012). Factors Influencing Perceived Benefits and User Satisfaction in Knowledge Management Systems. *Information Systems (ILAIS) Conference*.

Kazakov, Y., Krötzsch, M., & Simančík, F. (2012). ELK: A Reasoner: Architecture and Evaluation. *OWL Reasoner Evaluation Workshop*.

Kim, Y. G., Suh, J. H., & Park, S. C. (2008). Visualization of Patent Analysis for Emerging Technology. *Expert Systems with Applications*, *34*(3), 1804–1812.

Klösgen, W., & Zytkow, J. M. (2002). *Handbook of data mining and knowledge discovery*. Oxford University Press.

Lassila, O., & McGuinness, D. (2001). The Role of Frame-Based Representation on the Semantic Web. *Linköping Electronic Articles in Computer and Information Science*, *6*(5).

Malhorta, Y. (2001). *From information management to knowledge management: beyond the "Hi-Tech Hidebound" systems*. Knowledge Management And Business Model Innovation, 115-134. IGI Publishing.

Malhotra, Y. (2005). Integrating Knowledge Management Technologies in Organizational Business Processes: Getting Real Time Enterprises to Deliver Real Business Performance. *Journal of Knowledge Management*, *9*(1), 7–28.

Malhotra, Y., & Galletta, D. F. (2003). Role of Commitment and Motivation in Knowledge Management Systems Implementation: Theory, Conceptualization, and Measurement of Antecedents of Success. *36th Annual Hawaii International Conference on System,* 1–10.

Manola, F., & Miller, E. (2004). *RDF Primer*. W3C Recommendation.

Melville, N. P. (2010). Information Systems Innovation for Environmental Sustainability. *MIS Quarterly*, *34*(1), 1–21.

Michalski, R. S., Bratko, I., & Bratko, A. (1998). *Machine Learning and Data Mining; Methods and Applications*. Morgan Kaufmann.

Minsky, M. (1995). A Framework for Representing Knowledge. *The Phycology Of Computer Vision*, 163–189.

Motik, B., Grau, B. C., Horrocks, I., Wu, Z., Fokoue, A., & Lutz, C. (2009). *OWL 2 Web Ontology Language Profiles*. W3C Recommendation.

Motik, Boris, Patel-Schneider, P. F., & Parsia, B. (2009). *OWL 2 Web Ontology Language Primer*. W3C Recommendation.

Noy, Natalya F, & McGuinness, D. L. (2001). *Ontology Development 101: A Guide to Creating Your First Ontology*. Available at http://protege.stanford.edu/conference/2005/slides/T1_Noy_Ontology101.pdf.

Noy, Natalya Fridman, Fergerson, R. W., & Musen, M. A. (2000). The Knowledge Model of Protégé-2000 : Combining Interoperability and Flexibility. *Knowledge Engineering and Knowledge Management Methods Models and Tools*, *1937*(1), 1–20.

Obitko, M. (2007). *Translations between Ontologies in Multi-Agent Systems*. PhD Tesis, Czech Technical University.

O'Connell, D. (2007). *System and Methods for Recycling of Cell Phones at the end of Live*. Patent number: US7251458B2.

Pal, S. K., Talwar, V., & Mitra, P. (2002). Web Mining in Soft Computing Framework: Relevance, State of the Art and Future Directions. *IEEE transactions on neural*, *13*(5), 1163–77.

Piatetsky-Shapiro, G., & Frawley, W. J. (1992). Knowledge Discovery in Databases: An Overview. *AI Magazine*, 13(3), 57.

R. Shearer, B. Motik, and I. H. (2008). HermiT: A highly-efficient OWL reasoner. *OWL: Experience and Directions, 26-27*.

Schmid, H. A., & Swenson, J. R. (1975). On the Semantics of the Relational Data Model. *International Conference on Management of Data - SIGMOD,* 211-223.

Schwartz, D. G., & Schreiber, Z. (2005). Semantic Information Management. *Encyclopedia of Database Technologies and Applications*, 593-600.

Seol, H., Lee, S., & Kim, C. (2011). Identifying New Business Areas Using Patent Information: A DEA and Text Mining Approach. *Expert Systems with Applications*, *38*(4), 2933–2941.

Shih, M.-J., Liu, D.-R., & Hsu, M.-L. (2010). Discovering Competitive Intelligence by Mining Changes in Patent Trends. *Expert Systems with Applications*, *37*(4), 2882–2890.

Sirin, E., Parsia, B., Grau, B., Kalyanpur, A., & Katz, Y. (2007). Pellet: A Practical OWL-DL Reasoner. *Web Semantics Science Services and Agents on the World Wide Web*, *5*(2), 51–53.

Smith, M. K., Welty, C., & McGuinness, D. L. (2004). *OWL Web Ontology Language Guide*. W3C Recommendation.

Soo, V.-W., Lin, S.-Y., Yang, S.-Y., Lin, S.-N., & Cheng, S.-L. (2006). A Cooperative Multi-agent Platform for Invention Based on Patent Document Analysis and Ontology. *Expert Systems with Applications*, *31*(4), 766–775.

Strassmann, P. A. (1997). *The Squandered Computer: Evaluating the Business Alignment of Information Technologies*. Information Economic Press.

Taduri, S., Lau, G. T., Law, K. H., Yu, H., & Kesan, J. P. (2011). An Ontology to Integrate Multiple Information Domains in the Patent System. *International Symposium on Technology and Society ISTAS*, 23–25.

Thalheim, B. (2000). *Entity-relationship Modeling: Foundations of Database Technology*. Springer.

Trappey, A. J. C., Trappey, C. V., Wu, C.-Y., & Lin, C.-W. (2012). A Patent Quality Analysis for Innovative Technology and Product Development. *Advanced Engineering Informatics*, *26*(1), 26–34.

Trappey, A., Trappey, C., & Wu, C.-Y. (2009). Automatic Patent Document Summarization for Collaborative Knowledge Systems and Services. *Journal of Systems Science and Systems Engineering*, *18*(1), 71–94.

Van Heijst, G., Schreiber, A. T., & Wielinga, B. J. (1997). Using Explicit Ontologies in KBS Development. *International Journal of Human-Computer Studies*, *46*(2-3), 183–292.

W3C. (2004). *Web Services Architecture*. W3C Recommendation.

W3C. (2006). *Extensible Markup Language (XML) 1.1 (Second Edition)*. W3C Recommendation.

Wanner, L., Brügmann, S., Diallo, B., Giereth, M., Kompatsiaris, Y., Pianta, E., Rao, G., et al. (2009). PATExpert: Semantic Processing of Patent Documentation. Available at http://recerca.upf.edu/patexpert.

Witten, I., & Frank, E. (1999). Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations. *Special Interest Group on Management Data, 31*(1), 76-77.

CAPÍTULO 3

THE IMPORTANCE OF CORPORATE WEBSITES FOR DIFFUSION OF ENVIRONMENTAL INNOVATIONS: THE INFLUENCE OF TRUSTING BELIEFS FOR RECRUITMENT PROCESSES

THE IMPORTANCE OF CORPORATE WEBSITES FOR DIFFUSION OF ENVIRONMENTAL INNOVATIONS: THE INFLUENCE OF TRUSTING BELIEFS FOR RECRUITMENT PROCESSES

3.1 Introduction

Monster.com is currently the world's largest online job-search website. With operating expenses in the range of US$1.3 million in 2008 (BusinessWire, 2008), Monster illustrates the emerging importance of suppliers for recruiting-related online services. Additionally, a growing number of firms have developed innovations to move their recruitment processes to an online context with only the support of their own websites. The online recruitment process consists of offering positions via the Internet while attempting to attract applications from the best candidates. Recruiting-related online innovations represent growing and high-potential opportunities for employers to broaden their recruiting efforts as well as reduce costs (Maurer & Liu, 2007). It is estimated that 55% of newly hired employees in the United States in 2006 were already managed from online sources (Cober & Brown, 2007), and the number is probably increasing.

The diffusion of environmental innovative approaches for online recruitment may experience bias, however, due to potential employees' lack of trust in firms offering positions online, particularly when the firms are small, operate in a risky industry, or are relatively unknown. The best candidates usually behave selectively in these contexts, avoiding risks such as lost time spent in filling out recruitment forms, sharing personal details with an unknown organization, or becoming involved in a long or frustrating process with an uncertain aim.

While literature on the adoption and diffusion of technological innovations has widely used the Theory of Reasoned Action (TRA) (Fishbein & Ajzen, 1975) to propose that users' beliefs influence their behavioral intentions (Hsu & Lu, 2004; Kolekofski & Heminger, 2003; Liker & Sindi, 1997; Lu & Lin, 2002), the trust beliefs have received minor attention. The focus thus far has mostly been on e-commerce. We propose that trust beliefs developed by users of a corporate website with regard to three characteristics of the firm–ability, integrity, and benevolence related to environmental strategy–influence their behavioral intention to accept a job offered online by that firm.

This chapter answers calls from the literature to pay specific attention to each of the three characteristics influencing trusting beliefs (Schoorman et al., 2007). Trust literature has been traditionally focused on the influence of perceived technical ability (and integrity) in e-commerce (Clarke, 2007; Martin, 2006; Simons, 1999). We show that each dimension can play a relevant role. This chapter contributes to help the extension of TRA by paying specific attention to the environmental innovation of the firms influencing trusting beliefs. We are especially innovative in the analysis of the dimension of benevolence.

This chapter also answers the calls in the literature for analyzing online relationships outside of the commercial sphere of customer-supplier (Gill & Mathur, 2007; Gill, 2008). An innovative context of analysis is also relevant from a theoretical point of view because trust literature especially highlights the importance of paying specific attention to the consequences in the final findings of the research context (Gefen & Heart, 2006; Pavlou & Dimoka, 2006; Serva & Fuller, 2004). Although the importance of online recruiting innovations has undergone scrutiny, we do not know of any previous papers analyzing how theoretical models in the literature may be useful in this context (Lee, 2005).

Our work began with intensive interviews, which included eight experts in the field of recruitment. These interviews were focused on how to analyze and simulate a real recruitment on line process. Our lab experiment involved 218 Master's students because their profiles are widely targeted by online recruitment campaigns. The subjects agreed to participate in a lab experiment in which they were offered a consulting position in a chemical firm that could only be known through its corporate website. To this end, we design a website with information about the environmental innovation strategy.

To summarize, this chapter makes three contributions to the literature on the diffusion of online innovations. First, we pay extra attention to the trusting beliefs extending the traditionally used TRA (and TAM -Technology Acceptance Model-) models for the analysis of innovations acceptance. Second, while the vast majority of previous works, analyzing how trusting beliefs influence on online adaptations, have been focused on the trust of online buyers (Chen & Dhillon, 2003; Jarvenpaa et al., 2000; Jones & Leonard, 2008; Kuan & Bock, 2007; Lee, et al., 2009; Tagliavini et al., 2001), literature has highlighted the importance of understanding the trust of agents who do not purchase online a firm's products or services (Gefen et al., 2008). These agents are particularly important because the vast majority of online users are merely retrieving information rather than making purchases (Choudhury & Karahanna, 2008). Previous works have already suggested differences between the use of corporate websites for information gathering and for making purchases (Choudhury & Karahanna, 2008; Pavlou & Fygenson, 2006). Third, research regarding the influence of trusting beliefs on the online diffusion of innovations has paid little attention to the specific implications of the different attributes of the trustee (Gefen et al., 2008), specifically of the environmental innovations. Our work finds that two attributes receiving less attention in previous literature – integrity and benevolence – show strong potential in the diffusion of recruiting-related online innovations.

This chapter includes implications of our results for practitioners' purposes. The practical importance of analyzing how trust works in initial job choice decisions is relevant because it significantly affects employees' job satisfaction, organizational commitment, and turnover intentions (Alder et al., 2006). This chapter also suggests implications for situations in which different agents (e.g., suppliers, local authorities, neighbors, or scholars) often use a corporate website before determining their level of trust in the firm.

The structure of this chapter will be as follows. We will begin with a review of the importance of the trust in online environments focusing on non-commercial decisions such as those related to the job market. We propose our three hypotheses analyzing if the online users' beliefs in the ability, integrity and benevolence of the environmental strategy of the firm influence their decision to trust when dealing with the decision of accepting a job. Then, we will describe our methodology and results. Finally, we will discuss our findings, limitations and implications for practitioners and scholars.

3.2 Theoretical Background and Hypotheses

The Theory of Reasoned Action (TRA) views a person's intention to perform a behavior as the immediate determinant of the action; at the same time, a person's beliefs are antecedent to behavior intent to adopt and use any system (Ajzen & Fishbein, 1980). Literature has widely used TRA to study the adoption of the technological innovations analyzing the relationships between user beliefs and behavioral intention to use the innovations. Based on the TRA, a person's intention is a function of two basic determinants. The first one is a personal factor related to the individual's positive or negative evaluation of performing the behavior ("attitude toward the behavior"). The second determinant of intention is the person's perception of the social pressure put on him/her to perform or not to perform the behavior in question ("subjective norm"-which relates to the normative considerations) (Ajzen & Fishbein, 1980). Figure 3.1 illustrates our TRA description.

The TAM proposed by Davis (1989) and Davis et al. (1989) derived from TRA has been widely used in explaining the adoption and usage of information technologies (Gefen & Straub, 2000, 1997; Hsu & Lin, 2008; Ong et al., 2004; Venkatesh, 2000; Venkatesh & Davis, 2000). TAM predicts user acceptance of any technology to be determined by two factors: perceived usefulness and perceived ease of use.

TRA and TAM are both built on the assumption that human beings are normally quite rational and make systematic use of the information available to them; however, TRA and TAM might have more difficulties explaining decisions of trust in risky interactions with anonymous organizations. Trusting belief is the individual belief (not necessarily rationally founded) that the trustee has characteristics that would benefit the trustor (Mayer et al., 1995). These beliefs lead to trusting intention, which is the willingness or intention of the trustor to rely on the trustee. Finally, trusting intention leads to trusting behavior, which is the act of the trustor becoming vulnerable to the trustee in a situation of uncertainty. This cognitive process of trust formation has been shown to positively influence a person's intention to use e-commerce websites (Gefen et al., 2003a; Gefen et al., 2003b; McKnight et al., 2002; Wu & Chen, 2005).

Hence, trust implies the willingness to be vulnerable based upon positive expectations of the behavior of another under conditions of risk and interdependence (Rousseau et al., 1998). Three characteristics of the trustee have been analyzed as relevant to influence the consumers' trust in the usage of e-commerce: ability, integrity, and benevolence (McKnight et al., 1998). Ability means the potential of providing a quality output in time; integrity suggests that the firm will fulfill agreements as promised; and benevolence means a strong desire to serve buyer's needs aside from an egocentric profit motive (Scharl et al., 2001). Although these three dimensions have been accepted as important in the literature of e-commerce acceptance, ability has received much attention; the others, integrity and benevolence, have been accepted as marginal or aggregated (Scharl et al., 2001). Relevant literature has called for extending the disaggregated analysis of the three dimensions and analyzing it outside of commercial environments (Gill & Mathur, 2007; Gill, 2008).

These two remaining issues are relevant when information supplied by the corporate websites has exponentially grown in the last years, including a wide range of non-commercial aspects. Most firms have a corporate website, and it usually offers a broad range of information, such as the mission statement, strategic objectives, functional details, customers' profiles, history, and financial information, among other features. All these online issues developed by the firms to build their corporate website look well connected to engage the different stakeholders when the firm is dealing with emerging technologies (Au & Kauffman, 2008). Additionally, trust literature has highlighted that different contexts may raise theoretical concerns regarding implications of each trust dimension (Ganesan et al., 2003; Pavlou & Dimoka, 2006; Serva & Fuller, 2004).

The job choice process begins with an individual's evaluation of information obtained from recruitment sources (Gatewood et al., 1993). Although it is well known that the image of an organization affects the initial decisions of potential applicants, it is still unknown how candidates' beliefs affect theirs decisions (Gatewood et al., 1993; Rynes 1991). The importance of corporate sources of information (Herriot & Rothwell, 1981) may be easily accepted as greater for the development of initial trust (Koufaris & Hampton-Sosa, 2004), especially when the candidate lacks prior knowledge of the firm as common in the online context.

Potential employees may have corporate information before facing the decision of working in the firm; even so, they will probably want to check the corporate website to reinforce or modify their previous beliefs. However, globalization, the lack of geographical proximity, and the invisibility of many firms to media (especially those with small or medium size) often imply the absence of previous interaction between potential candidates and the firm. This underscores a maximum level of importance for the website of the firm.

Some authors have suggested that the Internet has revolutionized how people search jobs (BusinessWire, 2008). Corporate websites are one of the most used sources to attract employees. Data show that all Fortune 100 companies and 94% of all global companies use their websites for recruitment (Greenspan, 2003; Lee et al., 2009; Maurer & Liu, 2007). Although literature has offered descriptive evidence to suggest the importance of reinforcing this process using marketing tools (Maurer & Liu, 2007) or a perspective of human resource management (Parry & Tyson, 2008), previous works have failed to provide a model and evidence to explain how these things influence online users whose willingness to work in the firm may be increased (Maurer & Liu, 2007). Although a person's perception of social pressure is important to understand behavioral intention, the scope of this chapter focuses on the three dimensions of trust related to the personal norm and the different relationship of each one with the behavioral intention. Figure 3.1 illustrates our hypotheses.

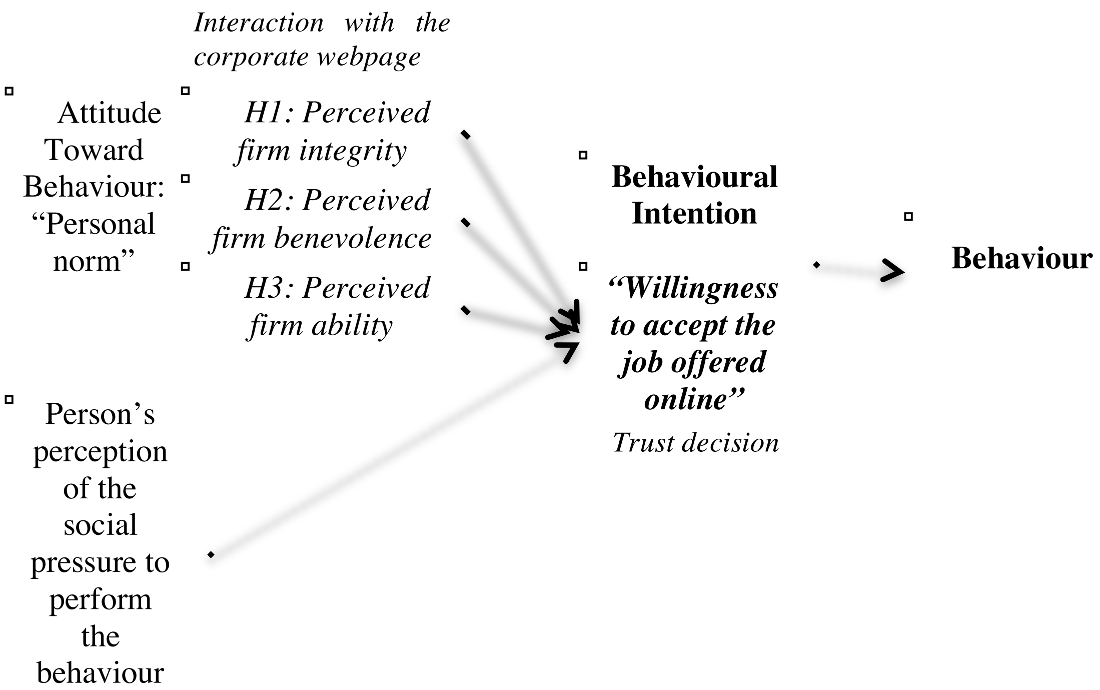

Figure 3.1 A description of TRA and its application to our analysis

3.2.1 The Firm's Integrity as Perceived by the Users of the Corporate Website

The consistency of the trustee's past actions, credible communications about the trustee from other parties, belief that the trustee has a strong sense of justice, and the extent to which the party's actions are congruent with his or her words all affect the degree to which the firm is judged to have integrity (Mayer et al., 1995). The literature on trust in online environments has focused on the buyers' concern regarding whether sellers will act opportunistically by not delivering the products (Gefen et al., 2008). However, when dealing with products or services which are not completely delimitated in advance, such as a hotel room, a tailored suit, or a job, online buyers are mostly concerned about inferring product quality and characteristics and buyers' trust in sellers is focused on product misrepresentation due to lack of seller integrity (Gefen et al., 2008). For instance, perceptions of the firm's integrity increase trust when deciding on a web-based recommendation agent because the customer believes that relying on the recommendation agents will provide truthful and objective recommendations (Komiak & Benbasat, 2006).

The willingness to accept a job implies that a trustee believes the firm will completely fulfill the promised conditions of employment. In the context of accepting a job when firm information is mainly coming from the corporate website, the firm integrity perceived by the online users will mean that the user believes that the firm is providing an objective panorama regarding the post and the firm. For instance, users assume that there is no hidden information about duties, salary, or activities.

Psychological implications of the integrity (versus more technically based roots of the firm's ability) might make more difficult the opportunities to reinforce the perceptions of firm's integrity using online information. Different alternatives might help to reinforce the online perception of integrity in this context, such as: the joint efforts of the firm with non-governmental organizations, voluntary collaborations with local authorities, or the support of prestigious customers, experts. If the online users perceive firm integrity, they will reduce the feeling of potential frauds when trusting in the firm. Thus we assert:

> ***Hypothesis 1 (Firm Integrity Hypothesis):*** *Increased firm integrity perceived by the users of the corporate website will increase their willingness to accept a job in the firm.*

3.2.2 The Firm's Benevolence Perceived by the Users of the Corporate Website

Benevolence is the extent to which a trustee is believed to want to do well on behalf of the trustor, aside from an egocentric profit motive (Mayer et al., 1995). In other words, benevolence is the perception that the trustee intends to act in the trustor's interest (McKnight et al., 1998), even if there is no extrinsic reward for the trustee. Trust emerges when trustors perceive trustees to be genuinely interested in the trustor's welfare and is motivated to seek joint gain (Doney et al., 1998).

Early works proposed benevolence as a basis for trust (Larzelere & Huston, 1980). However, benevolence has received little attention in trust literature (Rynes, 1991) and is particularly scarce in the literature on trust in online environments (David, Gefen et al., 2008). A few works have shown the potential of benevolence in some specific online environments. For instance, Pavlou and Dimoka (2006) showed that benevolence has a stronger effect on price premiums than on ability and credibility (what we call *integrity* in our work) in online auction marketplaces.

In the context of accepting a job when firm information is mainly coming from the corporate website, the firm benevolence perceived by the online users will mean that the user believes that relying on the firm will provide an appropriate environment for the multiple situations in which an explicit commitment was not made (Ganesan et al., 2003). These kind of situations may occur more often in an employment relationship because of the importance of discretional decisions regarding promotions, international mobility, sick or family leaves, or bonuses.

The lack of previous interactions with the firm makes the users' belief linked to the corporate website particularly important. The employment relationships involve some communal relationships in which identification with the other is particularly important, whereas completely market-based exchanges (e.g., e-commerce) may emphasize rational calculus, where characteristics of interactions are based upon economic exchange (Rousseau et al., 1998). Explicit considerations of such things as codes of conduct, employment policies, or description of positive experiences represent some of the online characteristics that might be useful to determine the trustee's beliefs regarding firm's benevolence. Thus, we will test:

> ***Hypothesis 2 (Firm Benevolence Hypothesis):*** *Increased firm benevolence perceived by the users of the corporate website will increase their willingness to accept a job in the firm.*

3.2.3 The Firm's Ability Perceived by the Users of the Corporate Website

Trustors must determine that trustees are capable of meeting their obligations and the trustor's expectations, assuming that trustees differ significantly in their competence to deliver on their promises (Doney et al., 1998). In a more descriptive way, *ability* is that group of skills, competencies and characteristics that enable a party to have influence within a specific domain (Mayer et al., 1995). Multiple works have analysed similar constructs based on the perceptions of the technical competence and, for instance, trust in the industrial salespeople increases as the customer gains the impression that the salesperson is dependable and competent (Swan et al., 1985).

Literature on trust in online environments has shown that users' belief regarding a firm's ability is a primary requirement for e-commerce. This belief affects the intentions to inquire about products, although it does not mean that the user will purchase the product (Gefen & Heart, 2006). The addition of ability to the antecedents of trust creates a framework of trust that is domain-specific (Doney et al., 1998). Hence, it is important to examine the characteristics of each specific situation because abilities to trust a firm are different for different decisions. For instance, trusting beliefs in the firm's ability increase when a specific product or service is suggested by a web-based recommendation agent because the customer feels that the recommendation agent will generate well-customized recommendations (Komiak & Benbasat, 2006).

In the context of accepting a job when firm information is mainly coming from the corporate website, the firm ability perceived by the online users will mean that the user believes that relying on the firm will provide an appropriate environment for their professional development. In our experience, the firm's ability to offer successful development is usually linked to positive outcomes for the employee, such as the existence of attractive products or processes, the availability of resources for exploring innovative ways, or the opportunities of reinforcing their own vita with a working experience in a prestigious firm.

Although the users of corporate online information do not have direct experience with the intrinsic quality of resources, products, services or processes in the firm, users may generate a trustee's beliefs regarding firm ability after scanning specific issues perceived as proxies of the firm competence. These issues may include the number of employees or its financial size, its locations, description of the products or processes of the firm, financial information about the firm, or customers profile among others. Thus, we state our third hypothesis:

> ***Hypothesis 3 (Firm Ability Hypothesis):*** *Increased firm ability perceived by the users of the corporate website will increase their willingness to accept a job in the firm.*

3.3 Methodology

3.3.1 Participants and Experimental Procedures

We used a series of intensive interviews with eight experts in the processing of recruitment to make our methodology effective. They include: three respondents who were working in consulting firms, three others who were directors of human resource management, and finally two university scholars. We also followed several online debates regarding the potential of online recruitment. One was from *The Economist*, which conducted an online debate in September 2009 involving different professional views. This enabled us to better understand the complexity of the problem. Finally, our own experience in the field of IT consulting and innovation research has provided relevant input in the design of the methodology. In this section, we detail the design and procedures that we used to test these hypotheses.

After our detailed review of multiple corporate websites in the industry, we built a corporate website for a fictional firm using as a model the corporate website provided by a medium firm in the sector. Our selection was intended to be as representative as possible of the majority of firms in the industry.

We recruited 218 subjects attending business courses at the Master's level of a large public university in Europe. The selection is appropriate because this profile is one of the most targeted by online campaigns of recruitment. Additionally the lab experiment is appropriate because it allows us to focus on the potential influence of specific variables that we included in the analysis. As usual in lab experiments, we asked for our subjects' formal written agreement to participate in the experiment in advance. To motivate the participants to view the lab experiment as a real situation, they were told before the experiment that one participant would receive €200 and that all of the participants would get extra credit in their business courses. All experimental sessions were conducted in the same computer laboratories. The participants were told not to talk to anyone, and they worked in individual cubicles under the supervision of the authors our assistants.

After reading the instructions, subjects completed a pre-questionnaire on their socio-demographics characteristics, Internet experience, general trust in the Internet and firms, and interest in different socio-economic issues. This information was compiled to control potential influences.

We then distributed a document with information about the experimental situation. Each subject was offered a consulting position in a chemical firm. The text mentioned general details about duties involved and offered the opportunity to analyze additional environmental information of the firm through its corporate website. Participants were randomly assigned to one of two groups. The situation described for each group was the same except that one description emphasized the business uncertainty of the chemical industry, due to a previous environmental accident. This treatment was provided to check the attention of the users to instructions. We included the same number of participants in each group to avoid bias in the results. This document was maintained by the subjects and could be read during the time that they filled out the subsequent questionnaire.

The corporate website that we designed was always available for participants in the lab experiment. The participants were asked to answer assuming that the firm and its vacancy were real, but its fictional nature was important to avoid any previous interaction with the firm by participants.

Each participant participated in the experiment individually and was allowed to take as much time as needed. We checked that all of them were using the corporate website before answering the online questionnaires. Our analysis concerning the subjects' perception about the firm's ability, integrity and benevolence and the willingness to accept a job in the firm is based on their responses to questions in these instruments. As in real life, it was always possible to come back to the corporate website before the subject answered any questions that we asked them.

Participants spent between 25 and 32 minutes for the whole experiment. Two participants were eliminated in the final sample for lack of attention to the task: they gave answers to the entire questionnaire using only 15 and 18 min for each one. The online questionnaire did not accept missing data; all the questions had to be answered. The final sample remaining was 218.

3.3.2 Measurements

Scales to measure each of the constructs in the model were developed based on previous literature, and existing scales were used where possible. The subjects' answers showed a reasonable dispersion in their distributions across the ranges. Seven-point Likert scales were used for all survey items, ranging from "Total Disagreement" to "Total Agreement." We selected and adapted various items from Gefen (2002) because of his detailed verification of the scales' validity, reliability and his similar effort to study ability, integrity and benevolence (Pavlou & Dimoka, 2006). These scales are closely linked with our research and with our definitions of the constructs of integrity, ability and benevolence related to environmental strategy. The items and descriptive statistics can be seen in Table 3.1 y 3.2.

Table 3.1 Items for the analyzed variables

Items	Description
Willingness1	My global willingness to accept the job in this chemical company is very positive
Willingness2	I will work for this company without doubts
Willingness3	I am determined to accept the proposal of my boss to work in this chemical firm
Integrity1	I believe that firm A will meet its environmental commitments as described in its website
Integrity2	I believe that firm A is doing the advance environmental processes that are described in its website
Integrity3	I believe A's environmental commitments deserve confidence
Integrity4	I believe that the environmental commitments of the firm A are credible
Integrity5	I think that A's commitment to develop environmental improvements is sincere
Ability1	I believe that A is a company with high technical capabilities in the environmental field
Ability2	I believe that firm A understands perfectly the environmental problems of the industry it works, and how to solve them
Ability3	I believe that firm A understands quite well the environmental problems of its products and knows how to solve them
Ability4	I believe that firm A has excellent resources to accomplish its environmental obligations
Ability5	I believe that the approaches of the managers of the firm A are really accurate for a good environmental management
Benevolence1	I believe that if there was any environmental problem, A will be willing to guarantee my health even if it would endanger his reputation
Benevolence2	I believe that A would even be willing to make economic sacrifices if needed to ensure my safety environment
Benevolence3	I believe that A would not act opportunistically in environmental issues; even it has the chance to do it
Benevolence4	I believe that A will take into account my environmental concerns anytime
Benevolence5	I believe that A will continue making significant environmental improvements, even if it has to reduce its benefits to get it

Table 3.2 Descriptive statistics

Variables	N	Mean	Standard deviation	Minimum	Maximum
Willingness1	218	4.88	1.314	1	7
Willingness2	218	4.55	1.421	1	7
Willingness3	218	5.12	1.432	1	7
Integrity1	218	4.61	1.150	1	7
Integrity2	218	4.74	1.070	2	7
Integrity3	218	4.28	1.130	1	7
Integrity4	218	4.63	1.174	1	7
Integrity5	218	4.59	1.160	2	7
Ability1	218	5.44	0.90	3	7
Ability2	218	5.42	0.872	3	7
Ability3	218	5.65	1.038	2	7
Ability4	218	5.10	1.011	3	7
Ability5	218	4.99	1.12	2	7
Benevolence1	218	4.08	1.227	1	7
Benevolence2	218	4.08	1.160	1	7
Benevolence3	218	4.04	1.104	1	7
Benevolence4	218	4.50	1.053	2	7
Benevolence5	218	4.01	1.211	1	7

Integrity. Several studies have developed scales to measure and evaluate integrity, we used five adapted items from Gefen (2002), and asked to subjects to express their belief on the honesty of the experimental firm. We developed a confirmatory factor analysis to validate our scales ($\chi 2$ = 5.89 with 5 d.f., RMSEA = 0.02, GFI = 1.00, AGFI = 0.99, CN = 556.90) and showed that the scale was one-dimensional and had a high reliability, with Cronbach's α =.894). Although the recommendations of the literature differ, there is a consensus that values of Cronbach's α of around 0.8 constitute "acceptable" or "sufficient" reliability.

Benevolence. Although literature has paid minor attention to the measurement of benevolence, we selected and adapted various items from Gefen's scale (Gefen, 2002). We used five items and asked subjects to express their level of agreement or disagreement with the intentions of the experimental firm to do well, aside from wanting to make a legitimate profit. We developed a confirmatory factor analysis to validate our scales ($\chi 2$ = 7.59 with 5 d.f., RMSEA = 0.06, GFI = 0.99, AGFI = 0.98, CN = 432.26) and showed that the scale was one-dimensional and had a high reliability with Cronbach's α =.781.

Ability. In this research, we developed five items also drawing on a scale used by Gefen (2002). We asked subjects to express their level of agreement or disagreement with the existence of different abilities in the firm. We developed a confirmatory factor analysis to validate our scales ($\chi 2$ = 1.62 with 4 d.f., RMSEA = 0.0, GFI = 1.00, AGFI = 1.00, CN = 1,778.24). The process of validation rejected the inclusion of the item Ability3 in the final scale and showed that the final scale was one-dimensional and had a high reliability, with Cronbach's α =.766.

Willingness to accept the job. Previous literature usually refers to the intentions to accept vulnerability in e-commerce situations (Jones & Leonard, 2008). Schoorman et al. (Schoorman et al., 2007) have used the term *trusting intentions* to mean the trust that employees have in their general manager. In our study, three items were used to measure the trust determined as the related intention of accepting the job: professional willingness to accept the job, personal willingness to accept the job, and global willingness to accept the job. We used the three items and asked to subjects to express their level of agreement or disagreement. We used an arithmetic mean of the three answers for our final measurement. Cronbach´s α = 0.836 for this scale.

Control variables. The prior literature about online environments has shown there are multiple factors influencing the trusting intentions. These include factors related to the characteristics of the websites as well as individual differences among participants (Stewart, 2003). We control for several factors in this research. As the socio-demographics features of the participants were similar and only one website was used, we controlled potential effects of propensity to trust websites and for propensity to trust firms. We used three items for each variable, all of them directly obtained from Stewart (2003). These control variables fit with trust literature, suggesting that propensity may be an important factor to decide trust in a target (McKnight et al., 2002). Cronbach´s α = 0.781 for the scales representing the propensity to trust websites construct and α = 0.754 for the propensity to trust firms construct.

3.4 Results

Table 3.3 shows the means, standard deviations and correlation between the variables analyzed.

Table 3.3 Means, Standard Deviation, and Correlation

	Variable	Mean	s.d.	1	2	3	4	5	6
1	Willingness to accept the job (trust)	4.94	.98	-					
2	Propensity to trust websites	4.21	.79	.36***	-				
3	Propensity to trust firms	4.11	.77	.40***	.68***	-			
4	Perceived integrity	4.57	.95	.49***	.39***	.41***	-		
5	Perceived benevolence	4.14	.84	.52***	.28***	.29***	.55***	-	
6	Perceived ability	5.32	.71	.35**	.27**	.38**	.56**	.34**	-

N = 218 for all variables

Significance levels: * p<.10 * *p<.05 * **p<.01

Table 3.4 offers the results of the regression analysis (Jaccard & Turrisi, 2003) and the estimated coefficients of the different models.

The control variables are included in the first model in Table 3.4 For the equation including control variables, the results indicate that 17.6% of the variance of willingness to accept the job was explained by a propensity to trust websites and a propensity to trust firms. In the second model, we introduced the users' belief regarding firm integrity, which contributed significantly to improve the fit. An additional 11.9% of the variance was explained by users' belief regarding firm integrity. Our results supported the Firm Integrity Hypothesis (H1) ($p<0.01$). For the subjects in our sample, there is a positive relationship between firm integrity perceived by the users of the corporate website and their willingness to accept a job.

In the third model, we added the direct effect of the benevolence, which contributed significantly to improve our results. The Firm Benevolence Hypothesis (H2) was also supported ($p<0.01$). This indicated a positive relationship between firm benevolence perceived by the users of the corporate website and their willingness to accept a job. Finally, the fourth model includes the effects of the users' belief regarding firm ability. The Firm Ability Hypothesis (H3), which postulated a positive relationship between firm ability perceived by the users of the corporate website and their willingness to accept a job, was not supported. Multicollinearity was not observed as a problem, since the tolerance indices and variance inflation factors for each regression model were at acceptable values. Variance inflation factor values were all within acceptable range, with values less than 5.

Regarding our control variables, propensity to trust firms showed a positive and significant influence in all the models that we analyzed. However propensity to trust the website did not show a significant influence for our context when other variables were integrated in the analysis. As expected, there were no significant differences between the groups with or without uncertainty of the chemical industry, in terms of their answers to the pre-questionnaire. There were different perceptions about the uncertainty of the situation in a control analysis following the presentation of the informational document though.

Table 3.4 Results of the Multiple Regression Analysis

Variable	Model 1	Model 2	Model 3	Model 4
Constant	2.591*** (.356)	1.731** (.360)	1.090** (.363)	.899** (.458)
Propensity to trust website	.196* (.104)	.103 (.098)	.079 (.093)	.084 (.093)
Propensity to trust firm	.372*** (.108)	.240** (.102)	.224** (.097)	.210** (.099)
Integrity		.393*** (.066)	.219 ** (.071)	.195** (.079)
Benevolence			.387*** (.076)	.386*** (.076)
Ability				.064 (.093)
F	22.92***	29.75***	31.40***	25.152***
R^2	.176	.294	.371	.372
Change in R^2		.119	.077	.001

Willingness to accept the job, a proxy for trust, is the dependent variable. The values correspond to the non-standardized regression coefficients, with the typical standard error in parentheses.

Significance levels: * p<.10 **p<.05 ***p<.01

3.5 Discussion, Implications, and Future Research

Our results are consistent with the Theory of Reasoned Action (TRA), which asserts that beliefs positively affect attitude (Fishbein & Ajzen, 1975). More specifically, our results agree with works showing that trusting beliefs positively affect trusting intentions (Schoorman et al., 2007) and with literature stating the specific influence of the three dimensions of trusting beliefs in the online environments (Larzelere & Huston, 1980): ability, integrity and benevolence. Our model offers a useful theoretical approach to better understand innovation-related needs that are associated with the diffusion of new IT services. Psychological, technological, and business effects appear here as joint determinants of the diffusion and effects of IT services. This approach also provides a “broader umbrella” theorizing by accounting for different dimensions related to the diffusion of technological innovations.

While prior research with online customers has already stated the importance of the three dimensions of trustworthiness, limited attention has been paid to understand how different users of corporate websites generate trust in the firm and the specific influence of each trusting dimension (Gefen et al., 2008). In addition, our work complements previous findings, showing the positive influence of the trusting beliefs about the firms' environmental strategy, on the willingness to accept a job in the firm.

Moreover, our work illustrates the use of lab experiments as appropriate methods that can address research questions in the diffusion of IT services. The potential of this methodology to generate a real and properly controlled background offers appealing opportunities to better understand situations with difficult visibility in a different context. For example, online users' perception of integrity or benevolence may be specifically measured; thus, we can better define their implications for the action.

Our methodology allows us to complement previous findings in e-commerce literature. While previous literature has widely shown the importance of users' beliefs regarding technical ability, our results showed a significant impact of benevolence and integrity. However, ability was not statistically significant for our context. Although we do not have enough evidence to nail down the differences between our hypothesis and our results with full explanations, we can suggest a potential explanation. The level and the influence of various dimensions of trusting beliefs may be contingent on the different environmental context of our study, in comparison of the context used in the work of others. The specific stage of the generation of trust in the online context, the different implications of the decisions, and the unique profiles of the online users all serve as relevant factors.

We propose three other ways to explain these differences, and all of them should be tested in the future to improve our theoretical understandings of the key relationships in our study context. First, we believe that although the perception of the firm's ability may be a prerequisite for subsequent steps, integrity and benevolence may be the determinants for a final intention to trust in the firm. Our experimental setting does not allow us to check this possibility though. So future research might analyse how, if levels of perceived ability are not enough, that ability alone might determine that the online users do not trust in the firm. However, when perceived ability is enough, this dimension may not guarantee the trust in the environmental strategy of the firm.

Second, we suggest that benevolence might be particularly important in situations where an identification with the firm over time is particularly necessary and where completely regulation-based exchanges are not possible. For instance, there are employment relationships in which the candidate cannot be provided with a dependable environment for future promotions, where there are potential risks of changes in specific conditions like business hours, or where there is a friendly organizational culture. However, ability and more rational planning might be more important for a single market-based exchange (Rousseau et al., 1998). The majority of the e-commerce situations match this situation well, especially when the risks involved are low and the product is well known. For instance, it is more difficult to understand how benevolence might be important to decide a customer's final trust in the firm for buying a book or a CD. However, perceptions of a firm's ability and integrity might be the most relevant in situations in which the relationships are based upon less tangible features, such as the firms' environmental strategy.

Finally, trust is likely to form via a capability process in societies where people perceive a large competence gap and show respect for individual qualifications, expertise, and accomplishment (Doney et al., 1998). Since our sample focused on a subset of well-educated people, especially in those fields related to management, this profile of online users could generate a narrower appraisal of the connection between ability and trust.

Our findings have also some practical implications. The World Economic Forum reported that the overall trust level for companies in 2006 had been the lowest since they began this analysis in 2003 (Schoorman et al., 2007), and the situation may be even worse after the world financial crisis at the beginning of the twenty-first century. Additionally, the specific problems of trust for online contexts might still increase these tensions. Our results suggest a better understanding of websites as a "*sensory emotive space where we are able to move the web from an emotionally flat environment to a space of rich interactions*" (Kambil, 2008). In this context, managers might want to utilize specific observations from our work. We include some of the more salient ones below.

The huge potential of the Internet for offering detailed information and tailored communication to users appears to be a relevant influence on stakeholder trust. The corporate website should facilitate the delivery of customized content for consumers (Scharl et al., 2001) but also should reflect the unique needs and individual preferences of other kinds of stakeholders. Our study illustrates how the potential to generate a high willingness to work in the firm is strongly dependent on its capabilities to generate candidates' trust on perceptions about the firm's environmental behavior through its corporate website. Managers should try to answer online a complete set of general and specific information about their corporate behavior and intentions. This information should consider the multiple situations in which online users may want to examine the corporate website to support final decisions about whether a user will trust the firm. It is important to keep in mind that diffusion of company information through new technologies such as corporate webpages reinforces the interest of multiple and differentiated stakeholders in the firm. These include, customers, suppliers, investors, industry experts, jobs seekers, scholars, green activists or neighbors, among others.

Corporate websites usually focus on technical features, such as the products and processes in the firm or financial details. However, our results show the importance of paying specific attention to factors with the potential to generate greater feelings of integrity and benevolence among corporate website users, such as environmental strategies. Social networks and the diffusion of technology innovations that support online interactions and multimedia may be especially useful for firms to increase the market perceptions of their integrity and their benevolence. Real employees or customers may offer more appealing testimonials than the corporate press services for multiple stakeholders. It is important to note that corporate information available online influences the confidence of stakeholders and affects how different dimensions of trust may be important to build that trust.

As in any empirical analysis, our research has limitations. The lab experiment was carried out using participants from the university student population who were young and well educated, but who had limited experience in the job market. These specific characteristics reduce the potential of a universal applicability of our results, since many corporate recruiting websites are targeted to industry professionals. However, the decisions of accepting jobs for consulting firms is often associated with a population that is similar to what we have in our sample. Hence, the sample that we analyzed fits well with the problem as it is normally encountered in the real world, and so it reinforces the strength of our analysis.

Despite the study's limitations, this chapter represents one of the first specific analyses of the effects on the relevance of the trustworthiness beliefs –ability, credibility, and benevolence– in the online environments. The current research has focused on the connection between a set of factors related to trustworthiness beliefs on environmental strategies, and an applicant's willingness to accept the job. It fits well with our objective for this chapter. Our research will benefit form the consideration of additional factors influencing the evolution of the willingness to accept the job offered online. The inclusion of additional variables would be useful for a better understanding of the final evolution of the dependent variable and increasing the variance explain by the regression. Additionally, our analysis uses a non-commercial situation involving the willingness of an applicant to accept a job in the firm in order to show the relevance of our arguments in the most common situation of the online environments, when online users gather information in the websites helping to delimitate their trusting beliefs.

Our research can be extended in multiple directions. First, it would be appropriate to reinforce the attention paid to research that has analyzed technical approaches that allow managers to determine and continually optimize the corporate website according to the specific corporate business processes and objectives (Neubauer & Stummer, 2009). Second, our hypotheses should be tested and compared with those for other situations that do not involve the acceptance of a job. Different environmental situations might also generate situations in which trustworthiness beliefs vary in their influence on trust. In any case, our lab results are specific to our sample population. So, further testing of our hypotheses will be useful with other types of populations and other methods. Third, our study explores initial trust, but influences on trusting beliefs might be different over time. For instance, the contingent effect of the general business environment on the dynamics of trust may be of particular importance in mature industries. Finally, the effects of mistrust might be also an attractive topic. The growing potential of social networks on the Internet may reinforce the importance of unsatisfactory experiences with the firm.

References

Ajzen, I., & Fishbein, M. (1980). *Understanding Attitudes and Predicting Social Behavior*. Prentice-Hall.

Alder, G. S., Noel, T. W., & Ambrose, M. L. (2006). Clarifying the Effects of Internet Monitoring on Job Attitudes: The Mediating Role of Employee Trust. *Information & Management*, *43*(7), 894–903.

Au, Y. A., & Kauffman, R. J. (2008). The Economics of Mobile Payments: Understanding Stakeholder Issues for an Emerging Financial Technology Application. *Electronic Commerce Research and Applications*, *7*(2), 141–164.

BusinessWire. (2008). Monster Worldwide Comments on Expectations for First Quarter 2008 Operating Expenses. Available at http://www.about-monster.com/archive/2008.

Chen, S. C., & Dhillon, G. (2003). Interpreting Dimensions of Consumer Trust in E-Commerce. *Information Technology and Management*, *4*(2), 303–318.

Choudhury, V., & Karahanna, E. (2008). The Relative Advantage of Electronic Channels: A Multidimensional View. *MIS Quarterly*, *32*(1), 179–200.

Clarke, J. (2007). Trust, Reputation, Integrity and Professionalism Reflections on Business Relationships. *Accountancy Ireland*, *39*(1), 54–58.

Cober, R., & Brown, D. (2007). *Direct Employers Association Recruiting Trends Survey*. Booz allen Hamilton. Available at http://www.jobcentral.com/pdfs/DEsurvey.pdf.

Davis, F D, Bagozzi, R. P., & Warshaw, P. R. (1989). User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. *Management Science*, *35*(8), 982–1003.

Davis, Fred D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, *13*(3), 319–340.

Doney, P. M., Cannon, J. P., & Mullen, M. R. (1998). Understanding the Influence of National Culture on Development of Trust. *Academy of Management Review*, *23*(3), 601–620.

Fishbein, M., & Ajzen, I. (1975). *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*. Addison-Wesley.

Ganesan, P., Garcia-Molina, H., & Widom, J. (2003). Exploiting Hierarchical Domain Structure to Compute Similarity. *ACM Transactions on Information Systems*, *21*(1), 64–93.

Gatewood, R. D., Gowan, M. A., & Lautenschlager, G. J. (1993). Corporate Image, Recruitment Image, and Initial Job Choice Decisions. *Academy of Management Journal*, *36*(2), 414–427.

Gefen, D, Karahanna, E., & Straub, D. (2003). Potential and Repeat E-consumers: the Role of and Trust vis-a-vis TAM. *IEEE Transactions on Engeneering Management*, *50*(3), 307–321.

Gefen, D. (2002). Reflections on the Dimensions of Trust and Trustworthiness Among Online Consumers. *ACM SIGMIS Database*, *33*(3), 38–53.

Gefen, D., Benbasat, I., & Pavlou, P. (2008). A Research Agenda for Trust in Online Environments. *Journal of Management Information Systems*, *24*(4), 275–286.

Gefen, D., & Heart, T. (2006). On the Need to Include National Culture as a Central Issue in E-Commerce Trust Beliefs. *Journal of Global Information Management*, *14*(4), 1–30.

Gefen, D., Karahanna, E., & Straub, D. W. (2003). Trust and TAM in Online Shopping: An Integrated Model. *MIS Quarterly*, *27*(1), 51–90.

Gefen, D., & Straub, D. (2000). The Relative Importance of Perceived Ease of Use in IS Adoption : A Study of E-Commerce Adoption. *Journal of the Association for Information Systems*, *1*, 1–28.

Gefen, D., & Straub, D. W. (1997). Gender Differences in the Perception and Use of E-Mail: An Extension to the Technology Acceptance Model. *MIS Quarterly*, *21*(4), 389–400.

Gill, A. S. (2008). The Role of Trust in Employee-manager Relationship. *International Journal of Contemporary Hospitality Management*, *20*(1), 98–103.

Gill, A. S., & Mathur, N. (2007). Improving Employee Dedication and Pro-social Behavior. *International Journal of Contemporary Hospitality Management*, *19*(4), 328–334.

Greenspan, R. (2003). *Job Seekers Have Choices*. ClickZ Networks. Available at http://www.clickz.com/clickz/news/1699182/job-seekers-have-choices.

Herriot, P., & Rothwell, C. (1981). Organizational Choice and Decision Theory: Effects of Employers' Literature and Selection Interview. *Journal of Occupational Psychology*, *54*(1), 17–31.

Hsu, C., & Lin, J. (2008). Acceptance of Blog Usage: The Roles of Technology Acceptance, Social Influence and Knowledge Sharing Motivation. *Information & Management*, *45*(1), 65–74.

Hsu, C.L., & Lu, H.P. (2004). Why do People Play On-line Games? An extended TAM with Social Influences and Flow Experience. *Information & Management*, *41*(7), 853–868.

Jaccard, J., & Turrisi, R. (2003). *Interaction Effects in Multiple Regression*. Sage Publications, Inc.

Jarvenpaa, S. L., Tractinsky, N., & Vitale, M. (2000). Consumer Trust in an Internet Store. *Information Technology & Management*, *1*(1-2), 45–71.

Jones, K., & Leonard, L. (2008). Trust in Consumer-to-consumer Electronic Commerce. *Information & Management*, *45*(2), 88–95.

Kambil, A. (2008). What is Your Web 5.0 Strategy? *Journal of Business Strategy*, *29*(6), 56–58.

Kolekofski, K., & Heminger, A. (2003). Beliefs and Attitudes Affecting Intentions to Share Information in an Organizational Setting. *Information & Management*, *40*(6), 521–532.

Komiak, S. Y. X., & Benbasat, I. (2006). The Effects of Personalization and Familiarity on Trust and Adoption of Recommendation Agents. *MIS Quarterly*, *30*(4), 941–960.

Koufaris, M., & Hampton-Sosa, W. (2004). The Development of Initial Trust in an Online Company by New Customers. *Information & Management*, *41*(3), 377–397.

Kuan, H., & Bock, G. (2007). Trust Transference in Brick and Click Retailers: An Investigation of the Before-online-visit Phase. *Information & Management*, *44*(2), 175–187.

Larzelere, R. E., & Huston, T. L. (1980). The Dyadic Trust Scale: Toward Understanding Interpersonal Trust in Close Relationships. *Journal of Marriage & the Family*, *42*(3), 595.

Lee, D., Kauffman, R. J., & Bergen, M. E. (2009). Image Effects and Rational Inattention in Internet-Based Selling. *International Journal of Electronic Commerce*, *13*(4), 127–166.

Lee, I. (2005). The Evolution of E-Recruiting: A Content Analysis of Fortune 100 Career Web Sites. *Journal of Electronic Commerce in Organizations*, *3*(3), 57–68.

Liker, J. K., & Sindi, A. A. (1997). User Acceptance of Expert Systems: a Test of the Theory of Reasoned Action. *Journal of Engineering & Technology Management*, *14*(2), 147–173.

Lu, H., & Lin, J. C. (2002). Predicting Customer Behavior in the Market-space: a Study of Rayport and Sviokla's Framework. *Information & Management*, *40*(1), 1–10.

Martin, J. (2006). Multiple Intelligence Theory, Knowledge Identification and Trust. *Knowledge Management Research Practice*, *4*(3), 207–215.

Maurer, S. D., & Liu, Y. (2007). Developing Effective E-recruiting Websites: Insights for Managers from Marketers. *Business Horizons*, *50*(4), 305–314.

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An Integrative Model of Organizational Trust. *Academy of Management Review*, *20*(3), 709–734.

McKnight, D. H., Choudhury, V., & Kacmar, C. (2002). Developing and Validating Trust Measures for e-Commerce: An Integrative Typology. *Information Systems Research*, *13*(3), 334–359.

McKnight, D. H., Cummings, L. L., & Chervany, N. L. (1998). Initial Trust Formation in New Organizational Relationships. *Academy of Management Review*, *23*(3), 473–490.

Neubauer, T., & Stummer, C. (2009). Interactive Selection of Web Services under Multiple Objectives. *Information Technology and Management*, *11*(1), 25–41.

Ong, C. S., Lai, J. Y., & Wang, Y. S. (2004). Factors Affecting Engineers' Acceptance of Asynchronous E-learning Systems in High-tech Companies. *Information & Management*, *41*(6), 795–804.

Parry, E., & Tyson, S. (2008). An Analysis of the Use and Success of Online Recruitment Methods in the UK. *Human Resource Management Journal*, *18*(3), 257–274.

Pavlou, P. A., & Fygenson, M. (2006). Understanding and Predicting Electronic Commerce Adoption: An Extension of the Theory of Planned Behavior. *MIS Quarterly*, *30*(1), 115–143.

Pavlou, P. A., & Dimoka, A. (2006). The Nature and Role of Feedback Text Comments in Online Marketplaces: Implications for Trust Building, Price Premiums, and Seller Differentiation. *Information Systems Research*, *17*(4), 392–414.

Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so Different after All: A Cross-Discipline View of Trust. *Academy of Management Review*, *23*(3), 393–404.

Rynes, S. L. (1991). Recruitment, Job Choice, and Post-Hire Consequences: A Nall for new Research Directions. *Handbook of Industrial and Organizational Psychology, 2*, 399–444.

Scharl, A., Gebauer, J., & Bauer, C. (2001). Matching Process Requirements with Information Technology to Assess the Efficiency of Web Information Systems. *Information Technology and Management*, *2*(2), 193–210.

Schoorman, F. D., Mayer, R. C., & Davis, J. H. (2007). An Integrative Model of Organizational Trust: Past, Present, and Future. *Academy of Management Review*, *32*(2), 344–354.

Serva, M. A., & Fuller, M. A. (2004). The Effects of Trustworthiness Perceptions on the Formation of Initial Trust: Implications for MIS Student Teams. *Journal of Information Systems Education*, *15*(4), 383–395.

Simons, T. L. (1999). Behavioral Integrity as a Critical Ingredient for Transformational Leadership. *Journal of Organizational Change Management*, *12*(2), 89–104.

Stewart, K. J. (2003). Trust Transfer on the World Wide Web. *Organization Science*, *14*(1), 5–17.

Swan, J., Trawick, I., & Silva, D. (1985). How Industrial Salespeople Gain Customer Trust. *Industrial Marketing Management*, *14*(3), 203–211.

Tagliavini, M., Ravarini, A., & Antonelli, A. (2001). An Evaluation Model for Electronic Commerce Activities within SMEs. *Information Technology and Management*, *2*(2), 211–230.

Venkatesh, V. (2000). Determinants of Perceived Ease of Use: Integrating Control, Intrinsic Motivation, and Emotion into the Technology Acceptance Model. *Information Systems Research*, *11*(4), 342–365.

Venkatesh, V., & Davis, F. D. (2000). A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. *Management Science*, *46*(2), 186–204.

Wu, I. L., & Chen, J. L. (2005). An extension of Trust and TAM Model with TPB in the Initial Adoption of on-line Tax: An Empirical Study. *International Journal of Human-Computer Studies*, *62*(6), 784–808.

CAPÍTULO 4

THE RELATIONSHIP BETWEEN PATENTED ENVIRONMENTAL INNOVATIONS AND FIRM PERFORMANCE: THE INFLUENCE OF THE INTERNATIONAL SCOPE

THE RELATIONSHIP BETWEEN PATENTED ENVIRONMENTAL INNOVATIONS AND FIRM PERFORMANCE: THE INFLUENCE OF THE INTERNATIONAL SCOPE

4.1 Introduction

The existing literature has often insisted on the positive consequences of corporative environmental innovations to improve both the sustainable approach of the firm and the firm's financial performance (Russo & Fouts, 1997; Sharma & Vredenburg, 1998), but green technology innovation has not always been successful (Marcus, 2005; Marcus & Fremeth, 2009). Most previous works have analyzed the general relationship between environmental innovation and firm performance, but only few have studied specific factors influencing this relationship. This chapter highlights the importance of paying specific attention to the implications of factors linked to managers' decisions to protect competitive advantages of the environmental first-mover firms in an international context.

To illustrate the importance of our analysis with an example, U.S. firms remain the technology leaders in the solar photovoltaic industry. However, Chinese firms' abilities to replicate and upgrade the most basic developments in the industry have already managed to dominate the production of crystalline silicon cells, representing 65% of the world's total capacity in 2009 and expected to grow by 20% each year (Usha & Schuler, 2011). This situation, like several other similar cases, helps to understand the relevance of analyzing whether the international patent approach can make a difference in the relationship between environmental innovation and firm performance.

From a resource-based view, a firm's patents are resources that may create a barrier to protect the rents generated from innovation efforts or otherwise prevent rivals from competing in the same innovative space (Grant, 1991). A patent is a temporary legal title protecting an invention by granting its owner exclusive rights over the manufacture, use, import, or sale of the underlying product or process (WIPO, 2004). Patents are illegal to copy without authorization but are public by nature because they are designed to simultaneously protect inventors and promote more innovation through the public information available in the patent application (Arora et al., 2001); consequently, rivals are often able to invent around relevant patents. In other words, whereas patents may be very effective for certain contexts or technologies (e.g., pharmaceutical formulas), patent protection is very porous, imperfect and unclear (Lemley & Shapiro, 2005), and appropriability from patents is determined by a firm's strategies and actions (Pisano, 2006; Somaya, 2012).

In this context, it is surprising that the implications of protecting corporate innovations by filing patents in an international context have received almost no attention in the existing management literature [see (Somaya, 2012; Hanel, 2006) for recent and exhaustive reviews of literature on patents and management]. Our research questions in this chapter are as follows: are patented corporate environmental innovations positively influencing firm performance; and do the managers' decisions regarding the international scope of the firms' patents influence the relationship between patented environmental innovations and firm performance and, if so, how?

Firms may have proprietary assets -such as technological know-how, patents, brands, or reputation- or internal skills –such as organizational innovation- that lead to a competitive advantage in international markets (Hitt et al., 2006; Lu & Beamish, 2004). A natural resource-based view proposes that innovation is one of the most positive outputs of being environmentally advanced (Hart, 1995). The degree to which innovations are protected should affect how firms profit from innovation (Teece, 1986).

Several studies on organizations and the natural environment have reported evidence of a generally positive relationship between advanced environmental strategies and financial performance [e.g., (Margolis & Walsh, 2003) for a review]. However, relatively little research has studied the effects of patent strategies on firm performance (Somaya, 2012), and no research has examined this phenomenon in the environmental arena. We want to focus on the decisions linked to the environmental innovations and contribute by analyzing the specific relationship between patented environmental innovations and financial performance and whether the international geographic scope influences this relationship.

Our work makes at least two theoretical contributions. First, we extend the resource-based view literature on environmental innovations (Russo & Fouts, 1997; Sharma & Vredenburg, 1998; Shrivastava, 1995), paying specific attention to the opportunities of patented innovations to enhance firm performance. We study the international scope of corporate environmental patents to answer the calls in previous studies to better understand the specific conditions allowing firms to improve financial performance using resources and capabilities emerging from environmental approaches (Aragón-Correa & Sharma, 2003).

Second, our analysis of the international approach of the patenting process complements two relevant fields. On the one hand, we contribute to the previous literature on the connections between international business and environmental issues, which is mostly focused on the advantages of being environmentally advanced in gaining privileged access to international markets (Darnall et al., 2008; Madsen, 2009;

Pinkse & Kolk, 2012; Pinkse et al., 2010), by analyzing the implications of the international strategies to protect a firm's environmental innovations. On the other hand, we also complement the previous patent literature on proprietary patent strategies (Lippman & Rumelt, 2003; Mazzoleni & Nelson, 1998), which uses the resource-based logic of patents as key competitive barriers to imitation but which has paid very limited attention to the effects of managers' internationalization strategies regarding patents.

The structure of this chapter includes three main parts. After this introduction, we will review the literature and propose our hypotheses. Later, we will describe our sample, methodology, and results. Finally, we will discuss our results and propose some implications for other scholars, practitioners and policy makers.

4.2 Theoretical Background and Hypotheses

4.2.1 Patented Environmental Innovations and Firm Performance

The resource-based view has been widely used to justify the relationship between a firm's environmental progress and its performance, arguing that proactive environmental strategies (meaning those that are preventive and voluntary beyond regulations) generate positive interactions between multiple organizational capabilities and the firm's performance (Hart, 1995; Marcus & Geffen, 1998). For example, proactive environmental strategies are linked to the existence of complex capabilities for continuous innovation, stakeholders' management, organizational learning, or a shared vision (Hart, 1995).

Corporate innovations related to the natural environment have gained in importance in recent decades; consequently, an increasing number of organizations have developed organizational capabilities related to green innovations, whether they are motivated by the increase in external pressures or whether they see environmental innovation as an strategic opportunity that presents long-term advantages rather than unrecoverable costs (Russo & Fouts, 1997).

The degree to which innovations are protected by legal and other mechanisms affects how firms profit from innovation (Teece, 1986). In this context, the role of patents in providing inventors with a proprietary business advantage in opportunities created by their innovation investments is broadly recognized in the economics of patents (Mazzoleni & Nelson, 1998). In addition, multiple works (Cohen et al., 2000) show that firms have various other reasons to file patents, such as blocking rivals from using the patented invention, using them as strategic bargaining chips (in cross-licensing agreements), or using them as a means to measure the internal performance of the firms' scientists and engineers.

The patent-related actions and decisions are generally expensive, and the typical globally prosecuted and maintained patents cost approximately $100,000 plus organizational costs –especially the time and attention of key technical and management personnel- which may exceeded the expenses noted above (Somaya, 2012). However, the use of patents to protect environmental innovations may be a relevant system to guarantee the advantages traditionally linked to organizations that carry out environmental innovations to obtain "first-mover advantages", including higher prices charged for their products, the development of new markets, and image enhancement (De Burgos Jiménez & Céspedes Lorente, 2001; Porter & Linde, 1995b; Shrivastava, 1995). It is important to recognize that environmental innovations may include tangible developments (e.g., products or technology) and intangible processes in the firm (e.g., employee motivation or shared values within the firm). Although the environmental innovations rooted in those intangible assets may be difficult for patents to protect (because patents require a specific and precise delimitation of the innovation), environmental innovations that are more directly linked to large investments to generate new products, processes or materials appear as excellent candidates for patent protection.

Whereas studies involving product and process patents abound, relatively few have attempted to measure the impact of these patents at the firm level (Martin & Mykytyn, 2009). Some literature in the last decade demonstrates that a firm's financial market value may be positively influenced by its patented innovations (Arora et al., 2008; Czarnitzki & Kraft, 2000; Hall et al., 2005; Reitzig, 2004), and other works have also found that patents can also increase a firm's market value [for a summary, see (Hanel, 2006)]. Whereas the previous literature usually confirms a general value-added impact in the short term, the longitudinal impact of the patents or the geographic scope of the patents were not evaluated in previous studies (Petruzzelli, 2011).

In the environmental arena, the different features of the patenting processes may be particularly appropriate for several reasons, especially compared with other approaches to protecting innovation (i.e., secrecy). First, environmental technological innovations often require high investment, and patents are a safe way to protect the generated returns of heavy investments. Second, the speed of changes and uncertainty regarding the future evolution of environmental regulation (Marcus et al., 2011) suggests the importance of protecting the innovative assets that may be appropriate for developing future requirements even when they are unclear today. Finally, the transversal nature of environmental innovations requires a high coordination of the firm's functional capabilities to delimit the specific features of the patented environmental innovation but also limits the possibility of protecting innovations against competitors by enabling secrets to emerge from a broad and heterogeneous group of collaborators.

These reasons suggest that the firm's environmental innovations and the process of patenting these innovations are not only additional costs that organizations must bear but also have positive effects that may increase firm performance. Therefore, we hypothesize the following:

> ***Hypothesis 1.*** *The firm's patented environmental innovations positively influences the firm's performance.*

4.2.2 International Scope of Patented Environmental Innovations and Firm Performance

Geographic scope can be defined as the set of locations of a firm's business activities (Goerzen & Beamish, 2003). A broader international scope of a firm is linked to a higher number of international locations where the firm operates within its value chain. Different works have showed how regional logics impact on organizational responses to an overarching market logic (e.g. Greenwood et al., 2009). In this chapter, we will pay specific attention to two relevant dimensions of the international scope of the firms' developments of patented environmental innovations: the international scope of the knowledge sourcing and the international scope of the regions selected for protection.

4.2.2.1 The International Scope of the Knowledge Sourcing of Patented Environmental Innovations

The international business literature has insisted that firms may gain access to new knowledge and generate more radical innovations by increasing the geographic scope of firms' activities (Barkema & Vermeulen, 1998; Bartlett & Ghoshal, 1989;

Kogut & Zander, 1992). Foreign activities reinforce the process of accessing knowledge resources that would be more difficult to find if the firm develops its activity only in a domestic market (Almeida & Kogut, 1999; Bartlett & Ghoshal, 1989; Nachum & Zaheer, 2005). When firms develop their activities within a broader geographic scope, they are able to access new lines of technological and market diversification that are reflected in local markets (Iwasa & Odagiri, 2004) and are able to obtain a greater diversity of flow of ideas, products, process and technologies (Håkanson & Nobel, 2001). To summarize, the existing international literature has clearly supported that firm's international activities provide rare knowledge and train value capabilities with the potential to generate a competitive advantage.

However, the existing literature has paid less attention to the specific influence of the scope of the international developments related to the process of generating innovations. Patents offer an excellent context for this analysis because patent documents store information about priority office, i.e., the country where the first application is filed before protection is extended (or not) to other countries. We find some opposing arguments in the previous literature when we want to discuss the potential influence of the scope of the international sourcing of patented innovations.

On the one hand, Lahiri (2010) notes that firms are heterogeneous in the benefits derived from the geographic heterogeneity of their research and development activities. Results in the U.S. global semiconductor manufacturing industry show that the technological diversity of resources and the intra-organizational linkages between research units significantly influence firms' ultimate ability to derive benefits from increasing the geographic scope of their research and innovation activities. Chung and Yeaple (2008) analyze the patent stocks of a sample of U.S. firms and find that the countries with industries with greater technological similarity to the U.S. are more attractive to the U.S. firms, suggesting that a broad international scope of the patenting process may not generate synergies. Kotabe et al. (2007) also analyze the geographic distribution of R&D through an analysis of patent data, and they suggest that many international disperse sourcing of knowledge may increase complexity beyond acceptable levels for effective communication and coordination.

On the other hand, different literature shows that firms with a broader international knowledge sourcing around the world can tap into different national systems of innovation (Cantwell, 2000). In this context, the heterogeneous types of knowledge emerging from different locations may be combined advantageously, supporting the argument for knowledge sourcing abroad to obtain “technological diversity” and increase the “knowledge portfolio” (Cantwell & Janne, 1999) to reinforce the potential to increase the positive performance influence of the patented innovations.

The evidence suggesting that countries are becoming more technologically specialized and divergent over time supports the international potential for generating a more diversified innovation portfolio (Archibugi & Pianta, 1992). Countries vary in the type and nature of their innovations and intellectual activity due to differences in their innovation systems, such as industrial property, trade policies, educational systems, and agencies promoting R&D. For example (Furman et al., 2002) show that the scale of the national R&D (R&D workforce, R&D spending) and its productivity (industrial property protection, openness to trade, sources of R&D funding) leads to differences in patent filing rates. Meanwhile, different countries also show substantial differences in terms of environmental priorities, regulations, and consumers’ interests (Christmann, 2004). In this context, reinforcing the opportunities of generating environmental innovations in different fields is particularly appealing because of the multiple dimensions involved in these environmental problems.

Comparing firms with and without international R&D (Penner-Hahn & Shaver, 2005) shows that firms with foreign R&D produce more patents and suggests the existence of a virtuous circle whereby this greater knowledge base can further foster benefits from increased international R&D. Furthermore, a broader scope of international sourcing of patents may help the process of adopting the entire group of innovations to multiple local markets, helping to overcome the fixed R&D cost hurdles. Finally, firms with foreign R&D are more successful in terms of higher sales growth rates due to new products than firms with only domestic R&D activities (Peters & Schmiele, 2011).

We propose that there exist positive benefits of a broader international sourcing of the patented environmental knowledge. Firms filing patents for environmental innovations in multiple international priority offices (instead of in just one location) have the advantage of being early movers in multiple heterogeneous contexts of environmental preferences and situations, but the combined effect of this heterogeneity allows better to protect the subsequent benefits against competitors' progresses in just one situation. In addition, the international sourcing of environmental knowledge allows a firm to gain legitimacy in multiple countries exponentially compared with a marginal increase from added patents when innovation is primarily connected to the same place.

We expect that a firm harnessing knowledge in different countries or regions to develop the uniqueness of their knowledge base increases the chance of positive synergies emerging from this knowledge. Our proposed hypothesis is as follows:

> ***Hypothesis 2.*** *A broader international scope of the knowledge sourcing of environmental patented innovation positively moderates the relationship between the firm's patented environmental innovations and firm performance.*

4.2.2.2 The International Scope of Protection of Environmental Patented Innovation

It is widely accepted in the existing international literature that the survival and success of a firm depend on the firm's capability to exploit its innovations in multiple international markets (Lu & Beamish, 2004; Subramaniam & Venkatraman, 2001; Zander, 1999). A firm's competitive advantage in global markets is dependent on the firm's ability to accumulate, exploit, recombine and innovate its set of firm-specific assets and to transfer these assets to the various nodes of its extended network (Hitt et al., 2006; Lu & Beamish, 2004).

Patents are specific assets protecting a firm's innovation and enhancing a firm's ability to appropriate the rewards of its innovation. A patent grants the exclusive right to commercially exploit the invention in the specific country where the patent is granted, and a patent must be filed in multiple countries to be used effectively in foreign markets. The process of protecting a patent in multiple countries must be developed individually in each country or through a combination of international and domestic processes (e.g., the European Patent Office process or the international system of the World Intellectual Property Organization).

However, filing a patent is costly both in terms of the costs of preparing the application (involving a complex legal process where specialized patent attorneys and internal research staff spend long time trying to describe the specific delimitation of the inventors contribution, filling the patent application, and answering the review of the examiners) and in terms of the administrative costs and fees associated with the grant and renewal procedure. Expanding a patent to a broader scope of international markets also entails some additional operational costs, such as those emerging from translations, potential differences in filing processes, the liability of foreignness, and the complexity emerging from the required coordination between a firm's activities in different institutional and cultural contexts (Gomes & Ramaswamy, 1999). In this context, firms usually internationally protect their innovations only when the managers have planned to derive some competitive benefits from exploiting their patents. Information on patent family size (number of patent offices where the patent is filed) has even been used as an indicator of the relevance of a specific patent held by a firm (Harhoff et al., 2003; Peters et al., 2012). We propose that a broader international scope of selected regions for the exploitation of a patented environmental innovation leads to a greater potential to influence the firm's financial performance for at least three different reasons.

First, the broader scope leads to scale economies linked to the higher volumes of customers and marketing economies (Carpano et al., 1994). Eckert et al. (2010) note that economies of scale emerging from internationalization decisions have positive consequences for shareholder value, especially if the firm owns intangible assets.

Second, the added value of patented environmental innovations will increase with the breadth of international exploitation because in these cases, the firm is able to appropriate a greater proportion of the total possible firm-specific asset's rents with similar costs (Kirca et al., 2011). The international protection and potential exploitation of patents reduce the research and development expenditures per unit (Carpano et al., 1994).

Both reasons are particularly relevant for the environmental field due to the high mobilization of the resources needed to analyze, test, and delimitate the environmental innovations (Shrivastava, 1995). Firm-specific assets explain the firm's performance better when the firm captures a higher proportion of the economic value that they create (Barney & Clark, 2007). Scale economies and efficiency may provide the opportunity of commercial viability to the patented environmental innovations. When the firm is able to find a broader international exploitation of its environmentally advanced approaches, the firm reinforces the firm's international legitimacy (Bansal, 2005; Kostova et al., 2008), transparency, reputation (Christmann, 2004), and internal coherence (Christmann & Taylor, 2006).

Third, in addition to previous benefits, filing environmental patents in a broader international scope may reduce risks. The risk spreading, the reduction of revenue fluctuations, and the increase in market power for a more effective lobbying pressure are all linked to a broader scope of internationalization (Lu & Beamish, 2004). The changing nature of citizens' environmental preferences, the quick evolution of the technological and scientific environmental knowledge, and the uncertainty regarding the evolution of environmental regulations (Marcus et al., 2011) are relevant features to understand the importance of being protected against risks through a broader international scope of the exploitation of patented environmental innovations.

For all these reasons, we propose that the patented environmental innovations have a higher positive impact on firm performance when they are exploited widely in the international markets; otherwise, the costs of the patented environmental innovations may provide fewer opportunities to balance the potential benefits. Therefore, we make the following hypothesis:

> ***Hypothesis 3***. *A broader international scope of patented environmental innovations positively moderates the impact of the firm's patented environmental innovations on the firm's performance.*

4.3 The Context: The Information and Communication Technology Industry and Environmental Patents

The Information and Communication Technology (ICT) industry is one of the most appropriate fields for measuring the connections between patents, international processes, and firm performance and is also appropriate for analyzing the emerging importance of patented environmental innovations. The importance of the ICT industry in the global economy, its inherent internationalization, the importance of innovation and patents in this industry reflected in its being the industry with the highest share of all patents filed under the Patent Cooperation Treaty (PCT) (Gurry et al., 2012), the existence of external pressures and the industry's stated willingness to go green are some of the reasons for this relevance. We will describe each of these factors below.

The ICT industry is an important contributor to the world economy, representing 6.4% of Gross Domestic Product in the U.S., 5% in the EU and 6.8% in Japan (European Commission, 2010a). Furthermore, the ICT drives 20% of global productivity growth. The ICT industry also generates relevant effects in other industrial activities. Currently, most firms depend on a complex ICT infrastructure to help them manufacture, sell, and distribute their products (such as computer networks, communication tools, or social networks), and the e-prefix, as in e-business, e-government, e-health or e-recruitment, among others, has become part of everyday life.

ICT is by definition an international sector. The companies leading the ICT sector have long ago understood that the only way to survive is to become global. All of the major Internet and hardware companies have subsidiary offices in several countries for reasons such as production costs, cheap manpower, or market expansion. Even the telephone companies, which have traditionally been national or government firms, have engaged in several global takeovers, mergers, and alliances.

Finally, the ICT manufacturing industry represents by itself a very relevant percentage of total R&D investment (e.g., it represents 25% of total R&D investment and 32.4% of total business-sector research employees in Europe (European Commission, 2010b). Consequently, innovation in this sector is huge, and new applications, gadgets, products, and improvements appear daily. In this context, the importance of patents is also growing. According to the European commission, 20% of all EU patents are in ICT technologies (European Commission, 2010b). This figure is even larger in the U.S., where ICT represents 50% of the patents applied for by U.S.-based inventors in 2006 and where ICT is the industry showing the largest growth in the number of patent applications in the period 2005- 2009 (Gurry et al., 2012).

The ICT firms used to negotiate according to the number of patents they owned in each technology (Hytönen et al., 2011). In recent years, the number of patent lawsuits in ICT has also been growing (Bilton, 2010; Raghu et al., 2007). Patent trolls, companies that do not invent anything and have no intention of exploiting inventions but instead buy patents only to enforce them, are also particularly active in this industry. This new environment has led to the reinforcement of the importance of patents in the industry (e.g., to illustrate, the business press has insisted that Google bought Motorola Mobility in 2011 primarily to own its patents).

Finally, the ICT industry has a growing impact on the natural environment. The exponential utilization of this industry's products by consumers and firms increases the relevance of waste pressure, the direct and indirect requirements of energy utilization (e.g., cooling the huge servers), or the required polluting materials, including the controversial plastic or brominated flame retardants (Greenpeace, 2011). Although the ICT industry has not traditionally been very concerned with environmental issues, the situation is changing rapidly. To illustrate, the Greenpeace ranking (Greenpeace, 2011), which evaluates mobile and PC manufacturers, recognizes that the evaluated firms continue to perform poorly but that there have been notable improvements since the first ranking in 2006. In this sense, the International Telecommunication Union is also very active in “Greening the Blue”, a United Nations initiative to manage telecommunication industry’s sustainability performance and to reduce and offset the environmental impact of industry.

4.4 Methodology

4.4.1 Sample and Data Collection

We analyzed 3,087 environmental patent applications for the 2005-2009 period including all of the environmental patent applications of the 79 ICT firms on the Financial Times Global 500 firms list. Information extracted from data on individual patents was then used to construct firm-level variables. We also obtain financial information for each firm. With this information, we built a five-year unbalanced data panel with 116 observations.

We consider environmental patents to be those that have been assigned the Y02 code. The Y02 code is the code used by the European Patent Office (EPO) to label patents that describe technologies or applications for the mitigation of or adaptations to climate change. As described in the previous section, the ICT industry was selected for three reasons: (1) this industry is technology-intensive and innovative, and patents are a common competitive tool; (2) internationalization and globalization are relevant; and

(3) environmental issues in the industry are receiving growing attention in the last decade. Although the propensity to file patents differs depending on the nature of the technology (Cohen et al., 2000), by focusing on just one industry we can better control for differences in technology characteristics, which is much more difficult in cross-industry studies (Arora & Fosfuri, 2000).

We use the Global Patent Index database to analyze the environmental patent applications in the selected firm, finding that 36 different firms in the sample have patented at least one environmental innovation between 2005 and 2009. A total of 44.4% of the firms were from North America (the United States and Canada), 22.2% were from Europe (Finland, France, Germany, Italy, the Netherlands, Sweden and Switzerland), and 33.3 % were from Asia. The Global Patent Index is an online database with 100 searchable fields and with 70 million patent records from over 80 patent offices.

Financial data come from the Compustat global database and Bloomberg (a financial services database that provides information on firms, covering all market sectors worldwide). In addition, we have considered country-level characteristics using information obtained from the World Bank.

4.4.2 Measurements

We measured firm performance in this work with the firms' Tobin Q, which is included in the Bloomberg database. Tobin Q has been defined as the market value of the firm's assets to their replacement cost. Tobin's Q reflects markets' perceptions regarding current and potential profitability and is also often used as a good proxy for long-term performance (e.g. Nayyar, 1993). These features are especially appropriate for our work because the market is especially sensitive to the potential profitability derived from green innovation and internationalization.

Patented environmental innovation at the firm level was measured by counting the number of environmental patents registered each year by the firms under study. We consider environmental patents to be those that have been assigned the Y02 code used by the EPO to label patents that describe technologies or applications for the mitigation of or adaptations to climate change. The variable environmental patented innovation exhibits positive skew. We transformed this variable to its logarithm, reducing the skewness of the variable, yielding multiple regression residuals, whose distribution better approximated normality (Hamilton, 2006; Hair et al., 2009).

The international scope of the knowledge sourcing of patented environmental innovations was measured by gathering information on the number of different foreign countries listed as the first inventing country for each patent (priority country). We build an index of entropy of the geographic scope of each firm's patents with the YO2 ECLA code each year using an adaptation of the general formula (Hitt et al., 1997; Kim et al., 1993; Li & Qian, 2005).

$$GEI = \sum_{i=1}^{n} P_i \, . \, Ln\left(\frac{1}{P_i}\right)$$

where Pi is the percentage of patents that were first filed in region i, and Ln (1/ Pi) reflects the weight given to each region. Following Rugman and Verbeke's (2004) and Rugman (2005) delimitation of the international markets, we classified the regions as a broad triad of regions (North America, Europe, and Asia). A high value of this GEI index shows that the firm has used a broader geographic scope of international regions to file its innovations the first time. The advantage of using this measure (instead of just counting the number of different regions) is that we consider both the number of regions to which a firm files a patent first and whether these regions have similar relative importance (Hoskisson et al., 1993).

The geographic scope of exploitation of environmental patented innovation was measured by gathering information on the number of different regions in which a patent is filed (the family patent). We consider the triad of regions (North America, Europe, and Asia) to measure the geographic scope of exploitation of environmental patented innovation. A logarithmic transformation was not necessary this time because the variable meets the test for normality of distribution - Shapiro-Wills and Kolmogorov-Smirnov - ($z=1.42$; $p>0.05$ and $\chi2 = 5.67$ $p > 0.05$, respectively) (Hair et al., 2009).

Size. Firm size has been shown to be an important determinant of firm-level environmental conduct (Aragón-Correa, 1998) and also a traditional control variable when analyzing the performance of international firms (Christmann, 2004). We control for firm size using the logarithm of the total net revenues reported on each firm's balance sheet.

Firm level of innovation. We use the total number of patents filed each year by the firm, expressed in hundreds of patents, as a proxy measure that helps us to determine the degree of innovation within the firm (Penner-Hahn & Shaver, 2005). Firm innovation may influence environmental innovation and final firm performance.

Domestic growth. We use gross domestic product (GDP) growth for each country. GDP is a relevant measure of economic activity in a country and is defined as the market value of all goods and services produced within the borders of a country in a year.

Region. We use dummy variables to consider possible differences between the regions included into the sample: North America, Europe, and Asia (the reference category in our analysis).

Years. We use dummy variables to consider possible differences between the years included in the sample (2005-2009). Year t (2009) is the reference category.

Because practices associated with the natural environment are often context-specific (Bansal, 2005), we limited our study to a single industry: the ICT industry. We improved the reliability of the results by analyzing the data over five years: 2005-2009. To test the model, we used time series cross-sectional data analysis. This methodology is superior to analyzing cross-sectional data for a single time period because it controls for the confounding effect of time-invariant and company-specific variables (Wiersema & Bowen, 1997). In addition, to assess hypotheses 2 and 3, we used a multiple and moderated regression analysis introducing the moderating effect as a multiplicative variable (Cohen & Cohen, 1984). To create the multiplicative terms, we fixed both the independent and the moderating variables to their means to avoid multicollinearity (Venkatraman, 1989). We used the software package STATA. The results of the Housman specification test suggested that a random-effects model was appropriate.

4.5 Results

Table 4.1 reports descriptive statistics and correlations for the variables examined in our study. To investigate a possible autocorrelation bias, the Wooldridge test for autocorrelation in panel data was applied (Wooldridge, 2002). No significant serial correlation was detected.

With regard to heteroscedasticity, we conducted a likelihood-ratio test comparing GLS estimates under the assumption of heteroscedasticity with GLS estimates under the assumption of homoscedasticity. The result showed a presence of heteroscedasticity; as a result, we perform robust to heteroscedasticity estimations.

The results from the fixed-effects GLS regression analyses are listed in Table 4.2. Model 1 presents the results of the regression with the control variables and serves as a baseline model. As shown in the model, three of the control variables are statistically significant: North America and Year t-1 and domestic growth. On the contrary, in the baseline model, firm size and the firm's level of innovation do not have statistical significance for our sample.

Model 2 includes the direct effect of the independent variables analyzed. Hypothesis 1 suggests that a higher number of patented environmental innovations increases firm performance. Model 2 shows a positive effect of the number of patented environmental innovations, but the coefficient of the variable is not statistically significant. Although it is not mentioned by the hypotheses in this work, the direct effect of the geographic scope of the knowledge on the financial performance is positive and significant; however, the direct effect of the geographic scope of this exploitation is not significant.

Model 3 completes the analysis, including the moderating relationships. Model 3 shows that the international scope of the knowledge sourcing of patented environmental innovations moderates the relationship between the patented environmental innovation and firm performance, as stated in our hypothesis. Figure 4.1 helps to identify the nature of the moderating relation and shows that the moderating effect is so strong that the sign of the relationship between patented environmental innovation and financial performance changes with different levels of the geographic scope of each firm's knowledge sourcing. Here, the relationship is negative for a high geographic scope of environmental knowledge sourcing and positive when this scope is lower. Hypothesis 2 suggests this moderating effect of the sourcing of the environmental knowledge, but we proposed a different sign of the effect.

Finally, Model 3 shows that the geographic scope of the exploitation of patented environmental innovations moderates the relationship between patented environmental innovation and financial performance. Figure 4.2 shows that the moderating effect is again so strong that the relationship between patented environmental innovation and financial performance is negative when the geographic scope of the environmental knowledge exploitation is low but that this relationship is positive when the geographic scope of the knowledge is higher. These results completely support hypothesis 3.

Table 4.1 Descriptive statistics and correlations of variables used in analysis [a]

Variable	Mean	SD	1	2	3	4	5	6	7	8	9	10	11	12
1. Financial performance	2.13	1.15												
2. Firm size (ln)	9.87	1.45	.07											
3. Firm level of innovation	2.16	2.38	-.22*	.25*										
4. Domestic growth	1.26	3.45	.18†	.41***	.03									
5. North America[b]			.42***	-.01	-.23*	-.07								
6. Europe [b]			-.02	.31**	.05	-.05	-.41***							
7. Yeart-4[b]			.06	-.09	.07	.18†	.01	.00						
8. Yeart-3[b]			.10	-.03	.09	.26**	-.02	-.05	-.21*					
9 . Yeart-2[b]			.07	.03	-.01	.27**	.04	-.00	-.25**	-.25**				
10. Yeart-1[b]			-.19*	.05	-.03	-.12	.02	.01	-.25**	-.24**	-.28**			
11. Environmental patented innovation (ln)	2.26	1.39	-.40***	-.12	.56***	-.13	-.40***	-.00	-.03	.07	-..00	-.03		
12. Geographic scope of the knowledge sourc.	17.71	24.19	.05	.18†	.35***	.07	-.23*	.35**	.10	.05	.06	-.14	.36***	
13. Geographic scope of the exploitation	1.84	0.59	-.33***	.30**	.07	.10	-.30**	-.01	.15†	.09	.02	00	.16†	.01

[a] n=116.Table contains Pearson's correlation coefficient. Significant at the †.10; * .05; ** .01; *** .001 level.
[b] Categorical variable, so means and S.D. not shown

Table 4.2 Random effects robust GLS regression results [a]

	Model 1		Model 2		Model 3	
Control variables						
Firm size	– 0.14	(0.12)	– 0.16	(0.12)	– 0.20†	(0.12)
Firm level of innovation	– 0.02	(0.03)	– 0.03	(0.04)	– 0.03	(0.05)
North America	1.13**	(0.41)	1.01**	(0.39)	1.08**	(0.39)
Europe	–0.42	(0.52)	0.27	(0.52)	0.34	(0.49)
Domestic growth	0.07†	(0.04)	0.08*	(0.04)	0.10*	(0.04)
Year t-4	0.01	(0.30)	–0.07	(0.34)	–0.25	(0.33)
Year t-3	0.01	(0.30)	–0.05	(0.34)	–0.32	(0.33)
Yeart-2	0.06	(0.32)	–0.02	(0.33)	–0.29	(0.32)
Yeart-1	–0.55**	(0.19)	–0.55**	(0.22)	–0.73***	(0.21)
Direct effects						
Environt. patented innovation			–0.10	(0.08)	–0.10	(0.08)
Geog. scope of the knowledge sourcing			0.01**	(0.00)	0.01***	(0.00)
Geog. scope of the exploitation			–0.13	(0.12)	0.01	(0.10)
Moderating effects						
Environt. Patented innovation X. Geog. scope of the Knowledge sourcing					–0.01**	(0.00)
Environt. Patented innovation X. Geog. scope of the exploitation					0.23**	(0.10)
Constant	3.04*	(1.25)	3.34*	(1.48)	3.84**	(1.19)
R^2 within	.43		.45		.53	
Wald Chi2 (dl)	64.73***	(9)	70.44***	(12)	94.55***	(14)
Δ F			1.26	(3, 103)	8.59**	(2,101)

[a]n=116. Table contains unstandardized regression coefficients.
Robust standard errors are in parentheses.
Significant at the †.10; * .05; ** .01; *** .001 level. Tobin Q is the dependent variable.

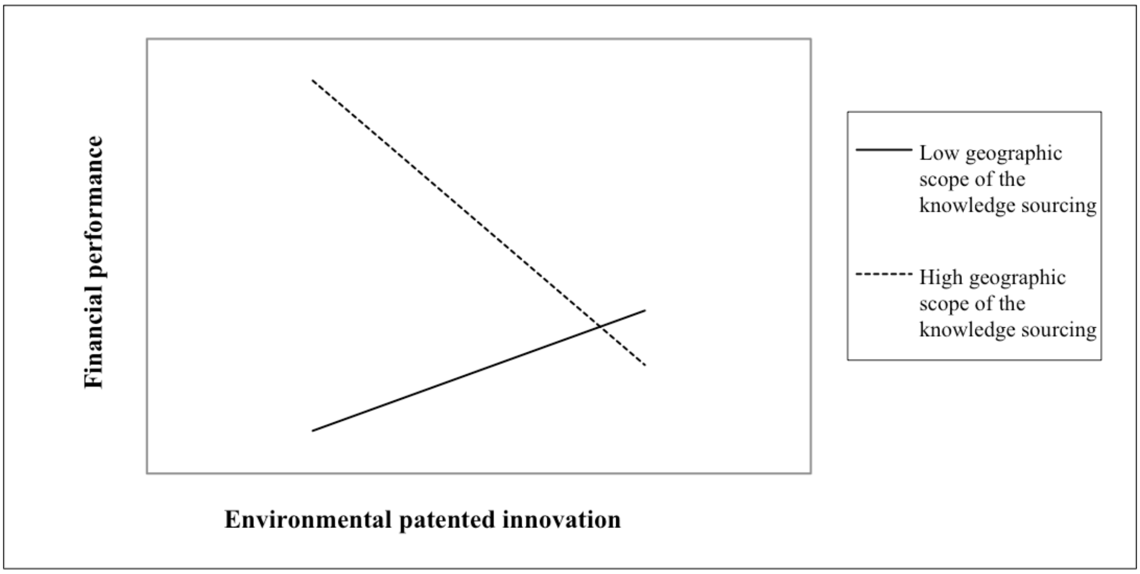

Figure 4.1 The combined effect of environment patented innovation and geographic scope of the knowledge

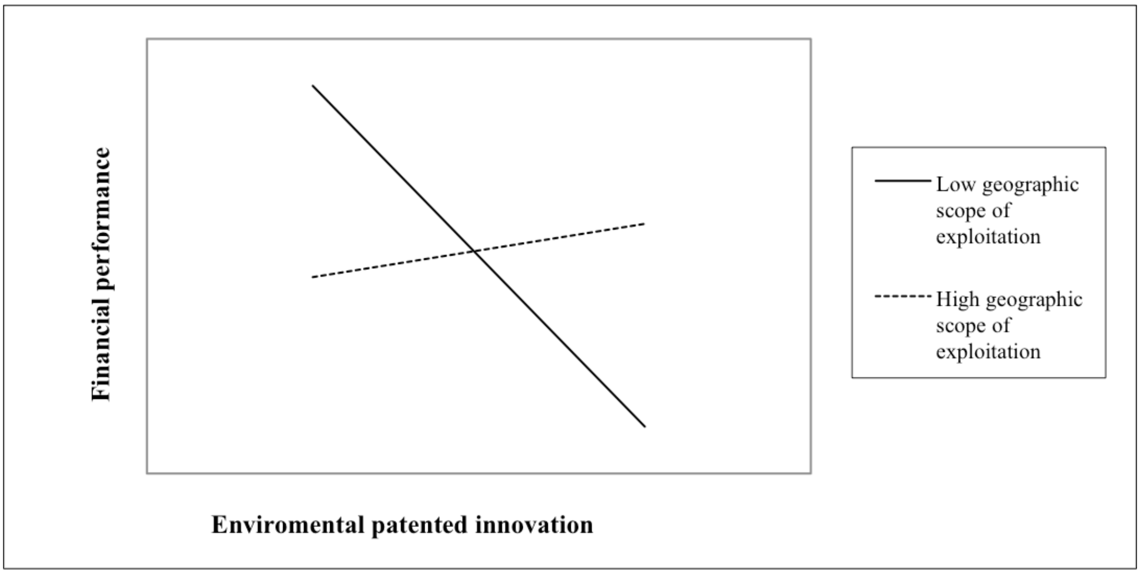

Figure 4.2 The combined effect of environment patented innovation and geographic scope of the exploitation

4.6 Discussion, Implications, and Future Research

Environmental innovations have generated increasing interest from policy makers and managers. The urgency of the environmental problems and the potential of environmental innovations to generate competitive advantages for firms and regions have been widely discussed (Porter & Linde, 1995a). However, the high requirements of investments to generate environmental innovations leads firms to require protection for their innovation returns over a long period in not only domestic but also international spheres.

Patents have gained growing attention at the societal, technological, political, and economical levels (e.g., David & Hall, 2006). Even when different systems of protecting environmental innovations may be available, in general, there are very few examples of economically significant inventions that have not been patented (Dernis & Guellec, 2005).

Our work has analyzed the implications of managers' selection of the international scope of the environmental patents on the relationship between the number of their firm's patented environmental innovations and its performance. Patents are among the most well-known tools of protecting the competitive advantage derived from a firm's innovations. However, the existing management literature has paid little attention to the potential influence of managers' selection of the international scope of the patents on their firm's financial performance. Our work contributes to extending the natural resource-based view (Hart, 1995) to analyze the importance of protecting a firm's environmental innovation against competitors and the influence of this innovation on firm performance. In contrast to previous works, which are mostly focused on managers' perception of the amount of generated innovation in the firm and its positive relationship with their firm's financial performance, we propose that patent protection and specific decisions regarding its international orientation - such as the international scope of the innovation sourcing or exploitation - are relevant to the sustainable generation of rents derived from the firms' innovative processes. We also complement the previous patent literature on proprietary patent strategies (Lippman & Rumelt, 2003), paying specific attention to the effects of managers' internationalization strategies regarding patents. More specifically, our results show several findings of interest.

First, we show a positive but non-statistically significant influence of the patented environmental innovations on the performance of our sampled firms. These results extend the previous literature finding the ability of environmental innovations to increase firm performance (Russo & Fouts, 1997; Sharma & Vredenburg, 1998; Shrivastava, 1995), paying specific attention to the opportunities of patented innovations. Although firms filing a patent must offer detailed and open information on the generated innovation and to make expenditures on the patenting process, the potential costs and knowledge spillover from available public information do not appear to be greater than the competitive advantages gained from securing the protection and exploitation of the generated returns of the innovations for the sampled firms.

The lack of a significant relevance of the identified positive relationship between the amount of patented environmental innovation and firm performance is an important issue to keep in mind. Although specific characteristics of the sampled industry may be influencing the lack of a significant importance of this relation between patent protection and a positive premium (Arora et al., 2008), we also believe that our results clearly suggest that analyzing complementary and contextual factors are relevant for a more precise delimitation of the connection between firms' environmental features and financial performance (Aragón-Correa & Sharma, 2003).

In fact, innovation in green technology is not always successful (Marcus, 2005; Marcus & Fremeth, 2009), and our results suggest that managers' decisions regarding how to exploit or protect these innovations (more than the innovations themselves) may be important for understanding the reasons. Our study of the selected international scope of the corporate environmental patents has tried to answer the calls in the existing literature to identify and analyze the conditions allowing firms to improve financial performance using environmental approaches.

Second, our results show that a broader international scope of knowledge sourcing of patented environmental innovation moderates the relationship between the number of patented environmental innovations and firm performance. Meanwhile, our results show a direction opposite to that originally hypothesized in hypothesis 2. When the sampled firms initially file to patent their innovations in more countries, a higher number of patented environmental innovations reduces the firms' financial performance. However, when the number of countries is lower, the connection between patented environmental innovations and firm performance is positive.

In this context, our results support that a wide international dispersion in the sourcing of knowledge may increase complexity beyond acceptable levels for effective communication and coordination without a proper reward (Kotabe et al., 2007). Our results suggest that the generation of a technological patent portfolio emerging from different countries does not result in synergies for the sampled firms. Even when countries vary in the type and nature of their innovation and industrial property activity due to differences in their innovation systems (Furman et al., 2002), these national differences may be not enough to generate differences in the environmental innovations in the global ICT industry selected for our sample.

To completely understand the differences between our findings and the original proposal regarding the sign of the moderation, it is also necessary to recognize that our measure of the sourcing of environmental knowledge may not be completely appropriate for our sample. The very global nature of the ICT industry may reveal that some sampled firms in the sampled industry are choosing the first place to file the patent for commercial reasons or due to a homogeneous standardized policy of filing first in an specific patent office (e.g., where the headquarter is) and not due to a connection with the R&D development process, as is common in most firms. It would be useful in future research to engage in a more detailed analysis of the process of patent generation for a more definitive understanding of the direction of the moderation.

Third, our results show that a broader international scope of selected regions for the exploitation of patented environmental innovation provides existing patents a greater potential to influence firms' financial performance. These results were predicted in hypothesis 3. Our work complements previous studies on the connection between international business and environmental issues (Darnall et al., 2008; Madsen, 2009; Pinkse & Kolk, 2012; Pinkse et al., 2010). It is particularly important that whereas the previous literature has emphasized the debate concerning whether international firms should use a standardized environmental approach in their current operations in different countries (Christmann, 2004; Dowell et al., 2000), our work focuses on a different perspective to analyze whether a broader international exploitation of specific patented environmental innovations makes financial sense.

Our results support the positive consequences of potential international scale economies (Carpano et al., 1994) and the efficiency of sampled innovations when it is possible for their international exploitation to lead to a greater proportion of all possible firm-specific innovations' rents with similar costs (Kirca et al., 2011). Our results also suggest that a broader international scope of exploiting the patented environmental innovations may reduce the risks of revenue fluctuations (Lu & Beamish, 2004), especially keeping in mind the changing nature of citizens' environmental preferences, the rapid development of technological and scientific environmental knowledge, and the uncertainty regarding the evolution of environmental regulations (Marcus et al., 2011).

Our developments in this chapter together have relevant implications for practitioners and policy makers. We strongly encourage the importance of developing specific approaches to protect the future exploitation of environmental innovations. Patents may be appropriate tools for managers' effective protection of many innovations. At the same time, we recognize that patents have some pros and cons related to the practical and specific requirements of granting a patent. In this context, our work emphasizes that it is important to obtain not only environmental patents but especially the specific conditions and management linked to these patents. For instance, our work highlights that a broader international scope of the places selected to exploit the patents appears to enhance the ability of a patent portfolio to improve the financial performance.

Policy makers have already signaled that environmental patents are now a priority in many Western countries (e.g., creating new specific codes for environmental patents). Our results show that not all firms are active in terms of generating patents in environmental innovation and policy makers should keep in mind the costs and difficulties of the process of filing and granting patents for different profiles of firms (e.g., large vs. small firms or firms with private development vs. universities or research centers). Our results also reflect interest in international approaches of patenting; therefore, regulators and policy makers may want to ease this process in at least two ways. First, in a scenario of growing international competence, regulators could offer collaboration with domestic firms in the international protection of their innovations. Second, local patent offices could offer a more coordinated effort to facilitate international protection. Even when there are already some transnational schemes available (e.g., World Intellectual Property Organization –WIPO- or EPO), these approaches do not yet offer a complete process, and the heterogeneity and complexity of the final international patenting process remain substantial.

Future research may want to pay more specific attention to the firm-level process of patent generation. It will be necessary to better understand the specific mechanics to make the decision of attempting the transformation (or not) of specific innovations in environmental patents. In addition, analyzing the degree and frequency of collaboration between different units of the firm developing environmental innovation and patent activity may represent a good proxy to understand the different knowledge loads and process complexity. When we have attempted to offer a more general perspective using data from multiple firms in a longitudinal period, future in-depth case studies may be more appropriate for understanding the specific features of the internal process.

References

Almeida, P., & Kogut, B. (1999). Location of Knowledge and the Mobility of Engineers in Regional Networks. *Management Science*, *45*(7), 905–917.

Aragón-Correa, J. A. (1998). Strategic Proactivity and Firm Approach to the Natural Environment. *Academy of Management Journal*, *41*(5), 556–567.

Aragón-Correa, J. A., & Sharma, S. (2003). A Contingent Resource-Based View of Proactive Corporate Environmental Strategy. *Academy of Management Review*, *28*(1), 71–88.

Archibugi, D., & Pianta, M. (1992). Specialization and Size of Technological Activities in Industrial Countries: The analysis of Patent Data. *Research Policy*, *21*(1), 79–93.

Arora, A, Ceccagnoli, M., & Cohen, W. (2008). R&D and the Patent Premium. *International Journal of Industrial Organization*, *26*(5), 1153–1179.

Arora, A., & Fosfuri, A. (2000). Wholly Owned Subsidiary Versus Technology Licensing in the Worldwide Chemical Industry. *Journal of International Business Studies*, *31*(4), 555–572.

Arora, A., Fosfuri, A., & Gambardella, A. (2001). Markets for Technology and their Implications for Corporate Strategy. *Industrial and Corporate Change*, *10*(2), 419–451.

Bansal, P. (2005). Evolving Sustainably: a Longitudinal Study of Corporate Sustainable Development. *Strategic Management Journal*, *26*(3), 197–218.

Barkema, H. G., & Vermeulen, F. (1998). International Expansion Through Start-up or Acquisition: A Learning Prespective. *Academy of Management Journal*, *41*(1), 7–26.

Barney, J. B., & Clark, D. N. (2007). *Resource-Based Theory: Creating and Sustaining Competitive Advantage*. *Journal of Marketing Management*. Oxford University Press.

Bartlett, C. A., & Ghoshal, S. (1989). *Managing Across Borders The Transnational Solution*. *Harvard Business School Press Books*. Harvard Business School Press.

Bilton, N. (2010). An Explosion of Mobile Patent Lawsuits - NYTimes.com. *New York Times*. *bits*. Available at http://bits.blogs.nytimes.com/2010/03/04/an-explosion-of-mobile-patent-lawsuits.

Cantwell, J. (2000). Technological Lock-in of Large Firms since the Interwar Period. *European Review of Economic History*, *4*(2), 147–174.

Cantwell, J., & Janne, O. (1999). Technological Globalisation and Innovative Centres: the Role of Corporate Technological Leadership and Locational Hierarchy. *Research Policy*, *28*(2-3), 119–144.

Carpano, C., J Chrisman, J., & Roth, K. (1994). International Strategy and Environment: An Assessment of the Performance Relationship. *Journal of International Business Studies*, *25*(3), 639–656.

Christmann, P. (2004). Multinational Companies and the Natural Environment: Determinants of Global Environmental Policy Standardization. *Academy of Management Journal*, *47*(5), 747–760.

Christmann, P., & Taylor, G. (2006). Firm Self-regulation Through International Certifiable Standards: Determinants of Symbolic Versus Substantive Implementation. *Journal of International Business Studies*, *37*(6), 863–878.

Chung, W., & Yeaple, S. (2008). International Knowledge Sourcing: Evidence from US Firms Expanding. *Strategic Management Journal*, *29*(11), 1207–1224.

Cohen, J., & Cohen, P. (1984). *Applied Multiple Regression: Correlation Analysis for the Behavioral Sciences*. (2nd ed.). Lawrence Erlbaum Associates.

Cohen, W. M., Nelson, R. R., & Walsh, J. P. (2000). *Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not)*. National Bureau of Economic Research. Available at http://www.nber.org/papers/w7552.

Czarnitzki, D., & Kraft, K. (2000). Haftungsregeln und Innovation, Legal Form and Innovation. *Journal of Economics and Statistics (Jahrbuecher fuer Nationaloekonomie und Statistik)*, *220*(5), 513–526.

Darnall, N., Henriques, I., & Sadorsky, P. (2008). Do Environmental Management Systems Improve Business Performance in an International Setting? *Journal of International Management*, *14*(4), 364–376.

David, P. A., & Hall, B. H. (2006). Property and the Pursuit of Knowledge: IPR Issues Affecting Scientific Research. *Research Policy*, *35*(6), 767–771.

De Burgos Jiménez, J., & Céspedes Lorente, J. J. (2001). Environmental Performance as an Operations Objective. *International Journal of Operations Production Management*, *21*(12), 1553–1572.

Dernis, H., & Guellec, D. (2005). *Using Patent Counts for Cross-country Comparisons of Technology Output*. Organisation for Economic Co-operation and Development. Available at http://www.oecd.org/dataoecd/26/11/21682515.pdf.

Dowell, G., Hart, S., & Yeung, B. (2000). Do Corporate Global Environmental Standards Create or Destroy Market Value? *Management Science*, *46*(8), 1059–1074.

Eckert, S., Dittfeld, M., Muche, T., & Rässler, S. (2010). Does Multinationality Lead to Value Enhancement? An Empirical Examination of Publicly Listed Corporations from Germany. *International Business Review*, *19*, 562–574.

European Commission. (2010a). *Europe's Digital Competitiveness Report*. European Commission. Available at http://ec.europa.eu/ceskarepublika/pdf/press/ks7ec2.pdf.

European Commission. (2010b). *The 2010 report on R&D in ICT in the European Union*. JCR Scientific and Technical Reports. European Commission. Available at http://ftp.jrc.es/EURdoc/JRC57808.pdf

Furman, J. L., Porter, M. E., & Stern, S. (2002). The Determinants of National Innovative Capacity. *Research Policy*, *31*(6), 899–933.

Goerzen, A., & Beamish, P. W. (2003). Geographic Scope and Multinational Enterprise Performance. *Strategic Management Journal*, *24*(13), 1289–1306.

Gomes, L., & Ramaswamy, K. (1999). An Empirical Examination of the Form of the Relationship Between Multinationality and Performance. *Journal of International Business Studies*, *30*(1), 173–187.

Grant, R. M. (1991). The Resource-Based Theory of Competitive Advantage: Implications for Strategy Formulation. *California Management Review,* (33), 114-135.

Greenpeace. (2011). *Guide to greener electronics. Version 17. Nov 2011*. Available at http://www.greenpeace.org/international/en/campaigns/climate-change/cool-it/Campaign-analysis/Guide-to-Greener-Electronics/?accept=84995a68d5f4ea201f3f289dd5fba946.

Greenwood, R., Diaz, A. M., Li, S. X., & Lorente, J. C. (2009). The Multiplicity of Institutional Logics and the Heterogeneity of Organizational Responses. *Organization Science*, *21*(2), 521–539.

Gurry, F., Fink, C., Khan, M., Bergquist, K., Lamb, R., Feuvre, B. Le, & Hao Zhou. (2012). *World Intellectual Property Indicators 2011*. WIPO Economics & Statistics Series. Available at http://www.wipo.int/ipstats/en/wipi.

Hair, J., Andersson, R. E., Tatham, R. L., & Black, W. (2009). *Multivariate Analysis* (6th ed.). Prentice Hall.

Haley, U. C. V., Schuler, D. A. (2011). Government Policy and Firm Strategy in the Solar Photovoltaic Industry. *California Management Review,* (54), 17-38.

Hall, B. H., Jaffe, A., & Trajtenberg, M. (2005). Market Value and Patent Citations. *Rand Journal of Economics*, *36*(1), 16–38.

Hamilton, L. C. (2006). *Statistic with Stata Updated for Version 9*. Thomson Brooks.

Hanel, P. (2006). Intellectual Property Rights Business Management Practices: A Survey of the Literature. *Technovation*, *26*(8), 895–931.

Harhoff, D., Scherer, F. M., & Vopel, K. (2003). *Exploring the Tail of Patented Invention Value Distributions*. Discussion Paper FS IV 97 - 27, Wissenschaftszentrum.

Hart, S. L. (1995). A Natural-Resource-Based View of the Firm. *Academy of Management Review*, *20*(4), 986–1014.

Hitt, M. A., Bierman, L., Uhlenbruck, K., & Shimizu, K. (2006). The Importance of Resources in the Internationalization of Professional Service Firms: The Good, the Bad, and the Ugly. *Academy of Management Journal*, *49*(6), 1137–1157.

Hitt, M. A., Hoskisson, R. E., & Kim, H. (1997). International Diversification: Effects on Innovation and Firm Performance in Product-diversified Firms. *Academy of management journal*, *40*(4), 767–798.

Hoskisson, R. E., Hitt, M. A., Johnson, R. A., & Moesel, D. D. (1993). Construct Validity of an Objective (Entropy) Categorical Measure of Diversification Strategy. *Strategic Management Journal*, *14*(3), 215–235.

Hytönen, H., Jarimo, T., Salo, A., & Yli-Juuti, E. (2011). Markets for Standardized Technologies: Patent Licensing with Principle of Proportionality. *Technovation*, *32*(9-10), 523–535.

Håkanson, L., & Nobel, R. (2001). Organizational Characteristics and Reverse Technology Transfer. *MIR Management International Review*, *41*(4), 395–420.

Iwasa, T., & Odagiri, H. (2004). Overseas R&D, Knowledge Sourcing, and Patenting: An Empirical Study of Japanese R&D Investment in the US. *Research Policy*, *33*(5), 807–828.

Kim, W. C., Hwang, P., & Burgers, W. P. (1993). Multinationals' Diversification and the Risk-return Trade-off. *Strategic Management Journal*, *14*(4), 275–286.

Kirca, A. H., Hult, G. T. M., Mena, J. A., & Miller, J. C. (2011). Firm-Specific Assets, Multinationality, and Financial Performance: A Meta-Analytic Review and Theoretical Integration. *Academy of Management Journal*, *54*(1), 47–72.

Kogut, B., & Zander, U. (1992). Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology. *Organization Science*, *3*(3), 383–397.

Kostova, T., Roth, K., & Dacin, M. T. (2008). Institutuonal Theory in the Study of Multinational Corporations: A critique and New Directions. *Academy of Management Review*, *33*(4), 994–1006.

Kotabe, M., Dunlap-Hinkler, D., Parente, R., & Mishra, H. A. (2007). Determinants of Cross-national Knowledge Transfer and its Effect on Firm Innovation. *Journal of International Business Studies*, *38*(2), 259–282.

Lahiri, N. (2010). Geographic Distribution of R&D Activity: How does it Affect Innovation Quality? *Academy of Management Journal*, *53*(5), 1194–1209.

Lemley, M. A., & Shapiro, C. (2005). Probabilistic Patents. *The Journal of Economic Perspectives*, *19*(2), 75–98.

Li, L., & Qian, G. (2005). Dimensions of International Diversification. *Journal of Global Marketing*, *18*(3-4), 7–35.

Lippman, S. A., & Rumelt, R. P. (2003). A Bargaining Perspective on Resource Advantage. *Strategic Management Journal*, *24*(11), 1069–1086.

Lu, J. W., & Beamish, P. W. (2004). International Diversification and Firm Performance: The S-Curve Hypothesis. *Academy of Management Journal*, *47*(4), 598–609.

Madsen, P. M. (2009). Does Corporate Investment Drive a "Race to the Bottom" in Environmental Protection? A Reexamination of the Effect of Environmental Regulation on Investment. *Academy of Management Journal*, *52*(6), 1297–1318.

Marcus, A. (2005). *Big Winners and Big Losers: The 4 Secrets of Long-term Business Success and Failure*. Wharton School Publishing.

Marcus, A. A., & Fremeth, A. R. (2009). Green Management Matters Regardless. *Academy of Management Perspectives*, *23*(3), 17–27.

Marcus, A., Aragón-Correa, J. A., & Pinkse, J. (2011). Firms, Regulatory Uncertainty, and the Natural Environment. *California Management Review*, *54*(1), 5-16.

Marcus, A., & Geffen, D. (1998). The Dialectics of Competency Acquisition: Pollution Prevention in Electric Generation. *Strategic Management Journal*, *19*(12), 1145–1168.

Margolis, J. D., & Walsh, J. P. (2003). Misery Loves Companies: Rethinking Social Initiatives by Business. *Administrative Science Quarterly*, *48*(2), 268–305.

Martin, N. L., & Mykytyn, P. P. (2009). Evaluating the Financial Performance of Business Method Patent Owners. *Information Systems Management*, *26*(3), 285–301.

Mazzoleni, R., & Nelson, R. R. (1998). The Benefits and Costs of Strong Patent Protection: A Contribution to the Current Debate. *Research Policy*, *27*(3), 273–284.

Nachum, L., & Zaheer, S. (2005). The Persistence of Distance? The Impact of Technology on MNE Motivations for Foreign Investment. *Strategic Management Journal*, *26*(8), 747–767.

Nayyar, P. R. (1993). Performance Effects of Information Asymmetry and Economies of Scope in Diversified Service Firms. *Academy of Management journal*, *36*(1), 28–57.

Penner-Hahn, J., & Shaver, J. M. (2005). Does International Research and Development Increase Patent Output? An Analysis of Japanese Pharmaceutical Firms. *Strategic Management Journal*, *26*(2), 121–140.

Peters, B., & Schmiele, A. (2011). *The Contribution of International R&D to Firm Profitability*. ZEW – Centre for European Economic Research. Available at http://ssrn.com/abstract=1740464.

Peters, M., Schneider, M., Griesshaber, T., & Hoffmann, V. H. (2012). The Impact of Technology-push and Demand-pull Policies on Technical Change – Does the Locus of Policies Matter? *Research Policy*, *41*(8), 1296–1308.

Petruzzelli, A. M. (2011). The Impact of Technological Relatedness, Prior Ties, and Geographical Distance on University–industry Collaborations: A Joint-patent Analysis. *Technovation*, *31*(7), 309–319.

Pinkse, J., & Kolk, A. (2012). Addressing the Climate Change—Sustainable Development Nexus. *Business Society*, *51*(1), 176–210.

Pinkse, J., Kuss, M. J., & Hoffmann, V. H. (2010). On the Implementation of a "Global" Environmental Strategy: The Role of Absorptive Capacity. *International Business Review*, *19*(2), 160–177.

Pisano, G. (2006). Profiting from Innovation and the Intellectual Property Revolution. *Research Policy*, *35*(8), 1122–1130.

Porter, M. E., & Van Der Linde, C. (1995a). Toward a New Conception of the Environment-Competitiveness Relationship. *Journal of Economic Perspectives*, *9*(4), 97–118.

Porter, M. E., & Van Der Linde, C. (1995b). Green and Competitive: Ending the Stalemate. *Harvard Business Review*, *73*(5), 120–134.

Raghu, T. S., Woo, W., Mohan, S. B., & Rao, H. R. (2007). Market Reaction to Patent Infringement Litigations in the Information Technology Industry. *Information Systems Frontiers*, *10*(1), 61–75.

Reitzig, M. (2004). Improving Patent Valuations for Management Purposes—Validating new Indicators by Analyzing Application Rationales. *Research Policy*, *33*(6-7), 939–957.

Rugman, A. M. (2005). *The Regional Multinationals: MNEs and "Global" Strategic Management*. Cambridge University Press.

Rugman, A. M., & Verbeke, A. (2004). A Perspective on Regional and Global Strategies of Multinational Enterprises. *Journal of International Business Studies*, *1*(35), 3–19.

Russo, M. V., & Fouts, P. A. (1997). A Resource-Based Perspective on Corporate Environmental Performance and Profitability. *Academy of Management Journal, 40*(3), 534-559.

Sharma, S., & Vredenburg, H. (1998). Proactive Corporate Environmental Strategy and the Development of Competitively Valuable Organizational Capabilities. *Strategic Management Journal*, *19*(8), 729–753.

Shrivastava, P. (1995). The Role of Corporations in Achieving Ecological Sustainability. *Academy of Management Review*, *20*(4), 936–960.

Somaya, D. (2012). Patent Strategy and Management: An Integrative Review and Research Agenda. *Journal of Management*, *38*(4), 1084–1114.

Subramaniam, M., & Venkatraman, N. (2001). Determinants of Transnational new Product Development Capability: Testing the Influence of Transferring and Deploying Tacit Overseas Knowledge. *Strategic Management Journal*, *22*(4), 359–378.

Teece, D. J. (1986). Profiting from Technological Innovation: Implications for Integration, Collaboration, Licensing and Public Policy. *Research Policy*, *15*(6), 285–305.

Venkatraman, N. (1989). The Concept of Fit in Strategy Research: Toward Verbal and Statistical Correspondence. *Academy of Management Review*, *14*(3), 423–444.

Wiersema, M. F., & Bowen, H. P. (1997). Empirical Methods in Strategy Research: Regression Analysis and the Use of Cross-section versus Pooled Time-series, Cross-section Data. *Statistical Models for Strategic Management*, 201–219.

WIPO. (2004). *WIPO Intellectual Property Handbook: Policy, Law and Use*. World Intellectual Property Organization (WIPO). Available at http://www.wipo.int/about-ip/en/iprm.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

Zander, I. (1999). How do you Mean "Global"? An Empirical Investigation of Innovation Networks in the Multinational Corporation. *Research Policy*, *28*(2-3), 195–213.

CAPÍTULO 5

ADDING SENSE TO PATENT ONTOLOGIES: A REPRESENTATION OF CONCEPTS AND REASONING

ADDING SENSE TO PATENT ONTOLOGIES: A REPRESENTATION OF CONCEPTS AND REASONING

5.1 Introduction

Innovation is an important driver of both firm success and national economic growth (Porter, 1990). In the context of a growing necessity for investments that develop innovation, patents are relevant tools for protection of the firm's innovative ideas. Patents are legal documents that protect the rights of the inventor of an industrial property. The total number of patent documents filed in 2010 worldwide is estimated at approximately 1.98 million (Gurry et al., 2012) and the cost of litigation related to patent violations is between € 400.000 and € 2.5 million in the USA and between €50.000 and €500.000 in Europe (Buzzacchi & Scellato, 2008).

The availability of online patent information has been also enhanced in recent years with attempts to provide more useful information for decision-making. A patent document provides different information about the innovation: the firm generating it, location, date, the technological field, and some information about the other patents that it cites, among others. All of these data that describe or are related to patent documents are called patent metadata (Giereth et al., 2006).

Patent documents are usually stored in large databases of patent offices (e.g., those maintained by the European Patent Office (EPO) and those under supervision of the United States Patent and Trademark Office (USPTO). These databases often use different data structures that make it difficult to interoperate or automatically and efficiently process the information contained therein. This difficulty is even more important in the context of the different objectives (informative and infringement searches, etc.) of the agents using these databases (e.g., patent lawyers, managers of start-up firms, scientists, and managers of competitors in the industry) (Foglia, 2007).

In this context, ontologies have proved to be useful for sharing information by providing formal, uniform and shareable representations about a domain. Several pieces of work have proposed the use of an ontology-based approach to represent patent metadata using the Web Ontology Language (OWL) (e.g., Taduri et al., 2011; Giereth et al., 2006). The main objective of these works is to provide a semantically, well-defined and homogeneous representation of the main types of patent metadata. The use of ontologies enables the representation of knowledge and the identification of context and dependency information more easily than using database-centric structures and interfaces (Giereth et al., 2006). Likewise, ontologies are increasingly being used in several domains related to innovation and patent registration (Ahmet et al., 2011), including economic and management business areas such as logistics (Anand et al., 2012), knowledge-commerce (Chen, 2011), team organization (Wi et al., 2011) or business processes (Noguera et al., 2010a; Pelegrina et al., 2010). Furthermore, the power of reasoning in ontologies allows different domains represented through ontologies, such as patents associated with economics and management studies, to be integrated (Seol et al., 2011).

The technological field of a patent document represents relevant metadata information about the patent. Patent databases use codes for technological fields that assume a hierarchical classification that delimits the categories a patent may pertain to or be associated with. This field is widely used in searches in databases to determine the field(s) in which a firm may infringe upon another company's industrial rights or where there exists a gap in the technology in which a company could innovate. However, in the patent databases, this hierarchy is not explicitly described and cannot be automatically processed by computers.

Previous patent ontologies have also represented technological fields by using codes; however, they do not fully exploit the formal representation of patent classification hierarchies of these technological fields, and basically mirror the technological patent codes of patent offices databases without leveraging further reasoning capabilities.

In summary, the motivations of this chapter in relation to patent information and ontologies are:

- Technological patent codes are defined in patent databases according to hierarchical classifications of technological patent fields, but these hierarchies are not explicitly specified for automatic processing by computers.

- Previous pieces of work on patent metadata ontologies do not represent the hierarchies of the technological patent codes either.

- Currently, it is not possible to infer transversal relationships between the different metadata that describe a patent document through the hierarchies of concepts of technological patent codes in current patent ontologies.

This chapter proposes a method to automatically build and populate patent metadata ontologies by indexing hierarchical codes, which can be retrieved from different patent repositories, by defining ontological categories which enrich the information retrieval process with new relationships, properties and enable the inference of new knowledge. In particular, the chapter first studies the characteristics of the hierarchical codes, such as the structure they follow. Second, we propose a representation of these hierarchical codes in OWL. And third, the hierarchy of the technological patent codes is translated from XML into the formal Web Ontology Language OWL 2 using XSLT (eXtensible Stylesheets Transformation Language) (W3C, 1999), according to the characteristics of the hierarchical codes, and the representation proposed.

Additionally, we show the power of reasoning that OWL offers in combination with the hierarchy of concepts (OWL classes) that we have created to infer new information in different fields through transversal relations among concepts from different (but related) domains. The aim is to benefit from linking patent information with external information by providing several methods to infer knowledge in different fields and to connect different knowledge sources.

To illustrate the applicability of our proposal, this chapter shows how new information can be inferred from the hierarchy of patent metadata concepts. In particular, the first case study demonstrates how a suitable ontological representation of patent metadata enables the automatic reclassification of patents when a new technological patent code appears. It will be illustrated with an example consisting in the introduction of a new technological patent field and corresponding code. In particular, a new technological field that will include environmentally friendly patents will be introduced. In the second case, a proposal is introduced to link two different knowledge domains by specifying new relationships between the representation of patent classifications and an external classification, the classification of industrial sectors provided by the United Nations Statistics Division (UNSD)[10], which can help in

[10] http://unstats.un.org/unsd/cr/registry/regcst.asp?Cl=27

the elaboration of financial studies that evaluate the innovation level of firms based on patent indicators.

The remainder of the chapter is organized as follows. Section 5.2 describes the related work. Section 5.3 explains how to translate the hierarchical technological patent codes into hierarchies of concepts and accomplish the population of the ontology. Section 5.4 presents the reasoning and the inference of new information from the hierarchical classes within the patent domain and in multiple domains. Finally, Section 5.5 concludes the chapter by discussing the contributions of the research.

5.2 Background

5.2.1 Patent Ontologies

Several patent ontologies have been proposed so far for a semantically well-defined and homogeneous representation for the major types of patent metadata. The most prominent examples are the ontology created within the European *Patexpert* project (Giereth et al., 2006; Wanner et al., 2009) and the *PatentOntology* from Stanford University (Taduri et al., 2011).

Patexpert was created to homogeneously represent different patent information from several EPO databases and to provide it with semantic meaning. However, Patexpert does not merge information retrieved from different patent offices. The patent metadata ontology has been populated by XSLT stylesheets. Unfortunately, the public version of this ontology[11] is not populated, but to the best of our knowledge this ontology does not automatically represent the semantics of the hierarchy of technological patent codes.

[11] http://mklab.iti.gr/project/patexpert

PatentOntology was developed to avoid the limitations of Patexpert when integrating heterogeneous domains (Taduri et al., 2011b). PatentOntology merges information from USPTO patent documents retrieved from the USPTO database with information from patent courts of USPTO from the LexisNexis database (Lexis, 2012). Although this ontology has been populated with a parser, it does not automatically retrieve the semantics of the hierarchy of technological patent codes and does not merge information from different patent offices. The technological patent codes classify innovations into fields of activity to facilitate the searches of interested agents. This technological field is widely used to delimit the scope of the searches and is one of the most used items of patent metadata (Foglia, 2007).

There are also patent ontologies based only on keywords found in patents (Choi et al., 2012; Seol et al., 2011; Soo et al., 2006). These ontologies allow retrieving information from the text of the patent documents, however they do not represent patent metadata and, hence, they do not exploit the potential of reasoning using the technological patent codes.

5.2.2 Technological Patent Classifications

Different classifications of technological patent codes exist, such as the US classification or the European ECLA (European CLAsification) and ICO (In Computer Only) classifications. The World Intellectual Property Organization (WIPO) has also defined the International Patent Classification (IPC), and most of the patent offices classify patents with their own classification as well as with IPC. This chapter will use the IPC codes for general application, and the ICO codes, which are based on the IPC codes and have the same underlying structure, for a specific classification of "green patent technology". We will explain this term in Section 5.4.1.

IPC codes are divided into 5 parts:

1. Section, which is represented by a letter (e.g., *A*, meaning *human necessities*, to *H*, for *electricity*).
2. Class, which is represented by 2 digits.
3. Subclass, which is represented by one letter.
4. Main group, represented by 1 to 3 digits.
5. Group, represented by at least 2 digits.

For example, the IPC code *H04L25/02* is composed of section (*H*), class (*04*), subclass (*L*), main group (*25*) and group (*02*). As such, the complete code *H04L25/02* means:

H	*electricity*
H04	*electric communication techniques*
H04L	*transmission of digital information*
H04L25	*baseband systems and*
H04L25/02	*shaping networks.*

5.2.3 Patent Ontologies Population

The ontologies need to be populated with information retrieved from patent repositories. There have been some efforts to automatize this process. In particular, certain work has extracted OWL documents from XML documents with the help of XSLT. While some of this work uses the XML schema (Bedini et al., 2010; Ferdinand et al., 2004), other work only creates the OWL model (Bosch & Mathiak, 2011) and others the OWL model and instances (Lacoste et al., 2011).

Other work has also been published on translating XML into OWL and the development of visual tools such as JXML2OWLMapper (Cardoso & Bussler, 2011; Rodrigues et al., 2008) or the online XMLtoOWL[12] tool (Bohring & Auer, 2005). These tools enable the visual assignment of XML labels of the XML instances or schema into OWL labels of the OWL model.

However, none of these works combines information retrieved from different sources or provides support for hierarchical codes. To fully exploit the IPC codes, this chapter proposes a method for extracting each part of the IPC codes and providing each of them with semantic. In particular we will retrieve the patent metadata in XML documents from the patent databases, and translate them into OWL classifiying the hierarchical codes in ontological categories.

5.3 An Approach for the Analysis of Hierarchical Codes Using Patent Ontologies

In this section, we present a patent description and analysis method based on an ontological representation of technological patent codes. The method begins by splitting the codes of the technological patent fields into their constituent parts (parts of the code) that aid in inferring a hierarchy of concepts as a final output; as far as we know, previous patent ontologies have not fully exploited the potential to automatically create a hierarchy of concepts in an ontology from the hierarchy represented in codes.

To automatically create a classification hierarchy of concepts in a patent ontology, this method includes the following steps:

1. The first step is to study the technological patent codes (IPC) and identify certain characteristics; this study considers two important characteristics of

[12] http://www.semanticscripting.org/XML2OWL_XSLT

technological patent codes, which although simple, they seem to have been disregarded in previous patent ontologies literature.

2. The second step is to provide a patent ontology for the analysis of the hierarchical technological patent codes, addressing the features found in step 1, which is covered in Section 5.3.2.

3. The third step is to propose an automatic mechanism to populate the patent ontology, which is discussed in Section 5.3.3.

5.3.1 Characteristics of the Technological Patent Codes

5.3.1.1 Importance of the Parts of the Codes

The first important characteristic that IPC codes exhibit is related to the meaning of each part of the codes. When dividing the code into its parts, the meaning of the individual parts can be different even when their representations are the same. For example, the IPC code *H03K3/03* means:

H	*electricity*
H03	*basic electronic circuitry*
H03K	*pulse technique in basic electronic circuitry*
H03K3	*circuits for generating electric pulses*
H03K3/03	*astable circuits for generating electric pulses*

This code is defined as "*dealing with astable circuits for generating electric pulses*". Dividing it into its parts -section (*H*), class (*03*), subclass (*K*), main group (*3*), subgroup (*03*)- we can see that the class and the subgroup have the same value (*03*). Although they have the same representation (*03*), they have different meanings because of their respective position in the code. One represents the class "*basic electronic* *circuits*", and the other one represents the subgroup "*astable circuits for generating electric pulses*".

Therefore, when translating IPC codes into a hierarchy of concepts, one characteristic is that the same individual (in the example, *03*) has different meaning depending on the part of the code to which it belongs (in the example, the class or the subgroup).

5.3.1.2 Importance of the Context of the Codes

Another important characteristic found in this study of IPC codes is that the meaning of the representation of one part of the code depends on the previous parts.

For example, lets consider the IPC code *H02K3/02,* where:

H	*electricity*
H02	*generation, conversion or distribution of electric power*
H02K	*dynamo electric machines*
H02K3	*details of winding in dynamo machines*
H02K3/02	*winding characterized by the conductor material*

By splitting the code *H02K3/02* into its parts -section (*H*), class (*02*), subclass (*K*), main group (*3*), and subgroup (*02*)- and comparing the codes in both examples, (*H03K3/03*) and (*H02K3/02*), we see that both of them have the subclass (*K*), but the meaning of the subclass (*K*) is different in each case because its meaning depends on the previous parts of the code.

Therefore, we have found another characteristic that must be taken into account when translating IPC codes into a hierarchy of concepts; that is, the need to put the code into context to understand the meaning of each part. In particular, it is necessary to consider not only the value of a particular part of the code, but also the values of the previous parts.

Hence, to create a hierarchy of concepts in the ontology, it is crucial to make an accurate translation of the hierarchical codes

5.3.2 HCOntology

This subsection provides a translation the IPC hierarchy of the technological patent codes into a hierarchy of concepts in an ontology. The concepts of the ontology are used to describe the semantics of the parts of the codes and to make their meaning explicit.

First, a hierarchical classification in the ontology has been designed taking into account all of the IPC sections (*IPC_A, IPC_B, IPC_C, IPC_D, IPC_E, IPC_F, IPC_G* and *IPC_H*). Next, for each IPC section, the subclasses for all IPC classes of the corresponding section (e.g., *IPC_H03*) have been created. In the same way, the IPC subclasses (e.g., *IPC_H03K*) and main groups (e.g., *IPC_H03K03*) have been created, and the subgroups are the individuals (e.g., *H03K03_03*), creating a large hierarchy of classes and individuals (from now on referred to as HCOntology – Hierarchical Codes Ontology; see Figure 5.1). The ontologies in the examples have been implemented by using the ontology editing tool Protégé (Noy et al., 2000; Sirin et al., 2007).

Thus, we make explicit the hierarchy found in the technological patent codes and enable automatic processing, filling the gaps in the previous patent metadata ontologies.

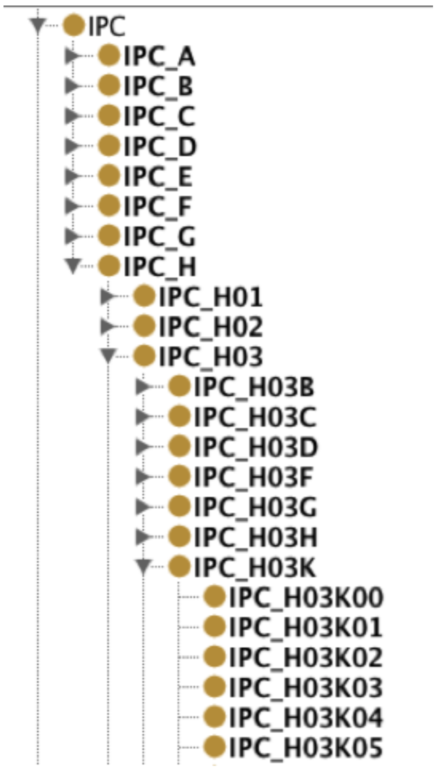

Figure 5.1 HCOntology in Protégé

This design is intended to address the two characteristics mentioned before in this section. The hierarchy of classes in HCOntology considers the first characteristic (Subsection 5.3.1.1) because one part of the code is a child class belonging to the previous part of the code, and therefore, the meaning of both parts are different. In the example of Subsection 5.3.1.1, the individual *H03K03_03* represents the subgroup (*03*), and *IPC_H03* represents the class (*03*). However, because the individual *H03K03_03* inherits *IPC_H03* and they belong to different classes, the meanings of both are well differentiated.

The proposed HCOntology also addresses the second characteristic. In the example of Subsection 5.3.1.2, the class *IPC_H03K* inherits the properties of the superclass *IPC_H03*, and another class, *IPCH02K*, inherits the meaning of the superclass *IPC_H02*. Therefore, the meanings of the two codes are well differentiated because the hierarchy of classes takes into account the values of the previous parts of the code.

Furthermore, HCOntology allows reasoning and using the classes and the semantics of their hierarchy. For example, the individual *H03K03_03* inherits all of the properties from its parent classes, *IPC_H03K03, IPC_H03K, IPC_H03* and *IPC_H*. Hence, this individual will be found in the searches of IPC codes *H03* because it belongs to this parent class.

After introducing the new hierarchy of classes, in the next section, we will implement an automatic mechanism that populates the ontology from query responses of the patent databases.

5.3.3 Populating the Patent Ontology and HCOntology

Patent databases provide their results to queries in different formats, one of which is XML. We took advantage of this result format to create and populate a patent ontology with the contents of a database using stylesheets.

Stylesheets are XSL documents that with the help of an XSLT processor can convert XML files into OWL files. This section presents a technical solution that automates the process of populating the ontology, parsing the query results from different databases in XML format, and the XML resulting files are converted into OWL by means of the corresponding XSL files and an XSLT processor. Figure 5.2 shows an overview of the system proposed. The first step of this process takes a XML file from the results of a search in the patent database. In the second step, a stylesheet to translate the XML tags into ontology classes and properties, including the IPC classes and subclasses, is implemented. In the third step, with the help of an XSLT processor, the XML individuals are translated into OWL individuals using our XSL file.

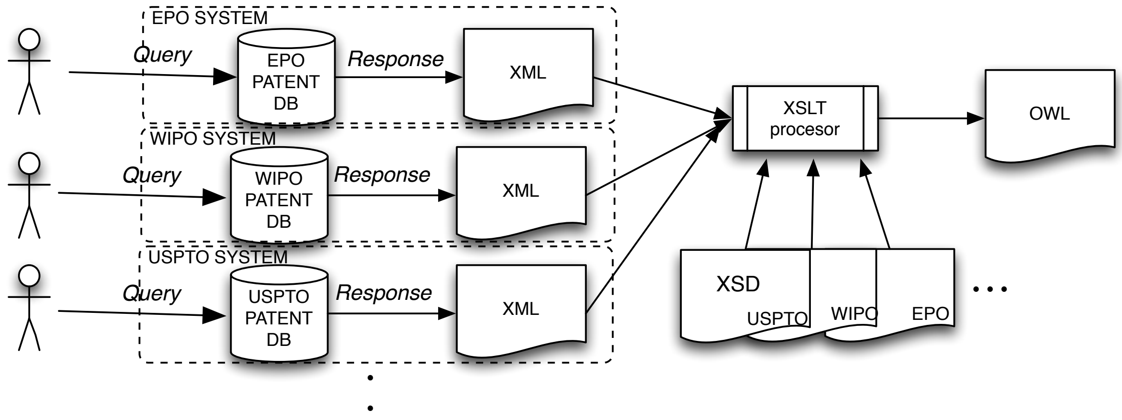

Figure 5.2 System overview

The second step requires to address the correspondence between the XML tags and the OWL classes using a stylesheet (Wanner et al., 2009), and to extract the individuals used to populate the ontology. In the case of HCOntology, we will also need XPATH (XML Path language) (Malhotra et al., 2007) to point at the different parts of the code. Additionally, this step allows the extraction of each part of the code one-by-one, navigates inside HCOntology, and inserts the IPC code in the class to which it belongs. We will describe this stylesheet in detail in next Subsection 5.3.3.1, and provide and example of stylesheet in Subsection 5.3.3.2.

5.3.3.1 Customization of Stylesheets for Hierarchical Codes

The resulting stylesheets will allow the automatic translation of XML instance documents (where the instances are hierarchical codes) into an OWL ontology model and instances with the process shown in Figure 5.2. Previous work has created XSD documents for translating XML instances into OWL models and/or instances (Giereth et al., 2006; Taduri et al., 2011). However, none of these proposals has studied the hierarchical codes. While other work translates one XML label (instance) into one OWL label (class or instance), our work, on the other hand, proposes the translation of one XML label (instance) into several OWL labels (several classes and one instance) following an implicit hierarchy in the XML label.

The customization of the stylesheets consists of the following steps:

1. Study the codes and define their parts, identifying the number of characters or digits in each part and whether they have separating characters
2. Study the structure of the labels in the XML and OWL files
   - 2.1. Location of the hierarchical codes in the XML file
   - 2.2. Detect the class from which implementing HCOntology in the OWL file
3. Write the XSL file
   - 3.1. Write the regular header of the OWL files with the namespaces. If HCOntology is implemented on top of an existing ontology, it is also needed to import the existing ontology in the header.
   - 3.2. Clean each label of the XML file, deleting unnecessary spaces, ensuring that each code is only written once in the OWL file, even if the code is repeated in several places in the XML file, etc.
   - 3.3. Create the hierarchical code structure
     - 3.3.1. Define the whole code except the last part as a subclass of the previous part of the code
     - 3.3.2. Repeat Step 3.3.1 for each part of the code until the first part of the code, and define the first part of the code as a subclass of the class of step 2.2
     - 3.3.3. Insert the individuals (the whole code) in the corresponding class of Step 3.3.1
4. Close the open labels

5.3.3.2 Example of Customization of Stylesheets

This section shows the applicability of the customization of the stylesheets proposed in the previous subsection. The EPO Open Patent Services database is used in this case study to download the XML file. We need to use the XPATH functions related to substrings to split the ICO codes into its parts and create the corresponding ontology classes, subclasses and individuals for each code. As we mention in Section 5.2, the ICO codes follows the same structure as the IPC codes. Following the previous section, creating a XSL file to represent the ICO codes needs:

1. Study the ICO codes (for example the ICO instance in the XML file, *Y02E10:20*). The ICO codes have the following structure:

    - Section: one character (*Y*)
    - Class: two digits (*02*)
    - Subclass: one character (*E*)
    - Main group: from one to three digits (*10*)
    - Group: at least two digits (*20*)
    - The mark ":" separates the main group from the group in the XML file

2. Study the structure of the labels in the XML and OWL files

    2.1. In the XML file, the ICO codes are stored under the labels: *<RESULT-LIST> <ROW> <ICO><p>*

    2.2. In the OWL file, implement HCOntology as subclasses of http://www.semanticweb.org/HCOntology#*ICO*

3. Build the XSL file

3.1. In the header of the OWL file, write, as usual in OWL files, the RDF namespace envelope and the ontology element (Smith et al., 2004)

3.2. Clean the labels in the XML file that contains the hierarchical code, delete the empty spaces and define each code only once

3.3. Create the hierarchy of classes of the ICO codes. For example, as shown in Figure 6.3, the code *Y02E10:20* should have an OWL instance *ICO_Y02E10-20*, with the previous OWL hierarchy of classes (*ICO_Y, ICO_Y02, ICO_Y02E, ICO_Y02E10*). Figure 6.4 shows the result of this step in the XSD file

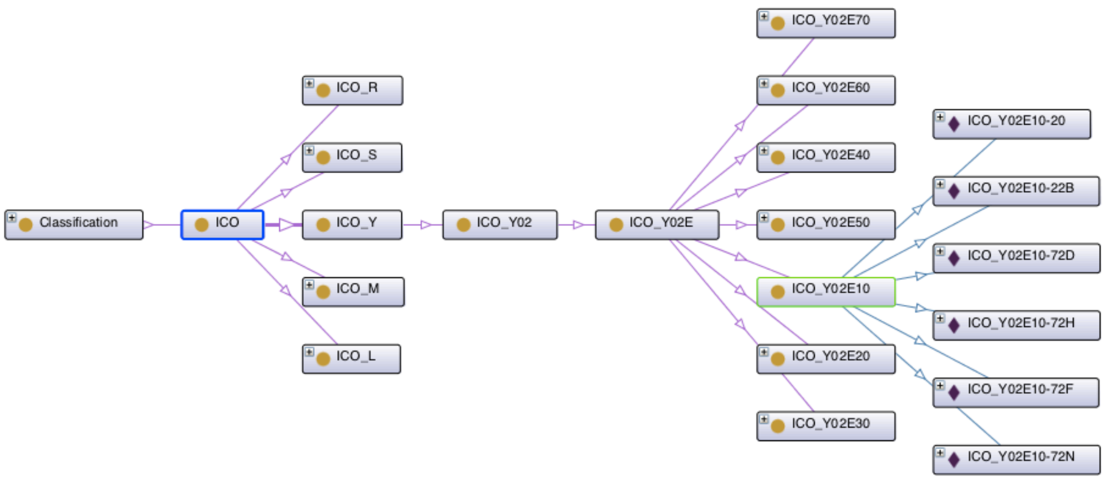

Figure 5.3 Hierarchy of classes for the ICO code Y02E10/20 in Protégé

3.3.1. Define the whole code except the last part of it as a subclass of the previous part of the code. For example, from the XML instance *Y02E10:20*, take the part *Y02E10* and make it a subclass of *Y02E*, thereby creating the class *ICO_Y02E10*. Afterwards, create the previous subclass (*ICO_Y02E*) and then define one as a subclass of the other (*ICO_Y02E10* subclass of *ICO_Y02E*)

3.3.2. Repeat Step "3.3.1" with the rest of the code using suitable XPATH functions. In the example, imply the creation of the class *ICO_Y0E* and define it as a subclass of *ICO_Y02*, create the class *ICO_Y02* and define it as a subclass of *ICO_Y*, create the class *ICO_Y* and define it as a subclass of *ICO*

3.3.3. Add the individual (the instance) in the corresponding class. In the example, imply by adding the instance *ICO_Y02E10-20* to the class *ICO_Y02E10*

```
<xsl:variable name="var8" select="concat('ICO_',substring-before(normalize-space(.),':'))"/>
    <owl:Class rdf:about="{$var8}">
        <rdfs:subClassOf rdf:resource="{concat('ICO_',substring(normalize-space(.),1,4))}"/>
    </owl:Class>
    <owl:Class rdf:about="{concat('ICO_',substring(normalize-space(.),1,4))}">
         <rdfs:subClassOf rdf:resource="{concat('ICO_',substring(normalize-space(.),1,3))}"/>
    </owl:Class>

    <owl:Class rdf:about="{concat('ICO_',substring(normalize-space(.),1,3))}">
         <rdfs:subClassOf rdf:resource="{concat('ICO_',substring(normalize-space(.),1,1))}"/>
    </owl:Class>

    <owl:Class rdf:about="{concat('ICO_',substring(normalize-space(.),1,1))}">
         <rdfs:subClassOf rdf:resource=" http://www.semanticweb.org/HCOntology#ICO "/>
    </owl:Class>
    <owl:NamedIndividual rdf:about="{concat('ICO_',substring-before(normalize-space(.),':'),
                                     '-',substring-after(normalize-space(.),':'))}">
         <rdf:type rdf:resource="{concat('ICO_', substring-before(normalize-space(.),'/'))}"/>
    </owl:NamedIndividual>
```

Figure 5.4 Excerpt of XSD file implementing step 3.3

3.4. Close all the labels that remain opened

With these stylesheets, the XSLT processor will automatically create a hierarchy of codes in OWL from a single code in XML, enabling future changes in the hierarchies. Technological patent codes represent technologies, and as the technologies evolve or new technologies appear, new technological patent codes appear. The method proposed could be used not only with current technological patent codes but also with future technological patent codes, without changing the stylesheet.

5.4 Exploiting the Ontology: Reasoning through Ontology Hierarchy

The advantages of having the hierarchical codes represented in a hierarchy of classes in OWL lies in that the meaning of all parts of the code are explicitly available. It is therefore possible to automatically process each part of the code.

These study cases will show the use of the meaning of parts of the code by inferring new information (or new classification of patents) by using parts of the technological patent codes. In particular, this section illustrates how the proposed method enables relationships to be created between concepts using the full semantics of the technological patent codes.

In particular, we will present two study cases. The first case occurs within the patent domain; we will show how to automatically reclassify patents when a new code appears (Section 5.4.1). The second case study will show how to automatically merge two knowledge domains: the patent domain and the financial domain (Section 5.4.2). In this case study, we will show how to automatically reclassify patents not only with the technological patent code of the patent domain but also with the industrial codes of the financial domain[13].

5.4.1 Introducing a New Technological Patent Code

As the technology evolves, new fields are discovered and patents must represent these new technological fields with their codes. Usually, the need for the new codes

13 http://unstats.un.org/unsd/cr/registry/regcst.asp?Cl=27

appears after a group of patents in this field have been filed and classified in other related technological fields. Our study case develops the same method used by the EPO to populate this new technological patent code, which is the same as the method used to search for patented inventions (Goebel & Hintermaier, 2011). The EPO explores patents by searching for particular IPC code, or a particular keyword found in patent documents, or a combination of both (IPC codes and keywords), to reclassify them with the new code.

For example, beginning with the *Y02* term, the European Patent Office (EPO) has new ICO (In-Computer-Only) codes and sub-classifications to cover "*technologies or applications for mitigation or adaptation against climate change*", or more commonly known as green patent technology. To prove the automatic reclassification of new codes, we reclassify patents as green patent technology using the same method as the EPO, with IPC codes and keywords, but we will use example patents that have not yet been classified as green patent technology by the patent authorities. With this approach, we will infer new information (ICO codes) derived from other information in the patent documents (as IPC codes or keywords).

To this end, we use OWL (Smith et al., 2004a) as the ontology language combined with the reasoner Pellet (Sirin et al., 2007), and we make use of the Protégé[14] tool (Anon, 2012b; Noy et al., 2000) for the implementation. This approach has been applied to PatentOntology.

The ICO codes were not implemented in PatentOntology. Therefore, before implementing the concepts, relationships and restrictions, a new class (*ICO*), and a new

14 http://protégé.stanford.edu

property (*hasICOCode*) with a domain (*PatentDocument*) and range (*ICO*), have been created, in accordance with description logics notation (Baader, 2003).

$$ICO \sqsubseteq Classification$$

$$Domain\ (hasICOCode, PatentDocument)$$

$$Range(hasICOCode, ICO)$$

Certain terms for the concepts and relationships in PatentOntology are misleading. For example, there is a property called "*hasIPCClass*" that links the patent with its IPC codes (individuals) and not the IPC Class (classes) as one might think. Therefore, for reasons of clarity, an equivalent object property called "*hasIPCCode*" is introduced.

Next, a hierarchical classification following the Y02 classification of ICO in the EPO databases has been defined, as well as a new class (*Y02_code_pending*) that represents the green patent technology without an official classification from the EPO.

The three scenarios of the method used by the EPO to reclassify new codes have been implemented. The main objective behind each scenario is the summarized as follows:

1. In the first scenario, we found an IPC code that can be directly classified as a green patent technology code. For example, the IPC class *H04N7/26* addresses television systems using bandwidth reduction. We can assume that all patents classified in this field are green patent technology because they attempt to reduce energy use when transmitting a television signal (Lorchat & Noel, 2003; Peng & Lu, 2000).

2. A second scenario is intended to find certain environmentally friendly keywords in the patents that have the same meaning in all of the technological fields and always lead to an environmental patent. For example, the keywords "*energy saving*", "*using less energy*", or "*reduce the energy*", etc., all refer to attempts to use less energy and therefore to be environmentally friendly, regardless of the field of the invention.
3. The third scenario is to have an IPC class that is a priori non-environmentally friendly (such as *H04 "transmission in electric communications"*) but when combined with certain keywords related only to this field (such as "*bit reduction"* or "*less overhead"*) will lead to green patent technology.

The scenarios are described in detail in the next subsections.

5.4.1.1 Reclassification Using IPC Codes

In this first scenario, we will show how to infer a new code based on an existing code. To illustrate this, we will reclassify all of the patents belonging to the IPC class *H04N7/26* "*transmission of television signal using bandwidth reduction*" into the new class *Y02_H04N7_26*, which we have created as a subclass of *Y02_code_pending*. This process includes the following steps:

1. Create the classes: *Y02_H04N7_26* (a subclass of *Y02_code-pending*) and *PatentH04N7_26* (a subclass of *Patent*).
2. Populate the ontology with the individuals with the IPC code *H04N7/26*.
3. Define the equivalent classes for the class *PatentH04N7_26* (see Figure 5.5) with Axioms 5.1 and 5.2. The first axiom will search for the patents with the

IPC code *H04N7_26*. The second axiom will add the value *Y02_H04N7_26* to the patents found.

$$PatentH04N7\_26 \equiv PatentDocument \sqcap \exists hasIPCCode.H04N7\_26$$

(Axiom 5.1)

$$PatentH04N7\_26 \equiv PatentDocument \sqcap \exists hasICOCode.\{Y02\_H04N7\_26Individual\}$$

(Axiom 5.2)

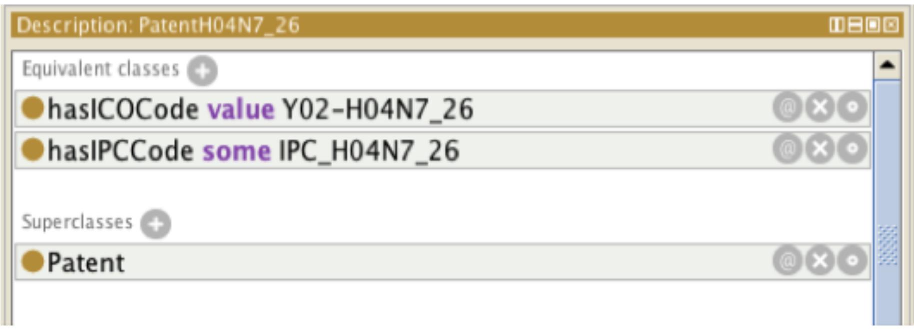

Figure 5.5 Equivalent classes that classified H04N7/26 patents as green patent technology

With these axioms, the reasoner will classify all of the patents with the IPC class *H04N7_26* into the class *PatentH04N7_26*. Next, it will link these patents with the individual *Y02-H04N7_26* of the class *ICO*.

Therefore, when someone searches for *Y02* patents, because the re-classified patents are from one subclass of *Y02*, the re-classified patents will show-up.

For this first scenario, we have inserted individual patents with the *H04N7/26* IPC code, such as *US794843* (see Figure 5.6), with the title "*Method of Implementing Improved Rate Control for a Multimedia Compression and Encoding System*" as well as the IPC codes *H04N7/24, H04N7/26* and *H04N7/30*. The patent addresses a method for rate control and bandwidth reduction. The code *H04N7/26* means '*bandwidth reduction* *of TV signals in communication transmission*', and therefore (as explained in Section 5.4.1), we can classify the patent as environmentally friendly.

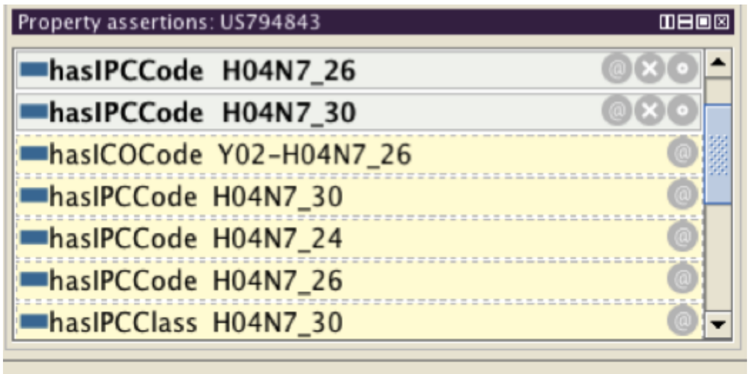

Figure 5.6 Patent US794843 reclassified

The reasoner has inferred the subclass of the class *Y02_code_pending*, *Y02-H04N7_26* for the patent *US794843* (see Figure 5.6). Therefore, the reasoner has reclassified the patents as green patent technology by making use of part of the codes, i.e., using the semantic meaning of the hierarchy found in HCOntology.

Thus, this approach enables the automatic reclassification of patents based on the IPC codes of the patents and can be easily implemented by a IT non-expert with a couple of lines. Furthermore, it is not necessary to change the application on top of the ontology. We have also shown how to create transversal relationships between different concepts.

5.4.1.2 Reclassification Using Keywords

In the second scenario, we will show how to reclassify patents based on transversal relations. In particular, the reasoner will infer a new reclassification of patents based on keywords found in the patent documents. All of the patents (regardless of the technological fields they are classified in) that have keywords related to "*energy-*

saving" (e.g., energy saving, reduce energy, power reduction, etc.) will be classified as green patent technology. This process includes the following steps:

1. Create the classes: *Y02_EnergySaving* (a subclass of *Y02_code-pending*), *Keywords* (and its subclass *EnvironmentalKeywords*) and *PatentEnergySaving* (a subclass of *Patent*).
2. Add the property *hasKeyword*, with range *Patent* and domain *Keyword*.
3. Populate the ontology with the individuals with the keyword "*energy-saving*". These are patents that we found with keywords related to energy saving, and we add to them the property *ObjectHasValue (hasKeyword energy-saving).*
4. Define the equivalent classes for the class *PatentEnergySaving* (see Figure 5.7). The first equivalent class, Axiom 5.3, will search for the patents with the keyword "*energy-saving*". The second equivalent class, Axiom 5.4, will add the value *Y02_EnergySavingIndividual* to the patents found.

$$PatentEnergySaving \equiv PatentDocument \sqcap \exists hasKeyword.\{energy - saving\}$$

(Axiom 5.3)

$$PatentEnergySaving \equiv PatentDocument \sqcap \exists hasICOCode.\{Y02\_EnergySavingIndividual\}$$

(Axiom 5.4)

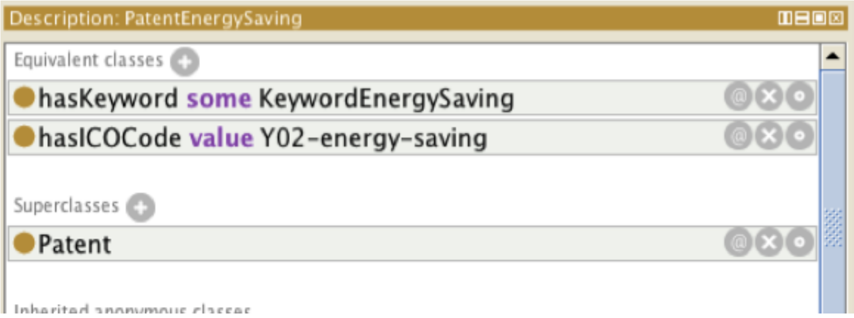

Figure 5.7 Equivalent classes that classify patents related to energy saving as green patent technology

With these axioms, the reasoner will classify all of the patents with the value "*energy-saving*" in the class Keyword in the class *PatentEnergySaving*. Next, it will link all of the patents that are in the class *PatentEnergySaving* with the individual *Y02-energy-saving* of the class *ICO*. Hence, when someone searches for *Y02* patents, because these patents are in a subclass of *Y02*, the re-classified patents will show-up.

For this second scenario, we have inserted individual patents dealing with energy saving, such as *WO2012010194* (see Figure 5.8) with the title "*Energy Saving in a Mobile Communications Network*", that deal with energy savings in a mobile communication network and with the IPC code *H04L5/00* "*affording multiple use of the transmission path in the transmission of digital information"*. In said patent, the IPC code is, a priori, non-environmental in nature, but the patent has in the text selected vocabulary related to *energy-saving*, and therefore we have inserted into the patent the property "*hasKeyword*" with the individual item "*energy-saving*".

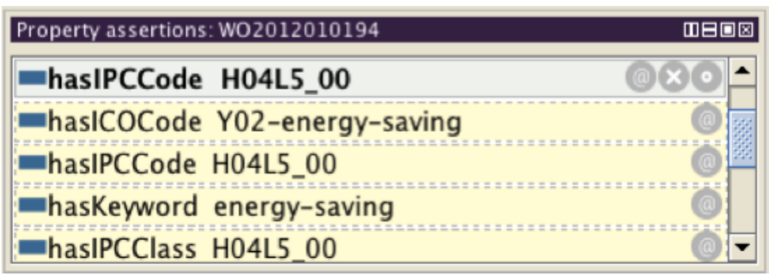

Figure 5.8 Patent WO2012010194 reclassified

The reasoner has inferred the subclass of the class *Y02_code_pending*, *Y02-energy-saving* for the patent *WO2012010194* (see Figure 5.8). Therefore, the reasoner has reclassified patents as green patent technology making use of the transversal relationships between concepts, in this case, the keywords and patent classification.

5.4.1.3 Reclassification Using IPC Codes and Keywords

This scenario will show how to reclassify patents based on two different concepts (in particular, keywords and IPC codes) that will be a combination of the previous scenarios. Specifically, the reasoner will infer a new re-classification of green patent technology based on the IPC code *H04* (and all the subcodes of *H04*) and keywords related to "*bit reduction*" because the energy consumed in the transmission of data in computer networks depends on the number of bits transmitted, among other things. We will link these patents to the *ICO* code *Y02_H04_BitReduction*, created as an individual of the class *Y02_tobeClassified*. In this scenario, the hierarchical classification of the IPC codes aids in reasoning regarding the *Y02* codes. In this case, by only selecting the class *H04*, all of the patents with IPC codes that begin with *H04* will be retrieved.

This process involves the following steps:

1. Create the classes: *Y02_H04BitReduction* (a subclass of *Y02_code-pending*) and *PatentH04BitReduction* (a subclass of *Patent*).
2. Populate the ontology with the individuals with IPC codes beginning with *H04* and keyword '*bit reduction*'. Add to them the property *ObjectHasValue (hasKeyword bit-reduction)*.
3. Define the equivalent classes for the class *PatentH04BitReduction* (see Figure 5.9). The first equivalent class, Axiom 5.5, will search for the patents

with the IPC codes *H04* and keyword "*bit-reduction*". The second equivalent class, Axiom 5.6, will add the value *Y02_ H04BitReduction* to the patents found.

$$PatentY02\_H04\_BitReduction$$
$$\equiv PatentDocument$$
$$\sqcap \exists hasICOCode.\{Y02\_H04BitReductionIndividual\}$$

(Axiom 5.5)

$$PatentY02\_H04\_BitReduction$$
$$\equiv PatentDocument \sqcap \exists hasIPCCode.H04$$
$$\sqcap \exists hasICOCode.\{bit - reduction\}$$

(Axiom 5.6)

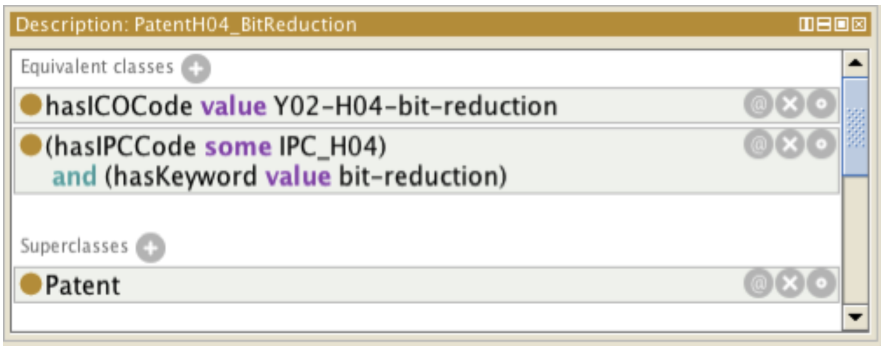

Figure 5.9 Equivalent classes that classify patents H04 related to bit reduction as green patent technology

The reasoner has classified all of the patents with the IPC code *H04* and the value "*bit-reduction*" in the class Keyword into this new class of *PatentH04_BitReduction*. Hence, when someone searches for *Y02* patents, because this individual is in a subclass of *Y02*, the re-classified patents will show-up. Next, all of the patents that are in the class *PatentH04_BitReduction* will be linked with the individual *Y02_H04-bit-reduction* of the class *ICO*.

The ontology has been populated with more patents according to the designed process. A number of these additional patents meet the criteria of this scenario, such as the patent *US20080107132* (see Figure 5.10) with the title "*Method and apparatus for transmitting overhead information*", and classified with the IPC code *H04J3/24*. The patent addresses a method that attempts to reduce the overhead in the transmission. This IPC code, *H04J3/24*, is defined as "*multiplex communications in which the allocation is indicated by an address*", and therefore an a priori non-green patent technology concept is appreciated. However, the patent belongs to the more general code *H04* and includes certain keywords such as "*bit-reduction*". Therefore, the patent meets the criteria shown in Figure 5.9.

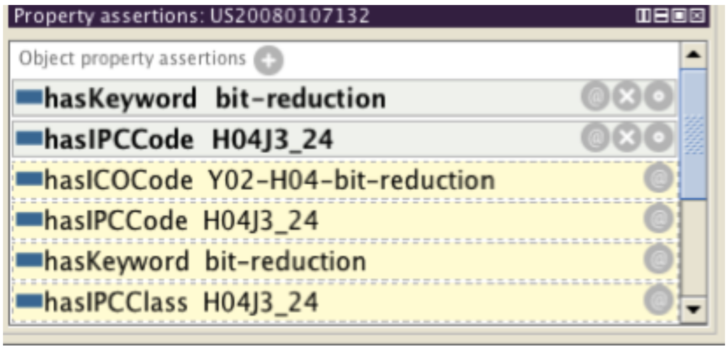

Figure 5.10 Patent US20080107132 reclassified

The reasoner has inferred the subclass of the class *Y02_code_pending*, *Y02-H04-bit-reduction* for the patent *US20080107132* (see Figure 5.10). Therefore, the reasoner has reclassified patents as green patent technology making use of part of the codes, i.e., using the semantic meaning of the hierarchy found in HCOntology combined with transversal relationships such as keywords and patent classification.

5.4.2 Reasoning with HCOntology in Heterogeneous Domains

In this section, we will present a case study that automatically merges two knowledge domains, in particular, the patent domain and the financial domain.

Multiple works have attempted to show the relationship between patent metadata and the financial situation of the firm (Oltra et al., 2010; Vries, 2012). The joint analysis of the patent and financial databases for specific fields of activities will lead to more robust results that avoids the perturbing effect of different contextual factors.

Financial databases include industrial codes of the companies used to classify the activities of the firms. There are several classification schemes of industrial codes (i.e. Standard Industrial Classification –SIC-, International Standard Industrial Classification –ISIS-, Nomenclature statistique des activités économiques dans la Communauté Européenne –NACE-) that allow business analysts to compare firms' economic activities on a statistical basis. However, as we have already mentioned, patent databases are classified by technological fields, and they are different from the industrial codes.

Hence, a link between the industrial codes and technological patent codes opens excellent opportunities for multiple analyses, in particular for the business and economics analyst. This need for matching industrial codes with the technological fields of patents has already been identified by the Organization for Economic Co-operation and Development (OECD) in their study of patent statistics (OECD, 2009).

For this study case, we will focus on the ISIC[15] codes (International Standard Industrial Classification) because they are known worldwide, and they are usually included in the databases regarding firm financial information. However, our findings can be applied to any industrial classification codes.

15 http://unstats.un.org/unsd/cr/registry/regcst.asp?Cl=27

Section 5.5.2.1 will show the internal structure of the ISIC codes, and Section 5.5.2.2 will present a case study of reasoning in heterogeneous domains using technological patent codes and industrial codes.

5.4.2.1 Industrial Codes and Mapping with Technological Fields

To develop the reasoning to automatically reclassify patents into industrial codes, this subsection will introduce the ISIC codes and previous efforts to match the industrial codes with the technological patent codes.

ISIC codes are organized hierarchically, which allows use of the advantages of the hierarchy concepts; however, the label of a specific code cannot be understood automatically by computers to provide semantic meaning to each part of the code.

ISIC codes[16] are divided into 4 parts:

- The first part is the section, which is represented by a letter from *A* to *U*.
- The second part is the division and is represented by 2 digits (from *01* to *99*). The *01* to *03* digits belong to *A*, the *04* to *09* digits belong to *B*, etc.
- The third part is the group, represented by 1 digit.
- The last part is the class, represented also by 1 digit.

For example, the code *2630* belonging to section *C (Manufacturing)*, has the division (*26*), the group (*3*) and the class (*0*). The entire code *2630* means:

26: *manufacture of computer, electronic and optical products*

263: *manufacture of communication equipment*

2630: *manufacture of communication equipment*

2630 is the same as *263* because this group (*263*) has only one class (*2630*).

16 http://unstats.un.org/unsd/cr/registry/regcs.asp?Cl=27&Lg=1&Co=0111

The first attempt made to match the industrial code with IPC codes was performed using the Yale Technology Concordance (YTC) (Kortum & Putnam, 1997). However, the industrial sectors and the IPC codes have evolved since then. More recently, the European Commission has published a report mapping industrial codes with corresponding IPC codes (Schmoch et al., 2003). This report focuses on section *C (manufacturing)* of the industrial codes and found 44 matching technological fields with industrial sectors.

The European commission report contains the most recent mapping; therefore, we will use this mapping for reasoning in the ontology with the IPC classes and to infer the industrial ISIC sectors of the patents.

5.4.2.2 Reasoning with HCOntology Merging Information Domains

In the following subsection, a case study for reasoning using different knowledge domains is described. In particular, different knowledge domains will be linked through various classification schemes, using the mapping between the different hierarchies of classifications. For that purpose, we will focus on the matched sector *35, telecommunication manufacturing,* of the European Commission report. This sector corresponds with code *263* of ISIC rev.4[17], and with 20 IPC codes, among them *G06F, G06G, G06K, H03K* and *H03L*. For clarity reasons, only these 5 IPC codes are shown in this chapter.

The HCOntology, developed in Section 5.4 of this chapter, will allow implementation of this reasoning and demonstrate the advantage of the semantic meaning of parts of the IPC codes.

17 The European commission report matches with the old ISIC rev3.1 code 332, which correspond to the new ISIC v.4 code 263 http://unstats.un.org/unsd/cr/registry/regso.asp?Ci=60

The implementation of this matching in the ontology includes the following steps:

1. Create the classes: *ISIC* (a subclass of *Classification*), a subclass *ISIC263* and *PatentISIC263* (a subclass of *Patent*).
2. Create the hierarchy of classes with the ISIC codes in a similar way as we have done with the IPC codes. The superclass is ISIC. Next, we create the subclasses sections (*ISIC_A* to *ISIC_U)*, then the subclasses (*ISIC_01* to *ISIC_99*) and populate them with individuals (e.g., *263*).
3. Add the property *hasISICCode* with domain *PatentDocument* and Range *ISIC*.
4. Populate the ontology with the individuals with the same 5 IPC codes.
5. Define the equivalent classes for the class *PatentISIC263* (see Figure 5.11). The first equivalent class will search for the patents containing at least one of the corresponding 5 IPC codes (see Axiom 5.7). The second equivalent class will add the value *ISIC_263* to the patents found (see Axiom 5.8).

$$ISIC \sqsubseteq Classification$$

$$Domain\ (hasISICCode, PatentDocument)$$

$$Range\ (hasISICCode, ISIC)$$

$$PatentISIC263 \equiv PatentDocument \sqcap \exists hasIPCCode.(G06G \sqcup G06K \sqcup H03K \sqcup H03L)$$

(Axiom 5.7)

$$PatentISIC263 \equiv PatentDocument \sqcap \exists hasISICCode.\{263\}$$

(Axiom 5.8)

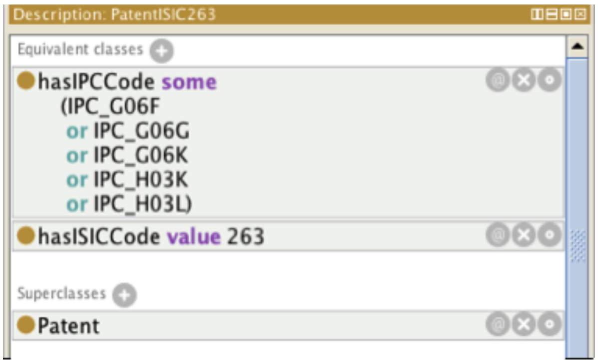

Figure 5.11 Definition of equivalent classes matching IPC codes with the ISIC code

The reasoner, with the first equivalent class, will classify all of the patents with the IPC codes beginning with *G06F* or *G06G* or *G06K* or *H03K* or *H03L* or *H03K* into the class *ISIC263*. Next, with the second equivalent class, all of the patents that are in the class *PatentISIC263* will be linked with individual *263* in the class *ISIC*.

For example, we have populated the ontology with the patent *US2012059512* that has the title "*Apparatus for dispensing solid pharmaceutical articles*" and with the IPC codes *G06F17/00* and *B65D83/00*. The first IPC code meets the criteria because it begins with *G06F*, and the second one does not meet the criteria. As shown in Figure 5.12, the reasoner has introduced the ISIC code *263* into the patent. The equivalent classes (defined above) and HCOntology allow inferences such as the one in this example.

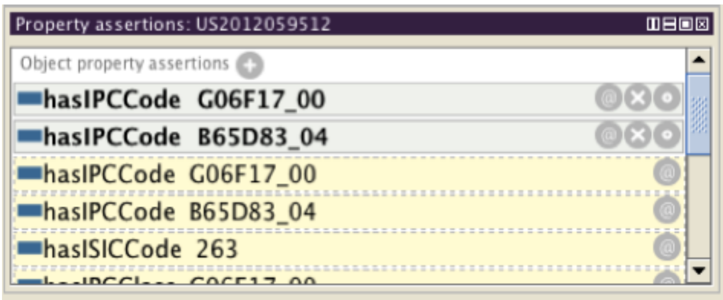

Figure 5.12 Reclassification of a patent with industrial code

Therefore, the semantic meaning of the hierarchy found in the different classification codes can be used to link two different knowledge domains resulting in better patent analyses. In this particular case, the financial studies could combine the information from financial databases with that of the patent databases. Therefore, it is possible to automatically use patents as a measure of innovation and relate them to data from the firms gathered in the financial databases.

5.5 Conclusions

A patent document provides different patent metadata about the innovation it protects such as its location, date or technological field. The technological field is one of the most used patent metadata item for innovation research. Information about patents is stored in large patent document databases without semantic annotations that allow further relationships and links between patents and other knowledge domains.

Ontology languages such as OWL have proven to be useful in the representation of knowledge; it also has also begun to be used to represent patent knowledge. However, patent ontologies based on this language are scarce (Taduri et al., 2011; Wanner et al., 2009), and in most cases, they just mirror the technological patent codes, without further leveraging reasoning capabilities. Therefore, the analysis of the transformation of these hierarchical codes into hierarchical concepts with semantic meaning remained to be addressed.

This work introduces a method to transform the representation of the hierarchical technological patent codes into a hierarchy of concepts, HCOntology, and discusses the importance and potential of this transformation. For that aim, this chapter has first identified the problems and potential solutions in the translation of technological patent codes into a hierarchy of concepts. A proposal has been devised to address these issues by defining a patent ontology for the analysis of the hierarchical technological patent code. And finally, a process has been designed for retrieving information from different patent repositories and indexing hierarchical codes in ontological categories. More specifically, the retrieval process includes automatically translate XML instances from different data repositories into OWL classes and instances. This proposal makes use of XSLT, and the stylesheets are built with the help of XPATH that splits the code into its structural parts. The stylesheets then rebuild the OWL classes from these parts of the code.

The proposed method enables the current and future emerging codes to be translated automatically from the patent databases to our ontology, without the need for any further implementation. The resulting hierarchy of concepts in OWL (which we called HCOntology) allows the exploitation of the information gathered in each part of the hierarchy, enriching the information retrieval process with new relationships, properties and inferred information. This exploitation requires the creation of transversal relationships between concepts that reclassify existing and future patent documents. The proposal provides benefits from two different perspectives: enriched knowledge can be obtained in an more automatically way for a better analysis of the patent information; new classifications (or changes in exiting ones) can be developed without the need of changes in the rest of applications that make use of the ontology.

Particularly, a patent ontology has been defined, and several techniques have been applied to exploit it especially taking into consideration the information represented in the hierarchical classification proposed (HCOntology). To this end, transversal relationships between concepts can be included and automatically processed by computers. The ontology language OWL has been used to implement and validate the proposal through two case studies.

In the first case study, we have simulated the introduction of a new code into a technological patent classification. In particular, the use of existing technological patent codes and/or specific keywords to generate new green patent technology codes has been described. This opportunity is especially appealing in the context of steadily increasing technological changes, in which multiple technologies (e.g., green patent technology) are developing quickly. Our analysis provides opportunities to automatically update the previous classification of the technological fields in an efficient way.

In the second case study, we have connected patents to other sources of knowledge for more enriched knowledge. In particular, the analysis has been carried out in terms of linking the firm’s financial knowledge with patent knowledge.

The approach has been applied to PatentOntology, but it could be used in any other patent ontology. The proposed method has been applied to the domain of patents, but it could be also applied to any domain with hierarchical codes. The hierarchical codes would therefore be enriched with semantics, enabling the definition of new relationships, properties and inferred information.

Future studies might explore additional opportunities for the OWL language to generate and exploit hierarchies of concepts to analyze data in different contexts. Extensions of the proposed patent ontology to analyze additional information (not provided in the patent database) may also be an attractive direction for future work.

References

Anand, N., Yang, M., Van Duin, J. H. R., & Tavasszy, L. (2012). GenCLOn: An Ontology for City Logistics. *Expert Systems with Applications,* 39(15), 11944-11960.

Baader, F. (2003). *The Description Logic Handbook: Theory, Implementation and Applications*. Cambridge University Press.

Bedini, I., Gardarin, G., & Nguyen, B. (2008). Deriving Ontologies from XML Schema. *Proceedings EDA 2008*.

Berners-Lee, B. T., Hendler, J., & Lassila, O. (2001). The Semantic Web. *Scientific American*, *284*(5), 34–43.

Bohring, H., & Auer, S. (2005). Mapping XML to OWL Ontologies. *Leipziger InformatikTage*, *72*, 147–156.

Bosch, T., & Mathiak, B. (2011). XSLT Transformation Generating OWL Ontologies Automatically Based on XML Schemas. *International Conference for Internet Technology and Secured Transactions* 660–667.

Buzzacchi, L., & Scellato, G. (2008). Patent Litigation Insurance and R&D Incentives. *International Review of Law and Economics*, *28*(4), 272–286.

Cardoso, J., & Bussler, C. (2011). Mapping between Heterogeneous XML and OWL Transaction Representations in B2B Integration. *Data & Knowledge Engineering*, *70*(12), 1046–1069.

Chen, T.-Y. (2011). Value Ontology-based Multi-aspect Intellectual Asset Valuation Method for Decision-making Support in K-commerce. *Expert Systems with Applications*, *38*(5), 5471–5485.

Choi, S., Kang, D., Lim, J., & Kim, K. (2012). A Fact-oriented Ontological Approach to SAO-based Function Modeling of Patents for Implementing Function-based Technology Database. *Expert Systems with Applications*, 39(10), 9129–9140.

Fallside, D. C., & Walmsley, P. (2004). XML Schema Part 0: Primer Second Edition. Available at http://www.w3.org/TR/xmlschema-0.

Ferdinand, M., Zirpins, C., & Trastour, D. (2004). Lifting XML Schema to OWL. *Lecture Notes in Computer Science*, 3140, 354–358.

Foglia, P. (2007). Patentability Search Strategies and the Reformed IPC: A Patent Office Perspective. *World Patent Information*, *29*(1), 33–53.

Giereth, M., St, A., Rotard, M., & Ertl, T. (2006). Application of Semantic Technologies for Representing Patent Metadata. *1st International Workshop on Applications of Semantic Technologies AST06*, *94*, 297–305.

Goebel, M. W. E., & Hintermaier, F. S. (2011). Searcher's Little Helper – ICO Index Terms. *World Patent Information*, *33*(3), 260–268.

Gruber, T. R. (1993). Toward Principles for the Design of Ontologies Used for Knowledge Sharing. *International Journal HumanComputer Studies*, *43*(5/6), 907–928.

Guarino, N., & Giaretta, P. (1995). Ontologies and Knowledge Bases: Towards a Terminological Clarification. *Towards Very Large Knowledge Bases Knowledge Building and Knowledge Sharing,* (1), 25–32.

Gurry, F., Fink, C., Khan, M., Bergquist, K., Lamb, R., Feuvre, B. Le, & Hao Zhou. (2012). *World Intellectual Property Indicators 2011. WIPO Economics & Statistics Series*. Available at http://www.wipo.int/freepublications/en/intproperty/941/wipo_pub_941_2011.pdf.

Kortum, S., & Putnam, J. (1997). Assigning Patents to Industries: Tests of the Yale Technology Concordance. *Economic Systems Research*, *9*(2), 161–176.

Lacoste, D., Sawant, K. P., & Roy, S. (2011). An efficient XML to OWL Converter. *Proceedings of the 4th India Software Engineering Conference 2011 ISEC11*, 145–154.

LexisNexis Website. (2012). Available at *www.lexisnexis.com/*.

Lorchat, J., & Noel, T. (2003). Energy Saving in IEEE 802.11 Communications using Frame Aggregation. *GLOBECOM '03. IEEE Global Telecommunications Conference,* 1296–1300.

Malhotra, A., Melton, J., & Walsh, N. (2007). XQuery 1.0 and XPath 2.0 Functions and Operators. *Working draft W3C*. Available at http://www.w3.org/TR/xpath-functions.

Noguera, M., Hurtado, M. V., Rodríguez, M. L., Chung, L., & Garrido, J. L. (2010). Ontology-driven Analysis of UML-based Collaborative Processes using OWL-DL and CPN. *Science of Computer Programming*, *75*(8), 726–760.

Noy, Natalya F, & McGuinness, D. L. (2001). Ontology Development 101: A Guide to Creating Your First Ontology. Available at http://www.ksl.stanford.edu/people/dlm/papers/ontology-tutorial-noy-mcguinness.pdf.

Noy, Natalya Fridman, Fergerson, R. W., & Musen, M. A. (2000). The knowledge model of Protégé-2000 : combining interoperability and flexibility. *Knowledge Engineering and Knowledge Management Methods Models and Tools*, *1937*(1), 1–20.

OECD. (2009). OECD Patent Statistics Manual. *Statistics*. Available at http://www.oecd.org/science/innovationinsciencetechnologyandindustry/oecdpatentstatisticsmanual.htm.

Oltra, V., Kemp, R., & Vries, F. de. (2010). Patents as a Measure of Eco-innovation. *International Journal of Environmental Technology and Management*, *13*(2), 130–148.

Pelegrina, A. B., Benghazi, K., Hurtado, M. V., Noguera, M., Barjis, J., Aalst, W., Mylopoulos, J., et al. (2010). A Framework for the Semantic Representation of Business Processes within Business Organizational Models. *Enterprise and Organizational Modeling and Simulation,* (63), 79–94.

Peng, W., & Lu, X.-C. (2000). On the Reduction of Broadcast Redundancy in Mobile ad hoc Networks. *Proceedings of the 1st ACM international symposium on Mobile ad hoc networking & computing*, 129–130.

Porter, M. E. (1990). *Competitive Advantage of Nations*. Free Press.

R. Shearer, B. Motik, and I. H. (2008). HermiT: A Highly-efficient OWL Reasoner. *Proceedings of the 5th International Workshop on OWL: Experiences and Directions (OWLED)*, 26-27.

Rodrigues, T., Rosa, P., & Cardoso, J. (2008). Moving from Syntactic to Semantic Organizations using JXML2OWL. *Computers in Industry*, *59*(8), 808–819.

Schmoch, U., Laville, F., Patel, P., & Frietsch, R. (2003). Linking Technology Areas to Industrial Sectors *Final Report to the European Commission*. Available at ftp://ftp.cordis.europa.eu/pub/indicators/docs/ind_report_isi_ost_spru.pdf.

Seol, H., Lee, S., & Kim, C. (2011). Identifying new Business Areas using Patent Information: A DEA and Text Aining approach. *Expert Systems with Applications*, *38*(4), 2933–2941.

Sirin, E., Parsia, B., Grau, B., Kalyanpur, A., & Katz, Y. (2007). Pellet: A Practical OWL-DL Reasoner. *Web Semantics Science Services and Agents on the World Wide Web*, *5*(2), 51–53.

Smith, M. K., Welty, C., & McGuinness, D. L. (2004). OWL Web Ontology Language Guide. *W3C Recommendation*. The World Wide Web Consortium (W3C). Available at http://www.w3.org/TR/owl-guide.

Soo, V.-W., Lin, S.-Y., Yang, S.-Y., Lin, S.-N., & Cheng, S.-L. (2006). A Coperative Multi-agent Platform for Invention Based on Patent Document Analysis and Ontology. *Expert Systems with Applications*, *31*(4), 766–775.

Soylu, A., De Causmaecker, P., Preuveneers, D.B., & Desmet, P. (2011). Formal Modelling, Knowledge Representation and Reasoning for Design and Development of User-centric Pervasive Software: A Meta-review. *International Journal of Metadata, Semantics and Ontologies*, *6*(2), 96–125.

Taduri, S., Lau, G. T., Law, K. H., Yu, H., & Kesan, J. P. (2011a). An Ontology to Integrate Multiple Information Domains in the Patent System. *IEEE International Symposium on Technology and Society ISTAS*, 23–25.

Taduri, S., Lau, G. T., Law, K. H., Yu, H., & Kesan, J. P. (2011b). Developing an Ontology for the U . S . Patent System, *Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times*, 157-166.

The Protégé Ontology Editor and Knowledge Acquisition System. (2012). Available at http://protege.stanford.edu/

Vries, D. (2012). *Leveraging Patents Financially. A Company Prespective*. Springer.

W3C. (1999). XSL Transformations (XSLT). Available at http://www.w3.org/TR/xslt

Wanner, L., Brügmann, S., Diallo, B., Giereth, M., Kompatsiaris, Y., Pianta, E., Rao, G., et al. (2009). PATExpert : Semantic Processing of Patent Documentation. Available at *http://recerca.upf.edu/patexpert/*.

Wi, H., Oh, S., & Jung, M. (2011). Virtual Organization for Open Innovation: Semantic Web based inter-organizational Team Formation. *Expert Systems with Applications*, *38*(7), 8466–8476.

CAPÍTULO 6

ANALYZING A FIRM'S INTERNATIONAL PORTFOLIO OF TECHNOLOGICAL KNOWLEDGE: A DECLARATIVE ONTOLOGY-BASED OWL APPROACH ON PATENT DOCUMENTS

ANALYZING A FIRM'S INTERNATIONAL PORTFOLIO OF TECHNOLOGICAL KNOWLEDGE: A DECLARATIVE ONTOLOGY-BASED OWL APPROACH ON PATENT DOCUMENTS

6.1 Introduction

In the current context of massive technological change and the important phenomenon of globalization, innovation is considered strategically important for the ability of firms to build and sustain competitive advantage thus creating value (Ciabuschi et al., 2011). A firm's patents are resources that may create a barrier to protect the income generated from innovation efforts or otherwise prevent rivals from competing in that innovative space (Grant, 1991). In this context, the number of patents filed worldwide grew by 7.2% in 2010 compared to a 5.2% increase in global gross domestic product (GDP) (Gurry et al., 2012).

Patent data are known to contain detailed information such as the applicant's name (or patenting firm), technological type, inventor, geographic location, which is particularly useful for economic development and business research. However, much of the useful information and knowledge that patents can provide requires sophisticated analysis through the connection of multiple data dimensions which are often not explicitly represented in patent databases. This situation generates serious difficulties for the automatic processing of information by computers and support for traditional patent analysis is therefore expensive and costly in terms of time and manpower (Trappey et al., 2012). In order to fully exploit patent metadata information and preprocess it for business analysis, this chapter proposes a method for exploiting the potential of identifying relevant relationships in OWL in order to broaden the information gathered in patent metadata. This chapter aims to formalize the reasoning shown in the previous chapter. The contribution of this chapter is to extend previous work on patent ontologies with a declarative method by means of the formal definition of key patent analysis indicators (henceforth referred to as KPAIs) in OWL.

In recent years, ontologies have become more and more popular in information system engineering (Noguera et al., 2010b; Panetto et al., 2012; Pelegrina et al., 2010; Rocca, 2012; Weissman et al., 2011) for describing information as they provide formal, uniform and shareable representations about a domain, and various patent ontologies have been proposed. Patent ontologies classify concepts (such as metadata found in most patent databases or keywords found in patent documents) hierarchically so that patent classifications may be developed and complex relationships between concepts may be examined.

One of the most used ontology languages is OWL (W3C, 2009). Much of the power of the OWL ontology, whereby model-theoretic semantics are based on description logics, relies on its reasoning capabilities (Smith et al., 2004). Reasoning in OWL allows certain verification procedures to be executed such as consistency checks, concept satisfaction and classification. More specifically, when relevant relationships are identified and represented between concepts in OWL, reasoners could classify concepts based on these relationships.

Our proposed method identifies and represents new relationships between concepts that allow KPAIs which are not explicitly specified in patent metadata to be inferred. The declarative nature of the proposed method involves first declaring the concepts and relationships which are then used by the system to automatically reclassify all present and future data.

By using a sample implementation, this research shows how OWL ontologies can represent new relationships between concepts gathered in patent documents. In particular, this research analyses information about international patent flows by implementing a classification which depends on patent internationalization. Such developments also illustrate the potential of ontologies for managing and researching the field of business.

Certain ontologies in fields other than patents have identified relevant relationships between concepts to infer new knowledge such as in the field of biological science (Wolstencroft et al., 2007) or in the field of context modeling (Bettini et al., 2010). However, none of these ontologies has proposed a method to identify new relationships and implement them in OWL to infer new knowledge, and no previous publication has implemented relationships specifically on the patent data domain. Although previous literature has already used artificial intelligence algorithms in the context of ontologies to analyze patents for detecting emerging information from specific keywords (Shih & Liu, 2010; Trappey et al., 2006), these ontologies are not declarative. Declarative methods complements artificial intelligence algorithms by a priori delimitating concepts in the ontology.

The information inferred from the proposed method could be used by academics and practitioners alike to improve their studies based on patents. This topic is especially relevant in the current context of competition in which firms are increasing their innovative activities organized at international levels (OECD, 2009). This study of the strategies based on the internationalization of innovative activities will allow firms to create a global competitive advantage (Almeida & Phene, 2004).

The remainder of the chapter is organized as follows: Section 6.2 provides a background study of patents, ontologies and patent ontologies; Section 6.3 describes the method of the proposal; Section 6.4 presents a case study; Section 6.5 complements our method with the description of its potential to modificate the declarative approach; Section 6.6 shows the performance of our proposal by running two reasoners in the context of our case study; and finally, Section 6.7 concludes the chapter by discussing the contributions of the research.

6.2 Background and Literature Survey

6.2.1 Patents

Patents are legal documents that describe technological inventions and protect the rights of the inventor of an industrial property to exploit the invention. Patent documents are organized into metadata fields which contain information such as the inventor, applicant, international technological classification, title, description and claims.

Patents, as we described in Chapter 1, are frequently used as indicators of technology and predictors of economic performance (Vries, 2012). They are also used to measure the inventiveness of countries, regions, firms or inventors (Almeida & Phene, 2004). One of the many uses of patent statistics is to measure the dynamics of the innovation process, (e.g. co-operation between industries or countries). These enable the patterns and intensity of international co-invention and foreign ownership of domestic inventions (and vice versa) to be tracked. This is particularly useful nowadays in the global market where it is common to use alliances between different geographical locations to obtain research synergies and to acquire new technological competences (OECD, 2009).

In conclusion, as stated in the OECD's patent statistics manual: "*patent data provide unique insights into the processes and outcomes of inventive activities (e.g. the location of inventive activities, inventive networks, emerging technologies, etc.). Used with other data, they support the analysis of other dimensions of innovation that are of policy interest, such as the role of intellectual property in economic performance, entrepreneurship, and the tracking of linkages in the science and technology system*" (OECD, 2009).

Traditional patent analysis, however, requires significant cost, time and manpower (Trappey et al., 2012), because they need to download the target patent documents into a customized database, and manually query the database in order to reclassify patent documents based on a new classification. Each time that new patent documents need to be added to the analysis, the customized database needs to be updated and a new query has to be designed and performed.

Algorithmic approaches are useful to analyze the big amount of information in the patent database. Multivariable analyses (e.g. factorial analysis) are usually used to extract key information without a priory fixed model, and a model is trained with the extracted information to predict consequences (Trappey et al., 2012). In declarative approaches, the model is a priory proposed, and any subsequent change needs only the change in the ontology while software tools may remain stable. In any case, the information and data that patents can provide require sophisticated analysis to connect multiple data dimensions and these are not explicitly represented in any patent database. As a result, it is extremely difficult for computers and statistical software applications to automatically process the information. This chapter proposes the development of a patent ontology to declare concepts and relationships and automatically infer information that is not explicitly written in the patent metadata.

5.2.2 Ontologies

Various definitions of ontology have been proposed in the literature (Guarino & Giaretta, 1995; Noy & McGuinness, 2001) and one of the most widely used is the "*specification of a conceptualization*" (Gruber, 1993a). As we introduced in Chapter 2, ontologies consist of a formal description of concepts (classes) and their properties, relationships, constraints and behavior (Gruber, 1993b; Noy & McGuinness, 2001). Ontologies describe a common vocabulary for each particular domain, enabling for example the automatic search for concepts, regardless of the terms used to represent the same concepts.

Ontologies are useful when dealing with large collections of information found in data repositories since they are able to maintain the information in a hierarchical organization of concepts and their subsumptions, like a taxonomic hierarchy (Beydouna et al., 2011). In terms of ontology languages, OWL (W3C, 2009) has become the de facto standard in the system and software engineering community.

OWL has a formal semantics based on description logics (Baader, 2003). There are several OWL-based reasoners such as Pellet (Sirin et al., 2007) or Hermit (Shearer et al., 2008) that perform verification procedures such as consistency checks, concept satisfaction and classification.

There are several proposals in the literature to create ontologies for any purpose, through the declaration of concepts and properties [see (Dahlem & Hahn, 2009) for a review], being the most cited the "*Ontology Development 101*" (Noy & McGuinness, 2001). However none of the previous literature deals with the method to create axioms and exploit the ontology with the power of reasoning, in order to enrich the way firms could analyze their innovation.

This chapter will use OWL-based reasoners and the most widely used visual tool to create ontologies, Protégé (Knublauch et al., 2004) to discuss and illustrate how ontologies may be efficient systems for representing patent information and reasoning, by identifying new contexts and dependency information not shown in the patent databases.

5.2.3 Patent Ontologies

As we described in Chapter 2, different ontologies about patents have been implemented in recent years. These ontologies allow different patent databases to be integrated with a homogeneous representation of semantics. Two main different groups of patent ontologies can be identified: patent ontologies that make use of artificial intelligence algorithms for data-mining and patent ontologies that make use of techniques that are based on the declaration of concepts.

In the first group are the ontologies that use the frequency of certain words found in patent documents to analyze technological trends or other useful information (Kim et al., 2008; Shih et al., 2010; Trappey et al., 2009). This type of work uses text data mining and intelligent algorithms to create patent ontologies. In the ontologies using artificial intelligence, patent data is analyzed and a model is derived from the data.

Other patent ontologies, meanwhile, have been developed with OWL which are based on keywords found in patents but not on the metadata (Choi et al., 2012; Seol et al., 2011; Soo et al., 2006). These ontologies mainly search new technological fields but do not contemplate patent metadata such as the applicant, inventor, etc. which is useful for business analysis. Moreover, these ontologies reclassify patent documents by querying the ontology. However, they do not provide a method to implement complex analysis, where auxiliary classes and properties are needed. These complex analyses are difficult to implement with queries, and the auxiliary classes are also difficult to be reused in other future complex analysis. In any case, the ontologies based on keywords and other patent information could be combined with patent metadata ontologies to increase the knowledge base, and enabling the information required for business analysis to be processed automatically.

In the second group, ontologies are structured models of concepts to represent the patent metadata (patent ontologies using concept declaration). Although these metadata-based patent ontologies represent the concepts of patent metadata, they do not identify relevant or complex relationships between these concepts in the ontology. The most relevant declarative patent ontologies based on patent metadata and represented in OWL are Patexpert, which is an ontology created within the European Patexpert project (Giereth et al., 2006; Wanner et al., 2009), and the PatentOntology, which was created at Stanford University (Taduri et al., 2011a).

Patexpert is a huge patent ontology developed over the core Suggested Upper Merge Ontology (SUMO). The ontology combines several smaller ontologies, such as PMO (Patent Metadata Ontology) that represents all the metadata used in most patent databases. Within the project, several applications using the ontology were developed.

The public ontology is not populated and mostly focuses on the delimitation of a global structure to include patent metadata information.

PatentOntology is a newer ontology created explicitly to integrate heterogeneous domains (Taduri et al., 2011a) and only for patents recovered from the USPTO (United States Patent and Trademark Office) database. PatentOntology simultaneously represents patent metadata and information from patent court cases. This combination of information helps companies to find not only the patent claims in a technological area but also to discover whether the claims are still valid or have been invalidated by a court. This ontology reduces the time required by companies to search for this information since all the information is in one knowledge base.

The method that we present in this chapter is related to the complex exploitation of the information in patent ontologies with the power of the reasoners, that allow inferring new knowledge, based on relationships between concepts of the ontology and not to the creation of an ontology, like previous works. Therefore our proposal could be applied to an ontology created by a declarative approach or over an ontology created through data mining analysis.

6.3 Proposal Enabling Explicit Relationships to be Made Between Concepts in Patent Ontologies

This methodological proposal is based on the declaration of concepts, properties and relationships between concepts, extending the creation of ontologies proposed by Noy and McGuinness (2001) and further developed by Darlington and Culley (2008) with further steps to declare classes and relationships for inferring new further knowledge with existing reasoning tools. In order to apply this method for analyzing relationships in ontologies, it is desirable to have an already implemented ontology, which we call the "*host ontology*". In this case, the proposal has been implemented over PatentOntology, although it could be implemented in any existing or new patent metadata ontology. The proposed steps of the declarative method are shown in Figure 6.1.

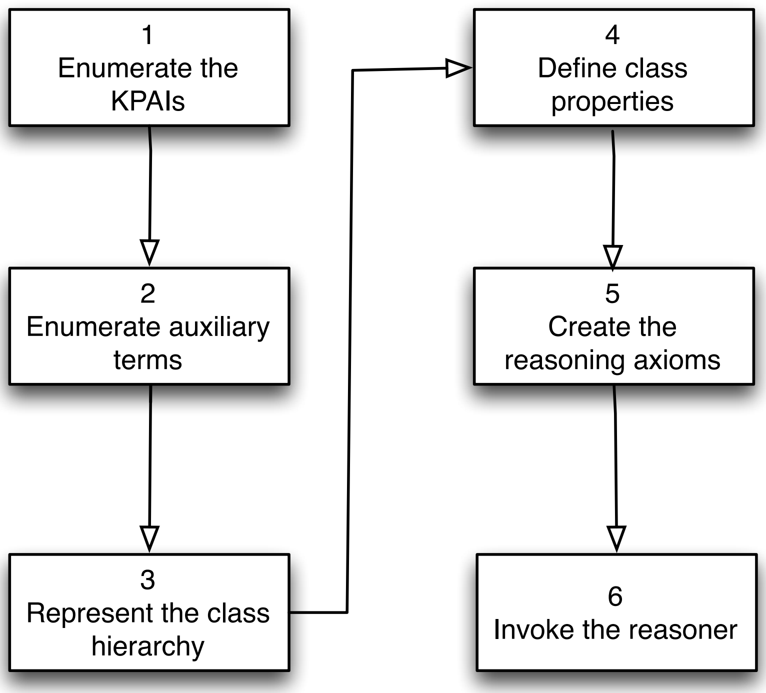

Figure 6.1 Method for analyzing patent ontology relationships

This proposal highlights the usefulness of representing relevant relationships in OWL in order to increase the knowledge gathered in ontologies or databases. Despite the recognized importance of identifying relevant relationships in the context of patent ontologies, previous publications have used the relationships between concepts in a limited way. The proposal also focuses on ontology specification processes to represent complex relationships in patent metadata ontologies in order to specify certain information and allow the inferred information to be automatically classified and processed by computers. These relationships will provide insights and outcomes of inventive activities (OECD, 2009) and together with other data could show the importance of industrial property strategies in economic performance (Somaya, 2012).

6.3.1 Step1. Enumerate the Key Patent Analysis Indicators (KPAIs)

When developing an ontology, the first step is to enumerate the important concepts in the ontology domain. This involves a conceptual mapping where experts in the target ontology domain identify the domain's important concepts through brainstorming or discussion and also involves definition of the concepts and identification of synonyms or terms defining the same concept. In a similar way, the proposed method for specifying relationships in the ontology firstly requires the full delimitation of classes delimitated by different concepts of the ontology: these classes are the key patent analysis indicators (KPAIs) for our analysis.

When the concepts in a patent ontology include information such as applicant, inventor, applicant country, the KPAIs will be complex concepts defined by specific characteristics of other concepts in the ontology. For example, KPAIs could be the different patterns that classify patents into an international research flow classification depending on the specific values of the applicants' and inventors' countries.

6.3.2 Step 2. Enumerate Auxiliary Terms

This step involves defining and declaring the auxiliary concepts (or classes in OWL) which are required to represent the complex relationships (or axioms) that will complete the KPAI definition. These auxiliary terms depend heavily on the analyst. The use of one set of classes or another depends on how the analyst sees the structure of the complex relationships. As Noy and McGuinness (2001) recognize in their general method for representing ontologies: "*there is no one correct way to model a domain –*

there are always viable alternatives". This assertion could also be applied to the method proposed to model new relationships between concepts.

6.3.3 Step 3. Represent the KPAIs and Auxiliary Terms in OWL Classes and Class Hierarchies

Once the KPAIs have been identified, this step consists in ordering the KPAIs hierarchically and representing them as ontology classes. This step matches the corresponding step in the development of an ontology. There are several ways of ordering the hierarchy: bottom-up, top-down, or a combination of the two. Although all of these can lead to a different hierarchy of classes, there is no clear better way to represent an ontology and it really depends on author preference. Nevertheless, for the representation of the KPAIs there is normally a reduced number of classes to organize and there is not much space for variance.

6.3.4 Step 4. Define Class Properties

Class properties represent the internal structure of concepts. This step declares the relationships or properties that the KPAIs (or classes representing each KPAI) need. These properties link one class to another, or one class with a type of *data: integer, dateTime, Boolean*, etc. In this step, property restrictions are also described, e.g. cardinality (how many values a property can have), range (permitted classes that a property links with) and domain (classes that a property describes).

6.3.5 Step 5. Create the Reasoning Axioms

In this main step of our proposed method, all the classes and properties represented in the previous steps and the classes and properties existing in the host ontology are combined so that complex relationships may be made between them. These relationships will allow patents to be classified according to a certain KPAI.

Unlike traditional databases, OWL follows the open world assumption (OWA), which implies that if something is not explicitly stated then we cannot assume that it is true or false, only that it is unknown. In order to infer new knowledge and reason about certain relationships, it is sometimes necessary to close this OWA (Grimm & Boris, 2005) by using closures axioms. When these axioms are applied over the properties, they limit the extension of the properties of a class.

This step allows the definition of the new classes created in Step 1 (KPAIs) based on existing or auxiliary classes or properties. This step also creates the auxiliary closure axioms needed to implement the reasoning.

6.3.6 Step 6. Invoke the Reasoner

This is the final step in which new information will be automatically inferred from the relationships (or axioms) identified above and is performed by OWL language reasoners. The reasoners will perform verification procedures such as KPAI consistency checks, concept satisfaction and more specifically classification, which automatically classifies instances into the newly created classes in Step 1 or KPAIs.

6.4 Case Study: Use of Reasoning to Analyze a Firm's International Patent Portfolio

In order to illustrate the applicability of the method, this chapter implements the proposal by classifying corporate patents into the five patterns of internationalization identified by the OECD (OECD, 2009). This analysis of the geographical location of innovative activities is important because of their implication in the technological capacity of countries, and in order to quantify the intensity and geographical patterns of these activities (Lahiri, 2010; Quintas et al., 2008).

When the number of international patents filed under the Patent Cooperation Treaty (PCT) grew by 5.7% in 2010 (Gurry et al., 2012), managers had to consider the specific geographical location of their firms and to decide on the specific location of the invention process. It is therefore important to analyze the connections between patenting processes and international knowledge flows of research and development. In this context, the importance of developing specific relationships between concepts to fully understand the situation of this internationalization is particularly important.

The OECD classification aims to study international flows of knowledge based on patent owner country and inventor country, thereby sorting patents according to whether the research has been conducted by subsidiaries of a multinational corporation, by joint ventures or by co-operation between companies, etc. The OECD classification therefore divides patents into five groups:

- Purely domestic ownership of foreign inventions (country A owner and country B inventor)
- Domestic ownership implying co-ownership with a single inventor (countries A and B owners but only country B inventor)
- Domestic ownership with co-invention (countries A and B inventors but only country A owner)
- Co-ownership jointly with co-invention (countries A and B inventors and owners)

- Cross-border ownership or inventorship with separate inventor and owner countries (A and B owners and C inventor)

Following the steps proposed in Section 6.3, the following subsections will show how to implement the OECD classification in an OWL patent ontology, and how the reasoners will reclassify current and future patent documents into the correspondent groups of the OECD classification.

6.4.1 Step1. Enumerate the KPAIs

This step enumerates the KPAIs identified by the OECD classification. In this case, there are five KPAIs, each corresponding to each of the five patterns of internationalization identified by the OECD.

- *PurelyDomesticForeignInventors*: representing purely domestic ownership of foreign inventions

- *CoDomesticSingleInventor*: representing domestic ownership implying co-ownership with a single inventor

- *DomesticWithCoinventors*: representing domestic ownership with co-invention

- *Jointly*: representing co-ownership jointly with co-invention

- *CrossBorder*: representing cross-border ownership or inventorship with distinct inventor and owner countries

6.4.2 Step 2. Enumerate Auxiliary Terms

This step describes various auxiliary terms that will help create the axioms in step 5. In this example in particular, one of the auxiliary terms is the number of countries where the different applicants reside. The applicants represent patent owners because patent ownership first belongs to the patent applicant. Another auxiliary term necessary for Step 5 is the number of different countries where the inventors reside.

6.4.3 Step 3. Represent the KPAIs and Auxiliary Terms in OWL Classes and Class Hierarchies

The parent class *InternationalOECD*, representing the OECD classification, is a subclass of the class Patent which belongs to the host ontology PatentOntology. The five KPAIs identified by the OECD are subclasses of *InternationalOECD*. Similarly, the auxiliary classes (*NumberApplicantCountries* and *NumberInventorCountries*) are also subclasses of the class Patent, which belongs to the PatentOntology. Both auxiliary classes have their subclasses (see Figure 6.2): one subclass involving only one country (*ApplicantCountry1* and *InventorCountry1*), another subclass involving exactly two countries (*ApplicantCountry2* and *InventorCountry2*), and the final subclass involving three or more countries (*ApplicantCountry3more* and *InventorCountry3more*). The class hierarchy followed in this example and implemented with the tool protégé is shown in Figure 6.2.

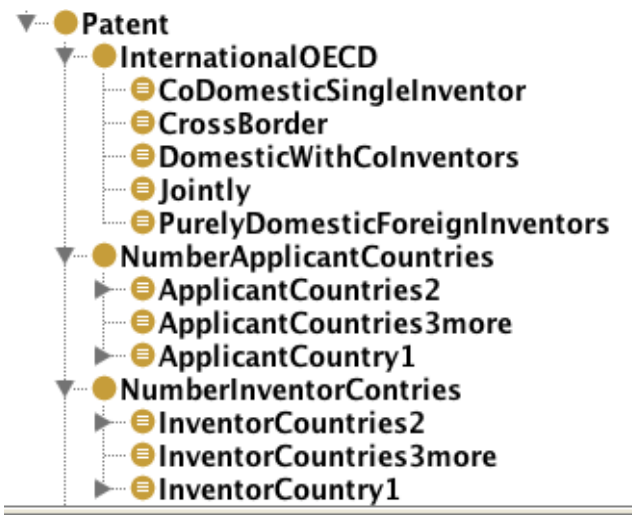

Figure 6.2 Representation of KPAI, and auxiliary terms as OWL classes in Protégé

6.4.4 Step 4. Define Class Properties

In order to define the relationships that will allow the KPAIs in this example to be inferred, certain class properties between patent and applicant countries and inventor countries must be defined. These class properties are:

- *hasApplicantCountry*. This property links patents with applicant countries. In accordance with description logic notation (Baader, 2003)

$$Domain\ (hasApplicantCountry, PatentDocument)$$

$$Range(hasApplicantCountry, State)$$

- *hasInventorCountry*. This property links patents with the inventor countries.

$$Domain\ (hasInventorCountry, PatentDocument)$$

$$Range(hasInventorCountry, State)$$

6.4.5 Step 5. Create the Reasoning Axioms

In order to represent the classes of the OECD classification that involve inventors residing in a single country, two different countries or three or more countries (see Figure 6.3), it is advisable to close the OWA so as to enable reasoning in some of the classes identified in Step 1. This step applies closure axioms over the classes *NumberOfApplicantCountries* and *NumberInventorsCountries* for the six different subclasses (*ApplicantCountry1, InventorCountry1, ApplicantCountry2, InventorCountry2, ApplicantCountry3more* and *InventorCountry3more*). For example, the class *ApplicantCountries2* needs a closure axiom so that the cardinality of the property "*hasApplicantCountry*" is exactly two countries. According to description logic notation:

$$InventorCountries2 \equiv \leq 2\ hasInventorCountry \sqcap \geq 2\ hasInventorCountry$$

(Axiom 6.1)

Figure 6.3 Closure axiom for the class InventorCountries2

Consequently, all the patents declared as individuals of this class have exactly two applicant countries. These closure axioms will enable us to determine the exact number of applicant or inventor countries and this will subsequently be used to create other axioms.

Having identified the closure axioms which will enable equivalent classes to be represented with negations and having differentiated the applicant countries from inventor countries, this step implements the axioms that represent the classes in Step 1. For example, the class *CrossBorder* is a class with two applicant countries and one inventor country where the inventor country is different from those of the applicants. For the sake of clarity in the figures and axioms, the case study has made a simplification here, which does not affect the global applicability of the illustration. The case study only includes the top five patent filing countries: USA, Japan, China, Germany and Korea (*US, JP, CN, DE* and *KR*). For the complete ontology (with all the countries), it is necessary to increase the number of countries to that defined by the World Patent International Organization (WIPO) in its Standard ST.3 (WIPO, 2011).

In description logic, this equivalent class is represented as:

$$
\begin{aligned}
CrossBorder \equiv\ & ApplicantCountries2 \sqcap InventorCountry1 \\
& \sqcap (\neg\big((\exists hasApplicantCountry.\{CNi\} \\
& \sqcap \exists hasInventorCountry.\{CNi\}) \\
& \sqcup\ (\exists hasApplicantCountry.\{JPi\} \sqcap \exists hasInventorCountry.\{JPi\}) \\
& \sqcup\ (\exists hasApplicantCountry.\{KRi\} \sqcap \exists hasInventorCountry.\{KRi\}) \\
& \sqcup\ (\exists hasApplicantCountry.\{USi\} \sqcap \exists hasInventorCountry.\{USi\}) \\
& \sqcup\ (\exists hasApplicantCountry.\{DEi\} \\
& \sqcap \exists hasInventorCountry.\{DEi\})\big))
\end{aligned}
$$

(Axiom 6.2)

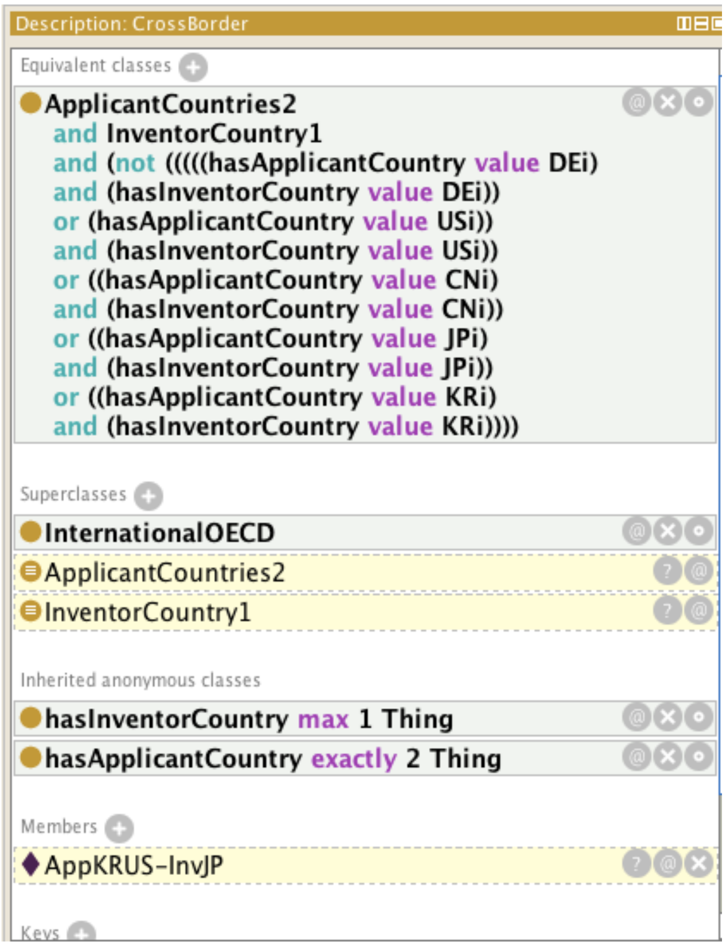

Figure 6.4 Class CrossBorder

6.4.6 Step 6. Invoke the Reasoner

Once all the classes, properties and relationships have been declared and represented, a OWL reasoner will infer the new knowledge derived from the implemented relationships. In the example, the reasoner Hermit has classified every patent document into its respective OECD class or KPAI. One example of this can be found in Figure 6.4, where the reasoner has inferred that the patent *AppKRUS-InvJP* (a patent with applicants in two countries *KR* and *US* and with inventors in one country *JP*) belongs to the class *CrossBorder*. Similar results have been achieved with the remaining patent documents, classified with the correspondent KPAI.

With these declarations of classes, properties and axioms, all the patents that will populate the ontology in the future will acquire the corresponding international classification.

6.5. Management of Changes

The proposed declarative ontology is particularly useful since it allows the classification to be efficiently adapted to future requirements. Our proposal allows an efficient way to implement modifications over the axioms. In a context of changing business and technological interests, the opportunities for a quick adaptation provides a robust alternative to analyze new situations and different interests.

It might be useful to consider different international corporate classifications, such as the one proposed by Rugman and Verbeke (2004) or the modifications subsequently proposed by Dunning et al., (2007). Even if the OECD changes its classification, the proposed method can be used to change or redefine the axioms, and the current and future patent documents will be reclassified accordingly. In this way, the declarative approach proposed allows patent documents to be analyzed without requiring further programming tasks to be performed. In the current context of increasing patent internationalization, for example, the OECD could be more restrictive in its classification by redefining *crossborder* to a class with two applicant countries and two inventor countries (instead of the single inventor country in the old *crossborder* definition) and where the inventor countries are different from those of the applicants. In this case, we must redefine *crossborder* by simply replacing *InventorCountry1* with *InventorCountry2* in Axiom 2 (see Axiom 6.3 and Figure 6.5).

$$
\begin{aligned}
&NewCrossBorder \\
&\qquad \equiv ApplicantCountries2 \sqcap InventorCountries2 \\
&\qquad \sqcap \Big(\neg\big((\exists hasApplicantCountry.\{CNi\} \\
&\qquad \sqcap \exists hasInventorCountry.\{CNi\}) \\
&\qquad \sqcup \ (\exists hasApplicantCountry.\{JPi\} \sqcap \exists hasInventorCountry.\{JPi\}) \\
&\qquad \sqcup \ (\exists hasApplicantCountry.\{KRi\} \sqcap \exists hasInventorCountry.\{KRi\}) \\
&\qquad \sqcup \ (\exists hasApplicantCountry.\{USi\} \sqcap \exists hasInventorCountry.\{USi\}) \\
&\qquad \sqcup \ (\exists hasApplicantCountry.\{DEi\} \\
&\qquad \sqcap \exists hasInventorCountry.\{DEi\})\big)\Big)
\end{aligned}
$$

(Axiom 6.3)

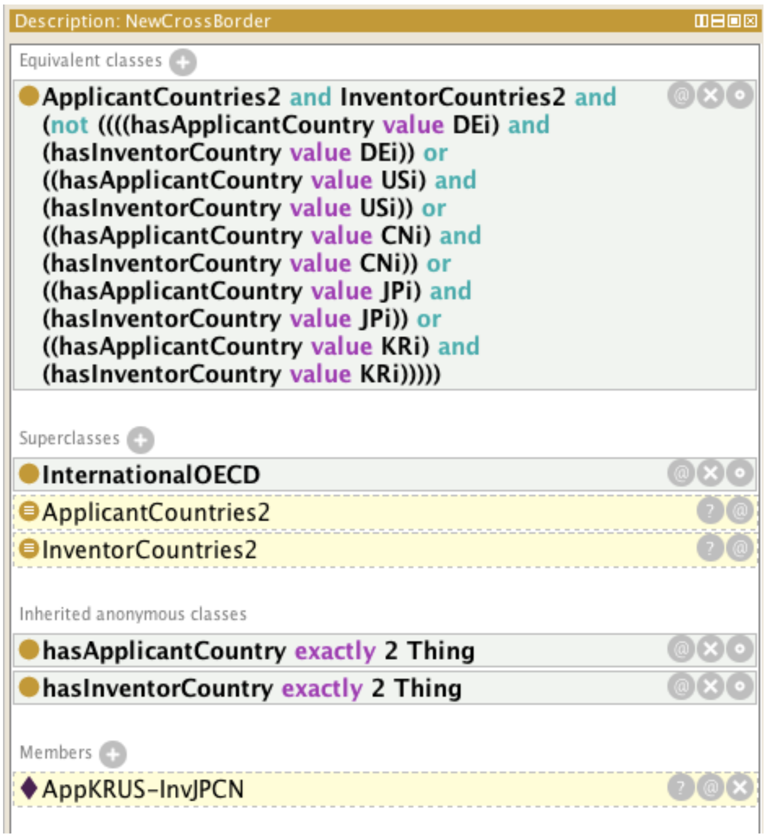

Figure 6.5 Redefinition of the class CrossBorder

In this case, instead of classifying the patent *AppKRUS-InvJP* as the *Crossborder* patent, the reasoner classifies the patent *AppKRUS-InvJPCN* (where Korea and USA are the applicant countries and Japan and China are the inventor countries) as a *crossborder* patent.

6.6 Evaluation of the Reasoning Process

In order to evaluate the feasibility and reasonability of the proposed method, this section shows the reasoning performance (Step 6 of the proposal method), with the two most popular tableaux-based reasoners: Pellet 2.2.0 and Hermit 1.3.5. These two reasoners are used with expressivity OWL-2 DL, and both as plug-ins of the editing ontology tool Protégé version 4.2.0 (build 249) with OWLAPI integrated.

The case study has been built on PatentOntology. Firstly, we have removed all the individuals from PatentOntology and populated it with the controlled individuals that we need for our analysis. Secondly, we have implemented our method by adding new classes, properties, individuals and axioms to PatentOntology.

The complete ontology is written in OWL-2 DL. The main ontology metrics are: it has 79 classes, 45 object properties, 21 individuals (other than patent documents), 73 subclass axioms, 12 equivalent classes axioms (all of them written for this chapter), 1 disjoint class axiom, 65 class assertion axioms and 51 object property assertion axioms. All the experiments are performed on a notebook running OS X 10.6.8 and Java 1.6 on an Intel® Core™2 Duo CPU at 1.86 GHz. Times are taken as the average over 10 independent runs as they were in previous work (Kazakov et al., 2011; Kang et al., 2012).

Table 6.1 shows the performance of the two reasoners when carrying out the most relevant tasks for our case study. These measures have been taken with different sizes of the ontology ranging from 20 to 50 patent documents. Time is shown in milliseconds (ms).

Table 6.1 Performance of the reasoners in the case study

		Pellet		Hermit	
Number of patent documents		20	50	20	50
Displayed Class Inferences	Unsatisfiability	1 .3	2.2	0 .8	1 .3
	Equivalent classes	3 .5	7.5	1 .7	1 .7
	Superclasses	1 .7	1.9	1 .6	1 .8
	Class Members	1 .8	2.8	1 .7	2 .2
	Inherited classes	0 .2	0.4	0 .3	0 .4
	Disjoint classes	230 .1	677.8	34	304 .2

Table 6.1 shows that performance of the two reasoners is feasible and enables the reasoners to determine all the subsumption relationships without incurring inconsistencies in a few seconds. The difference in times between the two reasoners shown in Table 6.1 could be due to the different underneath algorithms. The performance of the algorithms could be different, and the scope and the properties of the logic covered could be different, as well. In comparison with traditional patent analysis, the researcher's time is therefore significantly reduced. In traditional patent analysis, the researcher must manually retrieve patent documents that meet certain requirements (such as OECD classification) and perform the retrieval process whenever new patent documents are published. With our method, the researcher only needs to write a few axioms and the reasoners will automatically classify all the patent documents in a reduced time. Furthermore, these axioms will be valid for future patent documents in which case the reasoners will only take a matter of seconds to execute automatic reclassification.

6.7 Conclusions

Patents are good indicators of the processes and outcomes of inventive activities (e.g. the location of inventive activities, inventive networks or emerging technologies), the technological competitiveness of countries, and the cooperation between institutions, firms and countries in the patenting process. Since patents are stored in large databases where this co-operation is not explicitly connected, it is not therefore possible to automatically process these relationships by computer. Ontology languages such as OWL have proved useful for representing knowledge and creating new relationships which are not explicitly written in traditional databases.

This work highlights the usefulness of identifying and representing complex relationships in ontologies to automatically process information which is otherwise hidden in databases and to declaratively identify key analysis parameters.

The chapter first defined the specific steps required to represent complex relationships in OWL by representing the KPAIs in order to create new classifications which are not explicitly shown in databases. We also demonstrated and implemented our method using an example in which the KPAIs were the international patent classifications proposed by the OECD. In the current competitive context in which firms are increasing their innovative activities on an international level (OECD, 2009), it is particularly important to analyze the internationalization flows of patents. In our case study, the six steps of the method are represented. The case study uses the OWL class and property declarations and axioms (including closure axioms) to implement equivalent classes. These classes are then used by a reasoner to classify patents into their corresponding KPAI classes which in this particular case are the OECD classification classes.

Our research highlights the flexibility of ontologies when new descriptions of key concepts, or KPAIs, are proposed or needed. Reasoning enables relationships to be efficiently defined in the ontology so that all the information may be automatically reclassified within this new classification. A future combination of artificial intelligence algorithms and this reasoning proposal would be especially useful, not only for the analysis of patent metadata, but also to explicitly show the relationships of any data gathered in patents. An effective system to better understand a firm's international patent portfolio is also useful for managers in a context of growing internationalization.

This chapter has practical implications for engineers and researchers, enabling a systematic approach to analyze the complex and extensive patent information and reducing the time of the analysis by automatizing the classification defined by the axioms that we declare in the ontology. The interest of extending ontologies with the representation of complex relationships is particularly highlighted in this chapter. Managers could also benefit from the development of this proposal since it could result in more efficient ways to analyze the firm's patent portfolio and implement the correct patent filing policies. Although this chapter presents a method for representing complex relationships with patent data alone, future studies might also add financial data from different databases to the current ontology. With this data combination, managers and researchers could be more effective in identifying how international patenting decisions affect a firm's performance. Although the proposed method has been applied to a specific case study, this method could be used to any patent analysis, reinforcing its interest and applicability.

References

Almeida, P., & Phene, A. (2004). Subsidiaries and Knowledge Creation: the Influence of the MNC and Host Country on Innovation. *Strategic Management Journal*, *25*(89), 847–864.

Baader, F. (2003). *The Description Logic Handbook: Theory, Implementation and Applications*. Cambridge University Press.

Bettini, C., Brdiczka, O., Henricksen, K., Indulska, J., Nicklas, D., Ranganathan, A., & Riboni, D. (2010). A Survey of Context Modelling and Reasoning Techniques. *Pervasive and Mobile Computing*, *6*(2), 161–180.

Choi, S., Kang, D., Lim, J., & Kim, K. (2012). A Fact-oriented Ontological Approach to SAO-based Function Modeling of Patents for Implementing Function-based Technology Database. *Expert Systems with Applications*, *39*(10), 9129–9140.

Ciabuschi, F., Dellestrand, H., & Martín, O. M. (2011). Internal Embeddedness, Headquarters Involvement, and Innovation Importance in Multinational Enterprises. *Journal of Management Studies*, *48*(7), 1612–1639.

Dahlem, N., & Hahn, A. (2009). User-Friendly Ontology Creation Methodologies - A Survey. *Americas Conference on Information Systems AMCIS*.

Darlington, M. J., & Culley, S. J. (2008). Investigating Ontology Development for Engineering Design Support. *Advanced Engineering Informatics*, (*22)*, 112–134.

Dunning, J. H., Fujita, M., & Yakova, N. (2007). Some Macro-data on the Regionalisation/ Globalisation Debate: A Comment on the Rugman/Verbeke Analysis. *Journal of International Business Studies*, (*38)*, 177–199.

Ghassan B., Lopez-Lorca, A. A., García-Sánchez, F., & Martínez-Béjar, R. (2011). How do we Measure and Improve the Quality of a Hierarchical Ontology? *Journal of Systems and Software*, *84*(12), 2363–2373.

Giereth, M., St, A., Rotard, M., & Ertl, T. (2006). Application of Semantic Technologies for Representing Patent Metadata. *1st International Workshop on Applications of Semantic Technologies*, (*94)*, 297–305.

Grant, R. M. (1991). The Resource-Based Theory of Competitive Advantage: Implications for Strategy Formulation. *California Management Review,* (33), 114-135.

Grimm, S., & Boris, M. (2005). Closed World Reasoning in the Semantic Web through Epistemic Operators. *OWL: Experiences and Directions Proceedings*.

Gruber, T. R. (1993a). A Translational Approach to Portable Ontology Specifications. *Knowledge Acquisition*, *5*(2), 199–220.

Gruber, T. R. (1993b). Toward Principles for the Design of Ontologies Used for Knowledge Sharing. *International Journal Human Computer Studies*, *43*(5/6), 907–928.

Guarino, N., & Giaretta, P. (1995). Ontologies and Knowledge Bases: Towards a Terminological Clarification. *Towards Very Large Knowledge Bases Knowledge Building and Knowledge Sharing,* (1), 25–32.

Gurry, F., Fink, C., Khan, M., Bergquist, K., Lamb, R., Feuvre, B. Le, & Hao Zhou. (2012). *World Intellectual Property Indicators 2011*. WIPO Economics & Statistics Series. Available at http://www.wipo.int/ipstats/en/wipi.

Kang, Y. B. and Li, Y. F. and Krishnaswamy, S. (2012). Predicting Reasoning Performance Using Ontology Metrics. *Lecture Notes in Computer Science*, (*7649)*, 198–214.

Kazakov, Y., Krötzsch, M., & Simancík, F. (2011). Concurrent Classification of EL Ontologies. *Lecture Notes in Computer Science*, (*7031)*, 305–320.

Kim, Y. G., Suh, J. H., & Park, S. C. (2008). Visualization of Patent Analysis for Emerging Technology. *Expert Systems with Applications*, *34*(3), 1804–1812.

Knublauch, H., Fergerson, R.W., Noy, N.F., Musen, M. (2004). The Protégé OWL Plugin: An Open Development Environment for Semantic Web Applications. *Lecture Notes in Computer Science*, (*3298)*, 229–243.

Lahiri, N. (2010). Geographic Distribution of R&D Activity: How does it Affect Innovation Quality? *Academy of Management Journal*, *53*(5), 1194–1209.

Noguera, M., Hurtado, M. V., Rodríguez, M. L., Chung, L., & Garrido, J. L. (2010). Ontology-driven Analysis of UML-based Collaborative Processes using OWL-DL and CPN. *Science of Computer Programming*, *75*(8), 726–760.

Noy, N. F., & McGuinness, D. L. (2001). *Ontology Development 101: A Guide to Creating Your First Ontology*. University of Stanford. Available at http://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinness.html.

OECD. (2009). *OECD Patent Statistics Manual*. Available at http://www.oecd.org/science/innovationinsciencetechnologyandindustry/oecdpatentstatisticsmanual.htm.

Panetto, H., Dassisti, M., & Tursi, A. (2012). ONTO-PMD: Product-driven ONTOlogy for Product Data Management Interoperability within Manufacturing Process Environment. *Advanced Engineering Informatics*, *26*(2), 334–348.

Pelegrina, A. B., Benghazi, K., Hurtado, M. V., Noguera, M., Barjis, J., Aalst, W., Mylopoulos, J., et al. (2010). A Framework for the Semantic Representation of Business Processes within Business Organizational Models. *Enterprise and Organizational Modeling and Simulation* (63), 79–94.

Quintas, M., Vazquez, X., Garcia, J., & Caballero, G. (2008). Geographical Amplitude in the International Generation of Technology: Present Situation and Business Determinants. *Research Policy*, *37*(8), 1371–1381.

Shearer, R., Motik, B., & Horrocks, I. (2008). HermiT: A Highly-efficient OWL Reasoner. *Workshop on OWL: Experiences and Directions (OWLED*), 26-27.

Rocca, G. L. (2012). Knowledge Based Engineering: Between AI and CAD. Review of a Language Based Technology to Support Engineering Design. *Advanced Engineering Informatics*, *26*(2), 159–179.

Rugman, A. M., & Verbeke, A. (2004). A Perspective on Regional and Global Strategies of Multinational Enterprises. *Journal of International Business Studies*, *1*(35), 3–19.

Seol, H., Lee, S., & Kim, C. (2011). Identifying New Business Areas using Patent Information: A DEA and Text Mining Approach. *Expert Systems with Applications*, *38*(4), 2933–2941.

Shih, M., & Liu, D. (2010). Patent Classification using Ontology-based Patent Network Analysis. *Proceedings Pacific Asian Conference in Information systems*.

Shih, M. J., Liu, D. R., & Hsu, M. L. (2010). Discovering Competitive Intelligence by Mining Changes in Patent Trends. *Expert Systems with Applications*, *37*(4), 2882–2890.

Sirin, E., Parsia, B., Grau, B., Kalyanpur, A., & Katz, Y. (2007). Pellet: A Practical OWL-DL Reasoner. *Web Semantics Science Services and Agents on the World Wide Web*, *5*(2), 51–53.

Smith, M. K., Welty, C., & McGuinness, D. L. (2004). *OWL Web Ontology Language Guide*. W3C Recommendation. The World Wide Web Consortium (W3C). Available at http://www.w3.org/TR/owl-guide.

Somaya, D. (2012). Patent Strategy and Management: An Integrative Review and Research Agenda. *Journal of Management*, *38*(4), 1084–1114.

Soo, V. W., Lin, S. Y., Yang, S. Y., Lin, S. N., & Cheng, S. L. (2006). A Cooperative Multi-agent Platform for Invention Based on Patent Document Analysis and Ontology. *Expert Systems with Applications*, *31*(4), 766–775.

Taduri, S., Lau, G. T., Law, K. H., Yu, H., & Kesan, J. P. (2011a). An Ontology to Integrate Multiple Information Domains in the Patent System. *IEEE International Symposium on Technology and Society ISTAS*, 23–25.

Trappey, A. J. C., Hsu, F. C., Trappey, C. V., & Lin, C. I. (2006). Development of a Patent Document Classification and Search Platform using a Back-propagation Network. *Expert Systems with Applications*, *31*(4), 755–765.

Trappey, A. J. C., Trappey, C. V., Wu, C. Y., & Lin, C. W. (2012). A Patent Quality Analysis for Innovative Technology and Product Development. *Advanced Engineering Informatics*, *26*(1), 26–34.

Trappey, A., Trappey, C., & Wu, C. Y. (2009). Automatic Patent Document Summarization for Collaborative Knowledge Systems and Services. *Journal of Systems Science and Systems Engineering, 18(1),* 71–94.

Vries, D. (2012). *Leveraging Patents Financially. A Company Prespective*. Springer.

W3C. (2009). OWL - Ontology Web Language. W3C. Available at http://www.w3.org/TR/owl-features.

Wanner, L., Brügmann, S., Diallo, B., Giereth, M., Kompatsiaris, Y., Pianta, E., Rao, G., et al. (2009). PATExpert : Semantic Processing of Patent Documentation. Available at http://recerca.upf.edu/patexpert.

Weissman, A., Petrov, M., & Gupta, S. K. (2011). A Computational Framework for Authoring and Searching Product Design Specifications. *Journal Advanced Engineering Informatics*, *25*(3), 516–534.

WIPO. (2011). *Standard st.3 Recommended Standard on Two-letter Codes for the Representation of States, other Entities and Intergovernmental Organizations*. WIPO. Available at http://www.wipo.int/standards/en/part_03_standards.html

Wolstencroft, K., Stevens, R., & Haarslev, V. (2007). Applying OWL Reasoning to Genomic Data. *Semantic Web,* 225–248.

CAPÍTULO 7

CONCLUSIONES

CONCLUSIONES

7.1 Introducción

El presente capítulo presenta un resumen y una discusión de las principales aportaciones obtenidas gracias a los estudios realizados en los capítulos anteriores de esta tesis, haciendo referencia expresa a las implicaciones y beneficios resultantes de la investigación desarrollada, y futuras líneas de investigación que se abren al hilo de los objetivos alcanzados.

7.2 Conclusiones del Trabajo de Investigación

Este trabajo pretende aportar una mayor comprensión sobre los importantes vínculos existentes entre las innovaciones en gestión empresarial para la preservación del medio ambiente y las tecnologías de la información y las comunicaciones (TIC). Asimismo, el trabajo ayuda a entender el interés de analizar conjuntamente esos ámbitos y, de forma más específica, el soporte que pueden ofrecer los actuales métodos y técnicas TIC, y su implementación como herramientas y aplicaciones, a la gestión de los procesos para la innovación medioambiental. De este modo, podemos concluir que:

- Los resultados muestran que la utilización de las TIC en el ámbito de gestión medioambiental permite, no sólo reforzar el proceso interno de gestión de la empresa, sino también incrementar su reputación, transparencia y legitimidad en los distintos mercados donde actúan.

- El trabajo realizado en la tesis muestra que la dotación de semántica a la información sobre innovación medioambiental en la empresa, especialmente mediante las aplicaciones relacionadas con las ontologías, es una vía apropiada para la recopilación de datos de distintas fuentes de información, y para la inferencia de relaciones no explícitas en dichas fuentes.

A continuación, se resumen las conclusiones en relación con cada uno de los capítulos de la tesis.

El Capítulo 3 titulado *"The importance of corporate websites for diffusion of environmental innovation: The influence of trusting beliefs for recruiting processes"* estudia la confianza de los *stakeholders* hacia una empresa, en cuestiones relativas al medio ambiente, basándose en información consultada por los usuarios a través de una Web corporativa. Las principales conclusiones obtenidas son:

- Los resultados del estudio son consistentes con la Teoría de la Acción Razonada, que afirma que las creencias afectan positivamente la actitud (Fishbein & Ajzen, 1975).

- Se muestra la influencia específica de las tres dimensiones de la confiabilidad en los entornos on-line (Larzelere & Huston, 1980): habilidad, integridad y benevolencia.

- Los resultados complementan investigaciones previas (Gefen et al., 2008), mostrando la influencia positiva de la integridad y benevolencia en la disposición de los usuarios de la Web corporativa a aceptar un trabajo en la empresa.

- Esta investigación pone de manifiesto que el uso de experimentos de laboratorio es adecuado para abordar cuestiones de investigación en la difusión de información haciendo uso de las TIC.

El cuarto capítulo de la tesis se titula *“The relationship between patented environmental innovations and firm performance: The influence of the international scope”* y estudia la relación entre la internacionalización de la innovación medioambiental patentada y el desempeño financiero de la empresa. Las principales conclusiones son:

- Esta investigación analiza, en base a la perspectiva de recursos naturales (Hart, 1995), la importancia de la protección de la innovación medioambiental de una empresa y la influencia en su desempeño financiero.

- Los resultados complementan la literatura que analiza las estrategias de gestión de patentes (Lippman & Rumelt, 2003), prestando especial atención a la internacionalización de las mismas.

    - El trabajo desarrollado muestra que una mayor dispersión internacional en las fuentes de conocimiento de la innovación medioambiental patentada modera negativamente la relación entre el número de innovaciones medioambientales patentadas y el rendimiento financiero de la empresa. Es decir, una amplia dispersión internacional en el suministro de conocimiento puede aumentar la complejidad más allá de niveles aceptables (Kotabe et al., 2007).

    - Los resultados destacan que un mayor alcance internacional de la explotación de las patentes medioambientales fortalece la relación positiva entre el número de patentes y el desempeño financiero, pudiendo reducir los riesgos de las fluctuaciones de los ingresos derivados de las patentes (Lu & Beamish, 2004; Madsen, 2009; Pinkse & Kolk, 2012; Pinkse et al., 2010).

El capítulo quinto de la tesis *"Adding sense to patent ontologies: A representation of concepts and reasoning*" aborda el problema de la carencia de anotaciones semánticas en la información sobre patentes almacenada en grandes bases de datos. Estas anotaciones semánticas permitirían la creación de relaciones y vínculos entre las patentes y otros dominios del conocimiento. Las principales conclusiones son:

- Se ha definido una ontología que permite dotar de semántica a la información sobre patentes ampliando los desarrollos previos existentes (Taduri et al., 2011; Wanner et al., 2009).
    - Se han representado los códigos tecnológicos de patentes en OWL (HCOntology), extrayendo toda la información implícita que guardan en su estructura. Esta representación permite crear relaciones transversales entre los conceptos, que pueden ser procesadas automáticamente por máquinas, sin necesidad de realizar cambios en cualquier otra aplicación que haga uso de la ontología definida.
    - Se ha enriquecido el proceso de recuperación de la información, a través de la indexación automática de los códigos jerárquicos en sus partes estructurales, permitiendo que los futuros códigos emergentes puedan ser traducidos automáticamente.
- Se presentan los resultados de aplicar la ontología a dos casos de estudio que ayudan a validar la propuesta realizada a la vez que se resalta su potencial.
    - En el primero se pone de manifiesto la ventaja de tener los códigos jerárquicos para inferir nuevas relaciones en el dominio de las patentes.

En concreto, este caso de estudio destaca la flexibilidad que ofrece HCOntology para la inclusión de nuevos códigos tecnológicos.

- En el segundo caso de estudio se pone de relevancia la ventaja que ofrece HCOntology en la creación de relaciones entre dominios de distintos ámbitos, en particular, entre el dominio de las patentes y el dominio financiero.

El capítulo sexto, titulado "*Analyzing a firm's international portfolio of technological knowledge: A declarative ontology-based OWL approach on patent documents*", pone de manifiesto la utilidad de identificar y representar relaciones complejas en ontologías. Estas relaciones permiten procesar automáticamente información que de otro modo está oculta en bases de datos y facilitar la delimitación de forma declarativa de los principales parámetros de análisis. Las conclusiones de esta investigación son:

- Se propone un nuevo método para guiar a los analistas de patentes en el proceso de obtención de información relevante utilizando la ontología definida y herramientas asociadas. En este sentido, es posible identificar e incluir nuevas relaciones en la ontología definida con el lenguaje OWL permitiendo la inferencia automática, a través de un razonador basado en lógica descriptiva, de nueva información no explícita en las bases de datos de patentes, clasificando los documentos de patentes en base a las relaciones definidas.

- La investigación desarrollada pone de relieve la flexibilidad de las ontologías para inferir nueva información, ya que en concreto permite:

- Definir relaciones entre conceptos de la ontologías que son válidas para documentos de patentes actuales y futuros puesto que las relaciones correspondientes ya se incluyeron en la ontología, a diferencia de lo que ocurre en proceso de consultas en bases de datos.

- Facilitar el mantenimiento ya que los cambios en las relaciones definidas en la ontología no implican ningún cambio en aplicaciones que hagan uso de dicha ontología.

- El método propuesto implica la automatización de procesos y, por tanto, el uso de ontologías permite reducir tiempo en los análisis complejos de carteras de patentes.

7.3 Implicaciones y Beneficios de la Investigación

Los resultados de la investigación desarrollada en esta tesis tienen claras implicaciones y beneficios para académicos, gestores e ingenieros. Con carácter general destacamos que los resultados obtenidos permiten destacar la importancia de prestar especial atención al papel que la gestión de las TIC tiene a la hora de planificar e implementar las prácticas de gestión medioambiental de las organizaciones. Al mismo tiempo, los ingenieros responsables del desarrollo de aplicaciones podrían también considerar algunas de nuestras conclusiones para prestar atención preferente a avances orientados a un mejor servicio hacia los usuarios de los mismos.

A continuación se describen las implicaciones de forma más precisa:

El tercer capítulo de la tesis tiene implicaciones importantes en el ámbito de los problemas específicos de confianza hacia las empresas en entornos de Internet. Los resultados sugieren:

- La importancia de considerar las Webs corporativas como un "espacio sensorial emotivo en el que somos capaces de cambiar la Web corporativa de un ambiente emocionalmente plano a un espacio de interacciones ricas" (Kambil, 2008). En este contexto, los administradores pueden utilizar observaciones específicas realizadas en esta tesis doctoral en relación a la gestión de contenidos de la Web corporativa.

- Los gerentes deben tratar de especificar en la Web corporativa información general y específica acerca de su comportamiento empresarial y sus intenciones. Esta información debe considerar las múltiples situaciones en las que los usuarios pueden querer examinar la Web corporativa para apoyar sus decisiones finales en cuanto a confiar en la empresa.

- La importancia de prestar una atención específica a los factores con potencial para generar una mayor confianza en la empresa a través de las dimensiones de integridad y benevolencia.

El cuarto capítulo de la tesis plantea implicaciones importantes en el campo de la gestión de las carteras de patentes. De forma más específica se destaca que:

- Las empresas no sólo necesitan aumentar sus patentes medioambientales sino, sobre todo, gestionarlas adecuadamente. En este sentido, un alcance internacional más amplio en la explotación de las patentes, mejora la capacidad de una cartera de patentes para incrementar el resultado financiero de la empresa.

- La mayor parte de las empresas no son activas en la generación de patentes medioambientales y los responsables políticos deben analizar qué circunstancias pueden estar llevando a esa situación.

Los capítulos quinto y sexto de la tesis ponen de relevancia la importancia de aplicar técnicas y métodos basados en ontologías en el soporte a los procesos de gestión de la empresa, especialmente al análisis de la innovación en general, y de las patentes, en particular. De forma más específica se destaca que:

- El interés de extender las ontologías con la representación de relaciones complejas es especialmente relevante y tiene implicaciones prácticas para ingenieros e investigadores, ya que permite una mejor comprensión de las diferentes formas de analizar el gran volumen de información disponible sobre patentes.

- Los gestores de las empresas también pueden beneficiarse del desarrollo de esta propuesta, ya que podría dar lugar a formas más eficientes para analizar la cartera de patentes de la empresa e implementar las políticas de patentes correctas.

- La dotación de semántica a los códigos jerárquicos tecnológicos es relevante para los usuarios de los repositorios de patentes, ya que facilitará los análisis tanto de las patentes propias como de las patentes de competidores asociadas a un campo tecnológico concreto.

- Es importante destacar que aunque el método desarrollado se ha aplicado en el dominio de las patentes, podría igualmente ser aplicado a cualquier dominio con códigos jerárquicos. Los códigos jerárquicos enriquecidos con semántica permitirían una definición más flexible de nuevas relaciones y propiedades, además de inferir nueva información.

- La automatización en el proceso de recuperación de la información de las bases de datos de patentes y el volcado a la ontología, abre también perspectivas interesantes en un acceso más eficiente a grandes volúmenes de información. En nuestro caso particular, los gestores e investigadores podrán desarrollar análisis con datos de patentes e información obtenida de bases de datos financieras más eficientes, debido a la automatización del volcado de datos y al semi-procesamiento que esto implica.

7.4 Futuras Líneas de Investigación

Las conclusiones alcanzadas mediante el desarrollo de esta tesis doctoral son una base apropiada para profundizar más en el conocimiento de aquellas condiciones tecnológicas que ayudan a un mejor desempeño medioambiental y gestión de la innovación. Por ello, consideramos que nuestra investigación puede extenderse en varias direcciones.

- Avanzar en la investigación sobre las características de las tecnologías TIC que pueden permitir a los administradores determinar los contenidos y optimizar continuamente la Web corporativa de acuerdo a los procesos de negocio específicos y objetivos corporativos (Neubauer & Stummer, 2009), al tiempo que evaluar las implicaciones de su utilización.

- Analizar los antecedentes de la falta de confianza hacia una empresa, especialmente con el crecimiento de las redes sociales en Internet, que puede reforzar la difusión de las experiencias insatisfactorias con las empresas.

- Analizar factores tales como el grado y la frecuencia de la colaboración entre las diferentes unidades de la empresa involucradas en el desarrollo de la innovación medioambiental. Además, consideramos relevante el desarrollo de trabajos en los que se analice la complementariedad de los procesos internos de generación de innovación y los procesos de "*open-innovation*".

- Realizar una propuesta específica en la que se combinara la ontología de metadatos de patentes propuesta en esta tesis, con algoritmos de inteligencia artificial (Choi et al., 2012; Trappey et al., 2009; Trappey et al., 2012), que extraigan semántica del texto incluido por ejemplo en el título, resumen, descripción y reivindicaciones de los documentos de patentes. Esto podría suponer un avance enriquecedor en la ontología a la par que establecería un puente entre enfoques que hasta el momento se han venido desarrollando de manera independiente.

- Futuros estudios pueden explorar oportunidades del lenguaje OWL para generar y explotar las jerarquías de conceptos y analizar los datos en diferentes contextos distintos al dominio de patentes.

- En lo que se refiere al análisis de relaciones complejas con datos de patentes, los estudios futuros pueden también incorporar los análisis conjuntos de esos datos con los procedentes de datos financieros de diferentes bases de datos empresariales, con el objetivo de realizar un análisis riguroso que estudie los vínculos existentes entre las características de las patentes concedidas y el desempeño financiero de la empresa.

- Finalmente, proponemos abordar la conexión a través de SOAP ("*Service Object Access Protocol*") con las bases de datos de patentes de las oficinas de patentes en general, y de la Oficina Europea de Patentes (EPO), en particular. Esta conexión permitiría recuperar la información requerida en formato XML, pudiendo posteriormente aplicar las técnicas para la extracción de códigos tecnológicos propuestas en el desarrollo de la tesis. Esto posibilitaría el volcado completamente automático de los documentos de patentes desde las bases de datos a la ontología.

En definitiva, son varios los aspectos que requieren todavía una importante labor de los investigadores para tratar de comprender, y de mejorar, la utilización de herramientas TIC avanzadas por parte de las empresas buscando la mejora de sus niveles de innovación medioambiental en general y del desempeño financiero en particular. Esta tesis ha pretendido contribuir a la consecución de este objetivo general, así como abrir la puerta a futuros esfuerzos en este camino.

Referencias del Capítulo

Bosch, T., & Mathiak, B. (2011). XSLT Transformation Generating OWL Ontologies Automatically Based on XML Schemas. *International Conference for Internet Technology and Secured Transactions* 660–667.

Cardoso, J., & Bussler, C. (2011). Mapping Between Heterogeneous XML and OWL Transaction Representations in B2B Integration. *Data & Knowledge Engineering*, *70*(12), 1046–1069.

Choi, S., Kang, D., Lim, J., & Kim, K. (2012). A Fact-oriented Ontological Approach to SAO-based Function Modeling of Patents for Implementing Function-based Technology Database. *Expert Systems with Applications*, *39*(10), 9129–9140.

Doney, P. M., Cannon, J. P., & Mullen, M. R. (1998). Understanding the Influence of National Culture on Development of Trust. *Academy of Management Review*, *23*(3), 601–620.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behavior: an introduction to theory and research*. Addison Wesley.

Gefen, D., Benbasat, I., & Pavlou, P. (2008). A Research Agenda for Trust in Online Environments. *Journal of Management Information Systems*, *24*(4), 275–286.

Hart, S. L. (1995). A Natural-Resource-Based View of the Firm. *Academy of Management Review*, *20*(4), 986–1014.

Kambil, A. (2008). What is Your Web 5.0 Strategy? *Journal of Business Strategy*, *29*(6), 56–58.

Kotabe, M., Dunlap-Hinkler, D., Parente, R., & Mishra, H. A. (2007). Determinants of Cross-national Knowledge Transfer and its Effect on Firm Innovation. *Journal of International Business Studies*, *38*(2), 259–282.

Lacoste, D., Sawant, K. P., & Roy, S. (2011). An Efficient XML to OWL Converter. *Proceedings of the 4th India Software Engineering Conference*, 145–154.

Larzelere, R. E., & Huston, T. L. (1980). The Dyadic Trust Scale: Toward Understanding Interpersonal Trust in Close Relationships. *Journal of Marriage and the Family*, *42*(3), 595.

Lippman, S. A., & Rumelt, R. P. (2003). A Bargaining Perspective on Resource Advantage. *Strategic Management Journal*, *24*(11), 1069–1086.

Lu, J. W., & Beamish, P. W. (2004). International Diversification and Firm Performance: The S-Curve Hypothesis. *Academy of Management Journal*, *47*(4), 598–609.

Madsen, P. M. (2009). Does Corporate Investment Drive a "Race to the Bottom" in Environmental Protection? A Reexamination of the Effect of Environmental Regulation on Investment. *Academy of Management Journal*, *52*(6), 1297–1318.

Marcus, A. (2005). *Big Winners and Big Losers: The 4 Secrets of Long-term Business Success and Failure*. Wharton School Publishing.

Marcus, A. A., & Fremeth, A. R. (2009). Green Management Matters Regardless. *Academy of Management Perspectives*, *23*(3), 17–27.

Marcus, A., Aragon-Correa, J. A., & Pinkse, J. (2011). Firms, Regulatory Uncertainty, and the Natural Environment. *California Management Review*, *54*(1) 5-16.

Neubauer, T., & Stummer, C. (2009). Interactive Selection of Web Services under Multiple Objectives. *Information Technology and Management*, *11*(1), 25–41.

OECD. (2009). OECD Patent Statistics Manual. Available at http://www.oecd.org/science/innovationinsciencetechnologyandindustry/oecdpatentstatisticsmanual.htm.

Pinkse, J., & Kolk, A. (2012). Addressing the Climate Change—Sustainable Development Nexus. *Business Society*, *51*(1), 176–210.

Pinkse, J., Kuss, M. J., & Hoffmann, V. H. (2010). On the Implementation of a "Global" Environmental Strategy: The Role of Absorptive Capacity. *International Business Review*, *19*(2), 160–177.

Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so Different After All: A Cross-discipline View of Trust. *Academy of Management Review*, *23*(3), 393–404.

Russo, M. V., & Fouts, P. A. (1997). A Resource-Based Perspective on Corporate Environmental Performance and Profitability. *Academy of Management Journal*, *40*(3), 534-559.

Scharl, A., Gebauer, J., & Bauer, C. (2001). Matching Process Requirements with Information Technology to Assess the Efficiency of Web Information Systems. *Information Technology and Management*, *2*(2), 193–210.

Schoorman, F. D., Mayer, R. C., & Davis, J. H. (2007). An Integrative Model of Organizational Trust: Past, Present, and Future. *Academy of Management Review*, *32*(2), 344–354.

Sharma, S., & Vredenburg, H. (1998). Proactive Corporate Environmental Strategy and the Development of Competitively Valuable Organizational Capabilities. *Strategic Management Journal*, *19*(8), 729–753.

Shrivastava, P. (1995). The Role of Corporations in Achieving Ecological Sustainability. *Academy of Management Review*, *20*(4), 936–960.

Taduri, S., Lau, G. T., Law, K. H., Yu, H., & Kesan, J. P. (2011). Developing an Ontology for the U . S . Patent System. *Proceedings of the 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times*. 157-166.

Trappey, A. J. C., Trappey, C. V., Wu, C.-Y., & Lin, C.-W. (2012). A Patent Quality Analysis for Innovative Technology and Product Development. *Advanced Engineering Informatics*, *26*(1), 26–34.

Trappey, A., Trappey, C., & Wu, C.-Y. Automatic Patent Document Summarization for Collaborative Knowledge Systems and Services. *Journal of Systems Science and Systems Engineering,* 18(1), 71–94 (2009).

Veefkind, V., Hurtado-Albir, J., Angelucci, S., Karachalios, K., & Thumm, N. (2012). A New EPO Classification Scheme for Climate Change Mitigation Technologies. *World Patent Information*, 34(2), 106-111.

Wanner, L., Brügmann, S., Diallo, B., Giereth, M., Kompatsiaris, Y., Pianta, E., Rao, G., et al. (2009). PATExpert : Semantic Processing of Patent Documentation. *Available at ftp://ftp.cordis.europa.eu/pub/ist/docs/kct/patexpert-pr-poster-samt-2006-11-27_en.pdf*.

BIBLIOGRAFIA

Acs, Z. J., & Audretsch, D. B. (1987). Innovation, Market Structure, and Firm Size. *The Review of Economics and Statistics*, *69*(4), 567–574.

Aguilera-Caracuel, J., Aragón-Correa, J. A., Hurtado-Torres, N. E., & Rugman, A. M. (2011). The Effects of Institutional Distance and Headquarters' Financial Performance on the Generation of Environmental Standards in Multinational Companies. *Journal of Business Ethics*, *105*(4), 461–474.

Ajzen, I., & Fishbein, M. (1980). *Understanding Attitudes and Predicting Social Behavior*. Prentice-Hall.

Alavi, M., & Leidner, D. E. (2001). Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues. *MIS Quarterly*, *25*(1), 107.

Alberts L.K. (1993). *YMIR: An Ontology for Engineering Design*. PhD Tesis. University of Twente.

Alder, G. S., Noel, T. W., & Ambrose, M. L. (2006). Clarifying the Effects of Internet Monitoring on Job Attitudes: The Mediating Role of Employee Trust. *Information & Management*, *43*(7), 894–903.

Alexiev, V., Breu, M., & De Bruijn, J. (2005). *Information Integration with Ontologies : Experiences from an Industrial Showcase*. Wiley.

Almeida, P., & Kogut, B. (1999). Location of Knowledge and the Mobility of Engineers in Regional Networks. *Management Science*, *45*(7), 905–917.

Almeida, P., & Phene, A. (2004). Subsidiaries and Knowledge Creation: the Influence of the MNC and Host Country on Innovation. *Strategic Management Journal*, *25*(89), 847–864.

Anand, N., Yang, M., Van Duin, J. H. R., & Tavasszy, L. (2012). GenCLOn: An Ontology for City Logistics. *Expert Systems with Applications,* 39(15), 11944-11960.

Andersson, L. M., & Bateman, T. S. (2000). Individual Environmental Initiative: Championing Natural Environmental Issues in U.S. Business Organizations. *Academy of Management Journal*, *43*(4), 548–570.

Ankolenkar, A., Burstein, M., Hobbs, J., Lassila, O., Martin, D., McIlraith, S., … Zeng, H. (2001). DAML-S: A Semantic Markup Language For Web Services. *Proceedings of the First Semantic Web Working Symposium SWWS01 Stanford University*, 411–430.

Antoniou, G., & Harmelen, F. Van. (2009). Web Ontology Language: OWL. *Handbook on ontologies*, 91–110. Springer.

Aragón Correa, J. A., Hurtado Torres, N. E., & García Morales, V. J. (2005). Un Modelo Explicativo de las Estrategias Medioambientales Avanzadas para Pequeñas y Medianas Empresas y su Influencia en los Resultados. *Cuadernos De Economía y Dirección De La Empresa*, 25, 29-52.

Aragón-Correa, J. A. (1998). Strategic Proactivity and Firm Approach to the Natural Environment. *Academy of Management Journal*, *41*(5), 556–567.

Aragón-Correa, J. A., & Sharma, S. (2003). A Contingent Resource-Based View of Proactive Corporate Environmental Strategy. *Academy of Management Review*, *28*(1), 71–88.

Archibugi, D., & Pianta, M. (1992). Specialization and Size of Technological Activities in Industrial Countries: The analysis of Patent Data. *Research Policy*, *21*(1), 79–93.

Arora, A, Ceccagnoli, M., & Cohen, W. (2008). R&D and the Patent Premium. *International Journal of Industrial Organization*, *26*(5), 1153–1179.

Arora, A., & Fosfuri, A. (2000). Wholly Owned Subsidiary Versus Technology Licensing in the Worldwide Chemical Industry. *Journal of International Business Studies*, *31*(4), 555–572.

Arora, A., Fosfuri, A., & Gambardella, A. (2001). Markets for Technology and their Implications for Corporate Strategy. *Industrial and Corporate Change*, *10*(2), 419–451.

Au, Y. A., & Kauffman, R. J. (2008). The Economics of Mobile Payments: Understanding Stakeholder Issues for an Emerging Financial Technology Application. *Electronic Commerce Research and Applications*, *7*(2), 141–164.

Baader, F. (2003). *The Description Logic Handbook: Theory, Implementation and Applications*. Cambridge University Press.

Bansal, P. (2005). Evolving Sustainably: A Longitudinal Study of Corporate Sustainable Development. *Strategic Management Journal*, *26*(3), 197–218.

Bansal, P., & Hunter, T. (2003). Strategic Explanations for the Early Adoption of ISO 14001. *Journal of Business Ethics*, *46*(3), 289–299.

Bansal, P., & Roth, K. (2000). Why Companies Go Green: A Model of Ecological Responsiveness. *Academy of Management Journal*, *43*(4), 717–736.

Barkema, H. G., & Vermeulen, F. (1998). International Expansion Through Start-up or Acquisition: A Learning Prespective. *Academy of Management Journal*, *41*(1), 7–26.

Barney, J. B., & Clark, D. N. (2007). *Resource-Based Theory: Creating and Sustaining Competitive Advantage*. *Journal of Marketing Management*. Oxford University Press.

Bartlett, C. A., & Ghoshal, S. (1989). *Managing Across Borders The Transnational Solution*. *Harvard Business School Press Books*. Harvard Business School Press.

Bayardo, R. J., & Agrawal, R. (1999). Mining the Most Interesting Rules. *International Conference on Knowledge Discovery and Data Mining - KDD '99*, 145–154.

Bedini, I., Gardarin, G., & Nguyen, B. (2008). Deriving Ontologies from XML Schema. *Proceedings EDA 2008*.

Benbasat, I., and Zmud, R. W. (2003). The Identity Crisis Within the IS Discipline: Defining and Communicating the Discipline's Core Properties, *MIS Quarterly, 27*(2), 183-194.

Bermúdez-Edo, M., Hurtado-Torres, N., & Cordón-Pozo, E. (2008). The Influence of the Faculty Attitude on the Adoption of ICTs' Innovations for Learning Purposes: Analysis of a Shared Web. *Times of Convergence*. *Technologies Across Learning Contexts* (5192), 58–61.

Berners-Lee, B. T., Hendler, J., & Lassila, O. (2001). The Semantic Web. *Scientific American*, *284*(5), 34–43.

Bettini, C., Brdiczka, O., Henricksen, K., Indulska, J., Nicklas, D., Ranganathan, A., & Riboni, D. (2010). A Survey of Context Modelling and Reasoning Techniques. *Pervasive and Mobile Computing*, *6*(2), 161–180.

Bilton, N. (2010). An Explosion of Mobile Patent Lawsuits - NYTimes.com. *New York Times*. *bits*. Disponible en http://bits.blogs.nytimes.com/2010/03/04/an-explosion-of-mobile-patent-lawsuits.

Bohring, H., & Auer, S. (2005). Mapping XML to OWL Ontologies. *Leipziger InformatikTage*, *72*, 147–156.

Borst, W. N. (1997). *Construction of Engineering Ontologies for Knowledge Sharing and Reuse*. PhD Tesis Universiteit Twente.

Bosch, T., & Mathiak, B. (2011). XSLT Transformation Generating OWL Ontologies Automatically Based on XML Schemas. *International Conference for Internet Technology and Secured Transactions* 660–667.

Bound, J., Cummins, C., Griliches, Z., Hall, B. H., & Jaffe, A. (1982). *Who Does R&D and Who Patents?* University of Chicago Press.

Brickley, D., & Guha, R. V. (2004). *RDF Vocabulary Description Language 1.0: RDF Schema*. W3C Recommendation.

Brunnermeier, S. B., & Cohen, M. A. (2003). Determinants of Environmental Innovation in US Manufacturing Industries. *Journal of Environmental Economics and Management*, *45*(2), 278–293.

Burgos Jiménez, J. De, Céspedes Lorente, J. J., & Alvarez Gil, M. J. (2001). Un Análisis Exploratorio de las Estrategias Medioambientales y el Contexto Organizativo de los Hoteles Españoles. *Cuadernos de Economía y Dirección de la Empresa*, *8*, 5–32.

BusinessWire. (2008). Monster Worldwide Comments on Expectations for First Quarter 2008 Operating Expenses. Disponible en http://www.about-monster.com/archive/2008.

Buysse, K., & Verbeke, A. (2003). Proactive Environmental Strategies: A Stakeholder Management Perspective. *Strategic Management Journal*, *24*(5), 453–470.

Buzzacchi, L., & Scellato, G. (2008). Patent Litigation Insurance and R&D Incentives. *International Review of Law and Economics*, *28*(4), 272–286.

Cantwell, J. (2000). Technological Lock-in of Large Firms since the Interwar Period. *European Review of Economic History*, *4*(2), 147–174.

Cantwell, J., & Janne, O. (1999). Technological Globalisation and Innovative Centres: the Role of Corporate Technological Leadership and Locational Hierarchy. *Research Policy*, *28*(2-3), 119–144.

Cardoso, J. (2007). The Semantic Web Vision: Where Are We? *IEEE Intelligent Systems*, *22*(5), 84–88.

Cardoso, J., & Bussler, C. (2011). Mapping between Heterogeneous XML and OWL Transaction Representations in B2B Integration. *Data & Knowledge Engineering*, *70*(12), 1046–1069.

Carpano, C., J Chrisman, J., & Roth, K. (1994). International Strategy and Environment: An Assessment of the Performance Relationship. *Journal of International Business Studies*, *25*(3), 639–656.

Chandrasekaran, B., Josephson, J. R., & Benjamins, V. R. (1999). What are Ontologies, and Why do We Need Them? *IEEE Intelligent Systems*, *14*(1), 20–26.

Chaudhri, V. K., Farquhar, A., Fikes, R., Karp, P. D., & Rice, J. P. (1998). OKBC: A Programmatic Foundation for Knowledge Base Interoperability. *National Conference on Artificial Intelligence AAAI98*, 600–607.

Chen, D. Q., Mocker, M., Preston, D. S., & Teubner, A. (2010). Information Systems Strategy: Reconceptualization, Measurement, and Implications. *MIS Quarterly*, *34*(2), 233–259.

Chen, P. P. S. (1976). The Entity-Relationship Unified View of Data Model-Toward a Unified View Data. *ACM Transactions on Database Systems*, *1*(1), 9–36.

Chen, S. C., & Dhillon, G. (2003). Interpreting Dimensions of Consumer Trust in E-Commerce. *Information Technology and Management*, *4*(2), 303–318.

Chen, T. Y. (2011). Value Ontology-based Multi-aspect Intellectual Asset Valuation Method for Decision-making Support in K-commerce. *Expert Systems with Applications*, *38*(5), 5471–5485.

Choi, S., Kang, D., Lim, J., & Kim, K. (2012). A Fact-oriented Ontological Approach to SAO-based Function Modeling of Patents for Implementing Function-based Technology Database. *Expert Systems with Applications*, *39*(10), 9129–9140.

Choudhury, V., & Karahanna, E. (2008). The Relative Advantage of Electronic Channels: A Multidimensional View. *MIS Quarterly*, *32*(1), 179–200.

Christmann, P. (2000). Effects of "Best Practices" of Environmental Management on Cost Advantage: The Role of Complementary Assets. *Academy of Management Journal*, *43*(4), 663–680.

Christmann, P. (2004). Multinational Companies and the Natural Environment: Determinants of Global Environmental Policy Standardization. *Academy of Management Journal*, *47*(5), 747–760.

Christmann, P., & Taylor, G. (2001). Globalization and the Environment: Determinants of Firm Self-Regulation in China. *Journal of International Business Studies*, *32*(3), 439–458.

Christmann, P., & Taylor, G. (2006). Firm Self-regulation through International Certifiable Standards: Determinants of Symbolic Versus Substantive Implementation. *Journal of International Business Studies*, *37*(6), 863–878.

Chung, W., & Yeaple, S. (2008). International Knowledge Sourcing: Evidence from US Firms Expanding. *Strategic Management Journal*, *29*(11), 1207–1224.

Ciabuschi, F., Dellestrand, H., & Martín, O. M. (2011). Internal Embeddedness, Headquarters Involvement, and Innovation Importance in Multinational Enterprises. *Journal of Management Studies*, *48*(7), 1612–1639.

Clarke, J. (2007). Trust, Reputation, Integrity and Professionalism Reflections on Business Relationships. *Accountancy Ireland*, *39*(1), 54–58.

Cober, R., & Brown, D. (2007). *Direct Employers Association Recruiting Trends Survey*. Booz allen Hamilton. Disponible en http://www.jobcentral.com/pdfs/DEsurvey.pdf.

Codd, E. F. (1970). A relational Model of Data for Large Shared Data Banks. *Communications of the ACM*, *13*(6), 377–387.

Cohen, J., & Cohen, P. (1984). *Applied Multiple Regression: Correlation Analysis for the Behavioral Sciences*. (2nd ed.). Lawrence Erlbaum Associates.

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly*, *35*(1), 128–152.

Cohen, W. M., Nelson, R. R., & Walsh, J. P. (2000). *Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not)*. National Bureau of Economic Research. Disponible en http://www.nber.org/papers/w7552.

Cordano, M., & Frieze, I. H. (2000). Pollution Reduction Preferences of U.S. Environmental Managers: Applying Ajzen's Theory of Planned Behavior. *Academy of Management Journal*, *43*(4), 627–641.

Czarnitzki, D., & Kraft, K. (2000). Haftungsregeln und Innovation, Legal Form and Innovation. *Journal of Economics and Statistics (Jahrbuecher fuer Nationaloekonomie und Statistik)*, *220*(5), 513–526.

Dahlem, N., & Hahn, A. (2009). User-Friendly Ontology Creation Methodologies - A Survey. *Americas Conference on Information Systems AMCIS*.

Darlington, M. J., & Culley, S. J. (2008). Investigating Ontology Development for Engineering Design Support. *Advanced Engineering Informatics*, (*22)*, 112–134.

Darnall, N. (2006). Why Firms Mandate ISO 14001 Certification. *Business Society*, *45*(3), 354–381.

Darnall, N., Henriques, I., & Sadorsky, P. (2008). Do Environmental Management Systems Improve Business Performance in an International Setting? *Journal of International Management*, *14*(4), 364–376.

Davenport, T. H., & Prusak, L. (1997). *Information Ecology: Mastering the Information and Knowledge Environment*. Oxford University Press.

Davenport, T. H., & Prusak, L. (2000). *Working Knowledge: How Organizations Manage What They Know*. Harvard Business Press.

David, P. A., & Hall, B. H. (2006). Property and the Pursuit of Knowledge: IPR Issues Affecting Scientific Research. *Research Policy*, *35*(6), 767–771.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. *Management Science*, *35*(8), 982–1003.

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, *13*(3), 319–340.

De Burgos Jiménez, J., & Céspedes Lorente, J. J. (2001). Environmental Performance as an Operations Objective. *International Journal of Operations Production Management*, *21*(12), 1553–1572.

De-Miguel, A., Piattini, M., & Marcos, E. (1999). *Diseño de Bases de Datos Relacionales*. RA-MA.

Dernis, H., & Guellec, D. (2005). *Using Patent Counts for Cross-country Comparisons of Technology Output*. Organisation for Economic Co-operation and Development. Disponible en http://www.oecd.org/dataoecd/26/11/21682515.pdf.

Doney, P. M., Cannon, J. P., & Mullen, M. R. (1998). Understanding the Influence of National Culture on Development of Trust. *Academy of Management Review*, *23*(3), 601–620.

Dowell, G., Hart, S., & Yeung, B. (2000). Do Corporate Global Environmental Standards Create or Destroy Market Value? *Management Science*, *46*(8), 1059–1074.

Dunning, J. H., Fujita, M., & Yakova, N. (2007). Some Macro-data on the Regionalisation/ Globalisation Debate: A Comment on the Rugman/Verbeke Analysis. *Journal of International Business Studies*, (*38)*, 177–199.

Eckert, S., Dittfeld, M., Muche, T., & Rässler, S. (2010). Does Multinationality Lead to Value Enhancement? An Empirical Examination of Publicly Listed Corporations from Germany. *International Business Review*, *19*, 562–574.

Egri, C. P., & Herman, S. (2000). Leadership in the North American Environmental Sector: Values, Leadership Styles, and Contexts of Environmental Leaders and Their Organizations. *Academy of Management Journal*, *43*(4), 571–604.

European Commission. (2010a). *Europe's Digital Competitiveness Report*. Disponible en http://ec.europa.eu/ceskarepublika/pdf/press/ks7ec2.pdf.

European Commission. (2010b). *The 2010 report on R&D in ICT in the European Union*. JCR Scientific and Technical Reports. European Commission. Disponible en http://ftp.jrc.es/EURdoc/JRC57808.pdf

Fallside, D. C., & Walmsley, P. (2004). *XML Schema Part 0: Primer Second Edition*. W3C Recommendation. Disponible en http://www.w3.org/TR/xmlschema-0.

Fayyad, U., Piatetsky-Shapiro, G., Smyth, P., & Uthurusamy, R. (1996). *Advances in Knowledge Discovery and Data Mining*. The MIT Press.

Fensel, D., Horrocks, I., Harmelen, F. Van, Decker, S., Erdmann, M., & Klein, M. (2000). OIL in a Nutshell. *Knowledge Engineering Knowledge Management Methods, Models, and Tools*, *1937*(1937), 1–16.

Ferdinand, M., Zirpins, C., & Trastour, D. (2004). Lifting XML Schema to OWL. *Lecture Notes in Computer Science*, 3140, 354–358.

Fineman, S., & Clarke, K. (1996). Green Stakeholders: Industry Interpretations and Response. *Journal of Management Studies*, *33*(6), 715–730.

Fishbein, M., & Ajzen, I. (1975). *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*. Addison-Wesley.

Flannery, B. L., & May, D. R. (2000). Environmental Ethical Decision Making in the U.S. Metal-finishing Industry. *Academy of Management Journal*, *43*(4), 642–662.

Foglia, P. (2007). Patentability Search Strategies and the Reformed IPC: A Patent Office Perspective. *World Patent Information*, *29*(1), 33–53.

Freitas, A. A. (2003). A survey of Evolutionary Algorithms for Data Mining and Knowledge Discovery. *Advances in Evolutionary Computation* (136), 819–845.

Freitas, A. A. (1999). On Rule Interestingness Measures. *Knowledge-Based Systems*, *12*(5-6), 309–315.

Fundación-Entorno. (2011). *I+D+i y Cambio Climatico. Cómo aprovechar la capacidad tecnológica de la empresa española*. Fundación Entorno. Disponible en http://www.fundacionentorno.org/xtras/difusion/idi/FE_Publicacion_I+D+i.pdf

Furman, J. L., Porter, M. E., & Stern, S. (2002). The Determinants of National Innovative Capacity. *Research Policy*, *31*(6), 899–933.

Ganesan, P., Garcia-Molina, H., & Widom, J. (2003). Exploiting Hierarchical Domain Structure to Compute Similarity. *ACM Transactions on Information Systems*, *21*(1), 64–93.

Garcés Ayerbe, C., & Galve Górriz, C. (2001). Repercusión de las Inversiones en Protección del Medio Ambiente en la Productividad de las Empresas Españolas: Un Análisis Empírico. *Cuadernos de Economía y Dirección de la Empresa*, *8*, 33–50.

Garcia-Molina, H., Ullman, J. D., & Widom, J. (2009). *Database Systems: The Complete Book*. Prentice Hall.

Garrido, J. L. (2003). *AMENITIES: Una Metodología para el Desarrollo de Sistemas Cooperativos Basada en Modelos de Comportamiento y Tareas*. Tesis Doctoral. Universidad de Granada.

Garrido, J. L., Gea, M., Rodríguez, M. L. (2003). Requirements Engineering in Cooperative Systems.. *Requirements Engineering for Sociotechnical Systems*. 226-244.

Gatewood, R. D., Gowan, M. A., & Lautenschlager, G. J. (1993). Corporate Image, Recruitment Image, and Initial Job Choice Decisions. *Academy of Management Journal*, *36*(2), 414–427.

Gefen, D., Karahanna, E., & Straub, D. (2003). Potential and Repeat E-consumers: the Role of and Trust vis-a-vis TAM. *IEEE Transactions on Engeneering Management*, *50*(3), 307–321.

Gefen, D. (2002). Reflections on the Dimensions of Trust and Trustworthiness Among Online Consumers. *ACM SIGMIS Database*, *33*(3), 38–53.

Gefen, D., & Heart, T. (2006). On the Need to Include National Culture as a Central Issue in E-Commerce Trust Beliefs. *Journal of Global Information Management*, *14*(4), 1–30.

Gefen, D., & Straub, D. (2000). The Relative Importance of Perceived Ease of Use in IS Adoption : A Study of E-Commerce Adoption. *Journal of the Association for Information Systems*, *1*, 1–28.

Gefen, D., & Straub, D. W. (1997). Gender Differences in the Perception and Use of E-Mail: An Extension to the Technology Acceptance Model. *MIS Quarterly*, *21*(4), 389–400.

Gefen, D., Benbasat, I., & Pavlou, P. (2008). A Research Agenda for Trust in Online Environments. *Journal of Management Information Systems*, *24*(4), 275–286.

Gefen, D., Karahanna, E., & Straub, D. W. (2003). Trust and TAM in Online Shopping: An Integrated Model. *MIS Quarterly*, *27*(1), 51–90.

Ghassan B., Lopez-Lorca, A. A., García-Sánchez, F., & Martínez-Béjar, R. (2011). How do we Measure and Improve the Quality of a Hierarchical Ontology? *Journal of Systems and Software*, *84*(12), 2363–2373.

Giereth, M., St, A., Rotard, M., & Ertl, T. (2006). Application of Semantic Technologies for Representing Patent Metadata. *1st International Workshop on Applications of Semantic Technologies AST06*, *94*, 297–305.

Gill, A. S. (2008). The Role of Trust in Employee-manager Relationship. *International Journal of Contemporary Hospitality Management*, *20*(1), 98–103.

Gill, A. S., & Mathur, N. (2007). Improving Employee Dedication and Pro-social Behavior. *International Journal of Contemporary Hospitality Management*, *19*(4), 328–334.

Glimm, B., Horrocks, I., Motik, B., Shearer, R., & Stoilos, G. (2012). A Novel Approach to Ontology Classification. *Web Semantics: Science, Services and Agents on the World Wide Web*.

Goclenius, R. (1613). *Lexicon Philosophicum, Quo Tanquam Clave Philosophiae Fores Aperiuntur*. Vidua Matthiae Beckeri.

Goebel, M. W. E., & Hintermaier, F. S. (2011). Searcher's Little Helper – ICO Index Terms. *World Patent Information*, *33*(3), 260–268.

Goerzen, A., & Beamish, P. W. (2003). Geographic Scope and Multinational Enterprise Performance. *Strategic Management Journal*, *24*(13), 1289–1306.

Gomes, L., & Ramaswamy, K. (1999). An Empirical Examination of the Form of the Relationship Between Multinationality and Performance. *Journal of International Business Studies*, *30*(1), 173–187.

Gómez-Pérez, A., Fernández-López, M., & Corcho, O. (2004). *Ontological Engineering: with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web*. Springer.

Grant, R. M. (1991). The Resource-Based Theory of Competitive Advantage: Implications for Strategy Formulation. *California Management Review,* (33), 114-135.

Grant, R. M. (1996). Prospering in Dynamically-Competitive Environments: Organizational Capability as Knowledge Integration. *Organization Science*, *7*(4), 375–387.

Grau, B. C., Horrocks, I., Motik, B., Parsia, B., Patel-Schneider, P., & Sattler, U. (2008). OWL 2: The Next Step for OWL. *Web Semantics: Science, Services and Agents on the World Wide Web*, *6*(4), 309–322.

Greenpeace. (2011). *Guide to greener electronics. Version 17. Nov 2011.* Disponible en http://www.greenpeace.org/international/en/campaigns/climate-change/cool-it/Campaign-analysis/Guide-to-Greener-Electronics/?accept=84995a68d5f4ea201f3f289dd5fba946.

Greenspan, R. (2003). *Job Seekers Have Choices*. ClickZ Networks. Disponible en http://www.clickz.com/clickz/news/1699182/job-seekers-have-choices.

Greenwood, R., Diaz, A. M., Li, S. X., & Lorente, J. C. (2009). The Multiplicity of Institutional Logics and the Heterogeneity of Organizational Responses. *Organization Science*, *21*(2), 521–539.

Griliches, Z. (1990). Patent Statistics as Economic Indicators: A Survey. *Journal of Economic Literature*, *28*(4), 1661–1707.

Grimm, S., & Boris, M. (2005). Closed World Reasoning in the Semantic Web through Epistemic Operators. *OWL: Experiences and Directions Proceedings*.

Gruber, T. R. (1993a). A Translational Approach to Portable Ontology Specifications. *Knowledge Acquisition*, *5*(2), 199–220.

Gruber, T. R. (1993b). Toward Principles for the Design of Ontologies Used for Knowledge Sharing. *International Journal Human Computer Studies*, *43*(5/6), 907–928.

Grüninger, M., Atefi, K., & Fox, M. S. (2000). Ontologies to Support Process Integration in Enterprise Engineering. *Computational Mathematical Organization Theory*, *6*(4), 381–394.

Guarino, N. (1998). Formal Ontology in Information Systems. *First International Conference FOIS'98*, 3–15.

Guarino, N., (1997). Understanding, Building and Using Ontologies: A Commentary to "Using Explicit Ontologies in KBS Development". *International Journal of Human and Computer Studies*, *46*, 293–310.

Guarino, N., & Giaretta, P. (1995). Ontologies and Knowledge Bases: Towards a Terminological Clarification. *Towards Very Large Knowledge Bases Knowledge Building and Knowledge Sharing,* (1), 25–32.

Gurry, F., Fink, C., Khan, M., Bergquist, K., Lamb, R., Feuvre, B. Le, & Hao Zhou. (2012). *World Intellectual Property Indicators 2011. WIPO Economics & Statistics Series*. Disponible en http://www.wipo.int/ipstats/en/wipi/

Haase, P., Lewen, H., & Studer, R. (2008). The NeOn Ontology Engineering Toolkit. *World Wide Web*.

Hair, J., Andersson, R. E., Tatham, R. L., & Black, W., (2009). *Multivariate Analysis* (6th ed.). Prentice Hall.

Håkanson, L., & Nobel, R. (2001). Organizational Characteristics and Reverse Technology Transfer. *MIR Management International Review*, *41*(4), 395–420.

Haley, U. C. V.; Schuler, D. A. (2011). Government Policy and Firm Strategy in the Solar Photovoltaic Industry. *California Management Review,* (54), 17-38.

Hall, B. H., Griliches, Z., & Hausman, J. A. (1986). Patents and R&D: Is there a lag? *International Economic Review*, *27*(2), 265–283.

Hall, B. H., Jaffe, A., & Trajtenberg, M. (2005). Market Value and Patent Citations. *Rand Journal of Economics*, *36*(1), 16–38.

Hamilton, L. C. (2006). *Statistic with Stata Updated for Version 9*. Thomson Brooks.

Hanel, P. (2006). Intellectual Property Rights Business Management Practices: A survey of the Literature. *Technovation*, *26*(8), 895–931.

Harhoff, D., Scherer, F. M., & Vopel, K. (2003). *Exploring the Tail of Patented Invention Value Distributions*. Discussion Paper FS IV 97 - 27, Wissenschaftszentrum.

Hart, S. L. (1995). A Natural-Resource-Based View of the Firm. *Academy of Management Review*, *20*(4), 986–1014.

Hayes, P. (2004). *RDF Semantics*. W3C Recommendation.

Henriques, I., & Sadorsky, P. (1996). The Determinants of an Environmentally Responsive Firm: An Empirical Approach. *Journal of Environmental Economics and Management*, *30*(3), 381–395.

Herriot, P., & Rothwell, C. (1981). Organizational Choice and Decision Theory: Effects of Employers' Literature and Selection Interview. *Journal of Occupational Psychology*, *54*(1), 17–31.

Hitt, M. A., Bierman, L., Uhlenbruck, K., & Shimizu, K. (2006). The Importance of Resources in the Internationalization of Professional Service Firms: The Good, the Bad, and the Ugly. *Academy of Management Journal*, *49*(6), 1137–1157.

Hitt, M. A., Hoskisson, R. E., & Kim, H. (1997). International Diversification: Effects on Innovation and Firm Performance in Product-Diversified Firms. *Academy of Management Journal*, *40*(4), 767–798.

Hoffman, A. J. (1999). Institutional Evolution and Change: Environmentalism and the U.S. Chemical Industry. *Academy of Management Journal*, *42*(4), 351–371.

Horrocks, I. (2002). DAML+OIL: A Description Logic for the Semantic Web. *IEEE Data Engineering Bulletin*, *25*(1), 4–9.

Horrocks, I., Patel-Schneider, P. F., & Harmelen, F. (2003). From SHIQ and RDF to OWL: The Making of a Web Ontology Language. *Web Semantics Science Services and Agents on the World Wide Web*, *1*(1), 7–26.

Horrocks, I., Sattler, U., & Tobies, S. (2000). Practical Reasoning for Expressive Description Logics. *Logic Journal of the IGPL*, *8*(1705), 161–180.

Hoskisson, R. E., Hitt, M. A., Johnson, R. A., & Moesel, D. D. (1993). Construct Validity of an Objective (Entropy) Categorical Measure of Diversification Strategy. *Strategic Management Journal*, *14*(3), 215–235.

Hsu, C., & Lin, J. (2008). Acceptance of Blog Usage: The Roles of Technology Acceptance, Social Influence and Knowledge Sharing Motivation. *Information & Management*, *45*(1), 65–74.

Hsu, C. L., & Lu, H. P. (2004). Why do People Play On-line Games? An extended TAM with Social Influences and Flow Experience. *Information & Management*, *41*(7), 853–868.

Huber, G. P. (1991). Organizational Learning: The Contributing Processes and the Literatures. *Organization Science*, *2*(1), 88–115.

Hytönen, H., Jarimo, T., Salo, A., & Yli-Juuti, E. (2011). Markets for Standardized Technologies: Patent Licensing with Principle of Proportionality. *Technovation*, *32*(9-10), 523–535.

Iwasa, T., & Odagiri, H. (2004). Overseas R&D, Knowledge Sourcing, and Patenting: An Empirical Study of Japanese R&D Investment in the US. *Research Policy*, *33*(5), 807–828.

Jaccard, J., & Turrisi, R. (2003). *Interaction Effects in Multiple Regression*. Sage Publications, Inc.

Jaffe, A. B., & Palmer, K. (1997). Environmental Regulation and Innovation: A Panel Data Study. *The Review of Economics and Statistics*, *79*(4), 610–619.

Jarvenpaa, S. L., Tractinsky, N., & Vitale, M. (2000). Consumer Trust in an Internet Store. *Information Technology & Management*, *1*(1-2), 45–71.

Jennex, M. E., Smolnik, S., & Croasdell, D. (2012). Where to Look for Knowledge Management Success. *International Conference on System Sciences,* 3969–3978.

Jennings, P. D., & Zandbergen, P. A. (1995). Ecologically Sustainable Organizations: An Institutional Approach. *Academy of Management Review*, *20*(4), 1015–1052.

Johnstone, N., Haščič, I., Poirier, J., Hemar, M., & Michel, C. (2012). Environmental Policy Stringency and Technological Innovation: Evidence from Survey Data and Patent Counts. *Applied Economics*, *44*(17), 2157–2170.

Jones, K., & Leonard, L. (2008). Trust in Consumer-to-Consumer Electronic Commerce. *Information & Management*, *45*(2), 88–95.

Judge, W. Q., & Douglas, T. J. (1998). Performance Implications of Incorporating Natural Environmental Issues into the Strategic Planning Process: An Empirical Assessment. *Journal of Management Studies*, *35*(2), 241–262.

Kalyanpur, A., Parsia, B., Sirin, E., Grau, B., & Hendler, J. (2006). Swoop: A Web Ontology Editing Browser. *Web Semantics Science Services and Agents on the World Wide Web*, *4*(2), 144–153.

Kambil, A. (2008). What is Your Web 5.0 Strategy? *Journal of Business Strategy*, *29*(6), 56–58.

Kang, Y. B. and Li, Y. F. and Krishnaswamy, S. (2012). Predicting Reasoning Performance Using Ontology Metrics. *Lecture Notes in Computer Science*, (*7649)*, 198–214.

Karlinsky-Shichor, Y., & Zviran., M. (2012). Factors Influencing Perceived Benefits and User Satisfaction in Knowledge Management Systems. *Information Systems (ILAIS) Conference*.

Kassinis, G., & Vafeas, N. (2006). Stakeholder Pressures and Environmental Performance. *Academy of Management Journal*, *49*(1), 145–159.

Kazakov, Y., Krötzsch, M., & Simancík, F. (2011). Concurrent Classification of EL Ontologies. *Lecture Notes in Computer Science*, (*7031)*, 305–320.

Kazakov, Y., Krötzsch, M., & Simančík, F. (2012). ELK: A Reasoner: Architecture and Evaluation. *OWL Reasoner Evaluation Workshop*.

Kim, W. C., Hwang, P., & Burgers, W. P. (1993). Multinationals' Diversification and the Risk-return Trade-off. *Strategic Management Journal*, *14*(4), 275–286.

Kim, Y. G., Suh, J. H., & Park, S. C. (2008). Visualization of Patent Analysis for Emerging Technology. *Expert Systems with Applications*, *34*(3), 1804–1812.

King, A. A., & Lenox, M. J. (2000). Industry Self-Regulation without Sanctions: The Chemical Industry's Responsible Care Program. *Academy of Management Journal*, *43*(4), 698–716.

Kirca, A. H., Hult, G. T. M., Mena, J. A., & Miller, J. C. (2011). Firm-Specific Assets, Multinationality, and Financial Performance: A Meta-Analytic Review and Theoretical Integration. *Academy of Management Journal*, *54*(1), 47–72.

Klassen, R. D, & Whybark, D. C. (1999). The Impact of Environmental Technologies on Manufacturing Performance. *Academy of Management Journal*, *42*(6), 599–615.

Klassen, R. D., & McLaughlin, C. P. (1996). The Impact of Environmental Management on Firm Performance. *Management Science*, *42*(8), 1199–1214.

Klösgen, W., & Zytkow, J. M. (2002). *Handbook of Data Mining and Knowledge Discovery*. Oxford University Press.

Knublauch, H., Fergerson, R.W., Noy, N.F., Musen, M. (2004). The Protégé OWL Plugin: An Open Development Environment for Semantic Web Applications. *Lecture Notes in Computer Science*, (*3298)*, 229–243.

Kogut, B., & Zander, U. (1992). Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology. *Organization Science*, *3*(3), 383–397.

Kolekofski, K., & Heminger, A. (2003). Beliefs and Attitudes Affecting Intentions to Share Information in an Organizational Setting. *Information & Management*, *40*(6), 521–532.

Komiak, S. Y. X., & Benbasat, I. (2006). The Effects of Personalization and Familiarity on Trust and Adoption of Recommendation Agents. *MIS Quarterly*, *30*(4), 941–960.

Kortum, S., & Putnam, J. (1997). Assigning Patents to Industries: Tests of the Yale Technology Concordance. *Economic Systems Research*, *9*(2), 161–176.

Kostova, T. (1999). Transnational Transfer of Strategic Organizational Practices: A Contextual Perspective. *Academy of Management Review*, *24*(2), 308.

Kostova, T., Roth, K., & Dacin, M. T. (2008). Institutuonal Theory in the Study of Multinational Corporations: A critique and New Directions. *Academy of Management Review*, *33*(4), 994–1006.

Kotabe, M., Dunlap-Hinkler, D., Parente, R., & Mishra, H. A. (2007). Determinants of Cross-national Knowledge Transfer and its Effect on Firm Innovation. *Journal of International Business Studies*, *38*(2), 259–282.

Koufaris, M., & Hampton-Sosa, W. (2004). The Development of Initial Trust in an Online Company by New Customers. *Information & Management*, *41*(3), 377–397.

Kuan, H., & Bock, G. (2007). Trust Transference in Brick and Click Retailers: An Investigation of the Before-Online-Visit Phase. *Information & Management*, *44*(2), 175–187.

Lacoste, D., Sawant, K. P., & Roy, S. (2011). An Efficient XML to OWL Converter. *Proceedings of the 4th India Software Engineering Conference*, 145–154.

Lahiri, N. (2010). Geographic Distribution of R&D Activity: How does it Affect Innovation Quality? *Academy of Management Journal*, *53*(5), 1194–1209.

Lanjouw, J. O., & Mody, A. (1996). Innovation and the International Diffusion of Environmentally Responsive Technology. *Research Policy*, *25*(4), 549–571.

Larzelere, R. E., & Huston, T. L. (1980). The Dyadic Trust Scale: Toward Understanding Interpersonal Trust in Close Relationships. *Journal of Marriage & the Family*, *42*(3), 595.

Lassila, O., & McGuinness, D. (2001). The Role of Frame-Based Representation on the Semantic Web. *Linköping Electronic Articles in Computer and Information Science*, *6*(5).

Lee, D., Kauffman, R. J., & Bergen, M. E. (2009). Image Effects and Rational Inattention in Internet-Based Selling. *International Journal of Electronic Commerce*, *13*(4), 127–166.

Lee, I. (2005). The Evolution of E-Recruiting: A Content Analysis of Fortune 100 Career Web Sites. *Journal of Electronic Commerce in Organizations*, *3*(3), 57–68.

Lee, O. K., Sambamurthy, V., Lim, K., & Wei, K. K. (2008). *IT-Enabled Organizational Agility and Sustainable Competitive Advantage*. Disponible en http://ssrn.com/abstract=1249301.

Lemley, M. A., & Shapiro, C. (2005). Probabilistic Patents. *The Journal of Economic Perspectives*, *19*(2), 75–98.

Leonard, H. (1988). *Pollution and the Struggle for the World Product: Multinational Corporations, Environment, and International Comparative Advantage*. Cambridge University Press.

LexisNexis Website. (2012). Disponible en *www.lexisnexis.com/*.

Li, L., & Qian, G. (2005). Dimensions of International Diversification. *Journal of Global Marketing*, *18*(3-4), 7–35.

Liker, J. K., & Sindi, A. A. (1997). User Acceptance of Expert Systems: a Test of the Theory of Reasoned Action. *Journal of Engineering & Technology Management*, *14*(2), 147–173.

Lippman, S. A., & Rumelt, R. P. (2003). A Bargaining Perspective on Resource Advantage. *Strategic Management Journal*, *24*(11), 1069–1086.

Lorchat, J., & Noel, T. (2003). Energy Saving in IEEE 802.11 Communications using Frame Aggregation. *IEEE Global Telecommunications Conference,* 1296–1300.

Lu, H., & Lin, J. C. (2002). Predicting Customer Behavior in the Market-space: a Study of Rayport and Sviokla's Framework. *Information & Management*, *40*(1), 1–10.

Lu, J. W., & Beamish, P. W. (2004). International Diversification and Firm Performance: The S-Curve Hypothesis. *Academy of Management Journal*, *47*(4), 598–609.

Madsen, P. M. (2009). Does Corporate Investment Drive a "Race to the Bottom" in Environmental Protection? A Reexamination of the Effect of Environmental Regulation on Investment. *Academy of Management Journal*, *52*(6), 1297–1318.

Malhorta, Y. (2001). *From information management to knowledge management: beyond the "Hi-Tech Hidebound" systems*. Knowledge Management And Business Model Innovation, 115-134. IGI Publishing.

Malhotra, A., Melton, J., & Walsh, N. (2007). *XQuery 1.0 and XPath 2.0 Functions and Operators*. Working draft W3C. Disponible en http://www.w3.org/TR/xpath-functions.

Malhotra, Y. (2005). Integrating Knowledge Management Technologies in Organizational Business Processes: Getting Real Time Enterprises to Deliver Real Business Performance. *Journal of Knowledge Management*, *9*(1), 7–28.

Malhotra, Y., & Galletta, D. F. (2003). Role of Commitment and Motivation in Knowledge Management Systems Implementation: Theory, Conceptualization, and Measurement of Antecedents of Success. *Hawaii International Conference on System Science, 4,* 1–10.

Manola, F., & Miller, E. (2004). *RDF Primer*. W3C Recommendation.

Marcus, A. (2005). *Big Winners and Big Losers: The 4 Secrets of Long-term Business Success and Failure*. Wharton School Publishing.

Marcus, A. A., & Fremeth, A. R. (2009). Green Management Matters Regardless. *Academy of Management Perspectives*, *23*(3), 17–27.

Marcus, A., & Geffen, D. (1998). The Dialectics of Competency Acquisition: Pollution Prevention in Electric Generation. *Strategic Management Journal*, *19*(12), 1145–1168.

Marcus, A., Aragón-Correa, J. A., & Pinkse, J. (2011). Firms, Regulatory Uncertainty, and the Natural Environment. *California Management Review*, *54*(1), 5-16.

Margolis, J. D., & Walsh, J. P. (2003). Misery Loves Companies: Rethinking Social Initiatives by Business. *Administrative Science Quarterly*, *48*(2), 268–305.

Markman, G., Espina, M. I., & Phan, P. H. (2004). Patents as Surrogates for Inimitable and Non-Substitutable Resources. *Journal of Management*, *30*(4), 529–544.

Martin-Tapia, I., Aragon-Correa, J. A., & Senise-Barrio, M. E. (2008). Being Green and Export Intensity of SMEs: The Moderating Influence of Perceived Uncertainty. *Ecological Economics*, *68*(1-2), 56–67.

Martin, J. (2006). Multiple Intelligence Theory, Knowledge Identification and Trust. *Knowledge Management Research Practice*, *4*(3), 207–215.

Martin, N. L., & Mykytyn, P. P. (2009). Evaluating the Financial Performance of Business Method Patent Owners. *Information Systems Management*, *26*(3), 285–301.

Maurer, S. D., & Liu, Y. (2007). Developing Effective E-recruiting Websites: Insights for Managers from Marketers. *Business Horizons*, *50*(4), 305–314.

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An Integrative Model of Organizational Trust. *Academy of Management Review*, *20*(3), 709–734.

Mazzoleni, R., & Nelson, R. R. (1998). The Benefits and Costs of Strong Patent Protection: A Contribution to the Current Debate. *Research Policy*, *27*(3), 273–284.

McKnight, D. H., Choudhury, V., & Kacmar, C. (2002). Developing and Validating Trust Measures for e-Commerce: An Integrative Typology. *Information Systems Research*, *13*(3), 334–359.

McKnight, D. H., Cummings, L. L., & Chervany, N. L. (1998). Initial Trust Formation in New Organizational Relationships. *Academy of Management Review*, *23*(3), 473–490.

Melville, N. P. (2010). Information Systems Innovation for Environmental Sustainability. *MIS Quarterly*, *34*(1), 1–21.

Melville, N., Kraemer, K., & Gurbaxani, V. (2004). Review: Information Technology and Organizational Performance: An Integrative Model of IT Business Value. *MIS Quarterly*, *28*(2), 283–322.

Michalski, R. S., Bratko, I., & Bratko, A. (1998). *Machine Learning and Data Mining; Methods and Applications*. Morgan Kaufmann.

Minsky, M. (1995). A Framework for Representing Knowledge. *The Phycology Of Computer Vision*, 163–189.

Morrison, D. E., & Dunlap, R. E. (1986). Environmentalism and Elitism: a Conceptual and Empirical Analysis. *Environmental Management*, *10*(5), 581–589.

Motik, B., Grau, B. C., Horrocks, I., Wu, Z., Fokoue, A., & Lutz, C. (2009). *OWL 2 Web Ontology Language Profiles*. W3C Recommendation.

Motik, B., Patel-Schneider, P. F., & Parsia, B. (2009). *OWL 2 Web Ontology Language Primer*. W3C Recommendation.

Nachum, L., & Zaheer, S. (2005). The Persistence of Distance? The Impact of Technology on MNE Motivations for Foreign Investment. *Strategic Management Journal*, *26*(8), 747–767.

Nayyar, P. R. (1993). Performance Effects of Information Asymmetry and Economies of Scope in Diversified Service Firms. *Academy of Management journal*, *36*(1), 28–57.

Neubauer, T., & Stummer, C. (2009). Interactive Selection of Web Services under Multiple Objectives. *Information Technology and Management*, *11*(1), 25–41.

Newton, T., & Harte, G. (1997). Green Business: Technicist Kitsch? *Journal of Management Studies*, *34*(1), 75–98.

Noguera, M., Hurtado, M. V., Rodríguez, M. L., Chung, L., & Garrido, J. L. (2010). Ontology-driven Analysis of UML-based Collaborative Processes using OWL-DL and CPN. *Science of Computer Programming*, *75*(8), 726–760.

Noy, N. F., & McGuinness, D. L. (2001). *Ontology Development 101: A Guide to Creating Your First Ontology*. University of Stanford. Disponible en http://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinness.html.

Noy, N., Fergerson, R. W., & Musen, M. A. (2000). The Knowledge Model of Protégé-2000 : Combining Interoperability and Flexibility. *Knowledge Engineering and Knowledge Management Methods Models and Tools*, *1937*(1), 1–20.

NSF-US. (2012). *Science and Engineering Indicators*. Disponible en http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA278733.

Obitko, M. (2007). *Translations between Ontologies in Multi-Agent Systems*. PhD Tesis, Czech Technical University.

O'Connell, D. (2007). *System and Methods for Recycling of Cell Phones at the end of Live*. Patent number: US7251458B2.

OECD. (1994). *The Measurement of Scientific and Technological Activities Using Patent Data as Science and Technology Indicators*. Disponible en http://www.oecd-ilibrary.org/science-and-technology/the-measurement-of-scientific-and-technological-activities-using-patent-data-as-science-and-technology-indicators_9789264065574-en.

OECD. (2009). *OECD Patent Statistics Manual*. Disponible en http://www.oecd.org/science/innovationinsciencetechnologyandindustry/oecdpatentstatisticsmanual.htm.

Oltra, V., Kemp, R., & DeVries, F. (2010). Patents as a Measure of Eco-innovation. *International Journal of Environmental Technology and Management*, *13*(2), 130–148.

Ong, C. S., Lai, J. Y., & Wang, Y. S. (2004). Factors Affecting Engineers' Acceptance of Asynchronous E-learning Systems in High-tech Companies. *Information & Management*, *41*(6), 795–804.

Pakes, A., & Griliches, Z. (1984). Patents and R & D at the Firm Level : A First Look. Disponible en *http://www.nber.org/papers/w0561.pdf?new_window=1*.

Pal, S. K., Talwar, V., & Mitra, P. (2002). Web Mining in Soft Computing Framework: Relevance, State of the Art and Future Directions. *IEEE Transactions on Neural Networks*, *13*(5), 1163–77.

Panetto, H., Dassisti, M., & Tursi, A. (2012). ONTO-PMD: Product-driven ONTOlogy for Product Data Management Interoperability within Manufacturing Process Environment. *Advanced Engineering Informatics*, *26*(2), 334–348.

Park, W. G. (2010). On Patenting Costs. *The WIPO Journal*, *2*(1), 38-48.

Parry, E., & Tyson, S. (2008). An Analysis of the Use and Success of Online Recruitment Methods in the UK. *Human Resource Management Journal*, *18*(3), 257–274.

Pavlou, P. A., & Dimoka, A. (2006). The Nature and Role of Feedback Text Comments in Online Marketplaces: Implications for Trust Building, Price Premiums, and Seller Differentiation. *Information Systems Research*, *17*(4), 392–414.

Pavlou, P. A., & Fygenson, M. (2006). Understanding and Predicting Electronic Commerce Adoption: An Extension of the Theory of Planned Behavior. *MIS Quarterly*, *30*(1), 115–143.

Pelegrina, A. B., Benghazi, K., Hurtado, M. V., Noguera, M., Barjis, J., Aalst, W., Mylopoulos, J., et al. (2010). A Framework for the Semantic Representation of Business Processes within Business Organizational Models. *Enterprise and Organizational Modeling and Simulation* (63), 79–94.

Peng, W., & Lu, X. C. (2000). On the Reduction of Broadcast Redundancy in Mobile ad hoc Networks. *ACM international symposium on Mobile ad hoc networking & computing*, 129–130.

Penner-Hahn, J., & Shaver, J. M. (2005). Does International Research and Development Increase Patent Output? An Analysis of Japanese Pharmaceutical Firms. *Strategic Management Journal*, *26*(2), 121–140.

Peters, B., & Schmiele, A. (2011). *The Contribution of International R&D to Firm Profitability*. ZEW – Centre for European Economic Research. Disponible en http://ssrn.com/abstract=1740464.

Peters, M., Schneider, M., Griesshaber, T., & Hoffmann, V. H. (2012). The Impact of Technology-Push and Demand-Pull Policies on Technical Change – Does the Locus of Policies Matter? *Research Policy*, *41*(8), 1296–1308.

Petruzzelli, A. M. (2011). The Impact of Technological Relatedness, Prior Ties, and Geographical Distance on University–industry Collaborations: A Joint-patent Analysis. *Technovation*, *31*(7), 309–319.

Piatetsky-Shapiro, G., & Frawley, W. J. (1992). Knowledge Discovery in Databases: An Overview. *AI Magazine*, 13(3), 57.

Pinkse, J., & Kolk, A. (2012). Addressing the Climate Change—Sustainable Development Nexus. *Business Society*, *51*(1), 176–210.

Pinkse, J., Kuss, M. J., & Hoffmann, V. H. (2010). On the Implementation of a "Global" Environmental Strategy: The Role of Absorptive Capacity. *International Business Review*, *19*(2), 160–177.

Pisano, G. (2006). Profiting from Innovation and the Intellectual Property Revolution. *Research Policy*, *35*(8), 1122–1130.

Porter, M. E. (1990). Competitive Advantage of Nations. *Harvard Business Review 68*(2), 73-93.

Porter, M. E., & Van Der Linde, C. (1995a). Toward a New Conception of the Environment-Competitiveness Relationship. *Journal of Economic Perspectives*, *9*(4), 97–118.

Porter, M. E., & Van Der Linde, C. (1995b). Green and Competitive: Ending the Stalemate. *Harvard Business Review*, *73*(5), 120–134.

Quintas, M., Vazquez, X., Garcia, J., & Caballero, G. (2008). Geographical Amplitude in the International Generation of Technology: Present Situation and Business Determinants. *Research Policy*, *37*(8), 1371–1381.

Raghu, T. S., Woo, W., Mohan, S. B., & Rao, H. R. (2007). Market Reaction to Patent Infringement Litigations in the Information Technology Industry. *Information Systems Frontiers*, *10*(1), 61–75.

Ramus, C. A., & Steger, U. (2000). The Roles of Supervisory Support Behaviors and Environmental Policy in Employee "Ecoinitiatives" at Leading-Edge European Companies. *Academy of Management Journal*, *43*(4), 605–626.

Rappaport, A., & Flaherty, M. F. (1992). *Corporate Responses to Environmental Challenges: Initiatives by Multinational Management*. Quorum Books.

Reitzig, M. (2004). Improving Patent Valuations for Management Purposes—Validating new Indicators by Analyzing Application Rationales. *Research Policy*, *33*(6-7), 939–957.

Rocca, G. L. (2012). Knowledge Based Engineering: Between AI and CAD. Review of a Language Based Technology to Support Engineering Design. *Advanced Engineering Informatics*, *26*(2), 159–179.

Rodrigues, T., Rosa, P., & Cardoso, J. (2008). Moving from Syntactic to Semantic Organizations using JXML2OWL. *Computers in Industry*, *59*(8), 808–819.

Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so Different after All: A Cross-Discipline View of Trust. *Academy of Management Review*, *23*(3), 393–404.

Rugman, A. M. (2005). *The Regional Multinationals: MNEs and "Global" Strategic Management*. Cambridge University Press.

Rugman, A. M., & Verbeke, A. (1998a). Corporate Strategy and International Environmental Policy. *Journal of International Business Studies*, *29*(4), 819–833.

Rugman, A. M., & Verbeke, A. (1998b). Corporate Strategies and Environmental Regulations: an Organizing Framework. *Strategic Management Journal*, *19*(4), 363–375.

Rugman, A. M., & Verbeke, A. (2004). A Perspective on Regional and Global Strategies of Multinational Enterprises. *Journal of International Business Studies*, *1*(35), 3–19.

Russo, M. V., & Fouts, P. A. (1997). A Resource-Based Perspective on Corporate Environmental Performance and Profitability. *Academy of Management Journal*, *40*(3), 534–559.

Rynes, S. L. (1991). Recruitment, Job Choice, and Post-Hire Consequences: A Nall for new Research Directions. *Handbook of Industrial and Organizational Psychology, 2*, 399–444.

Sapienza, H. J., Autio, E., George, G., & Zahra, S. A. (2006). A Capabilities Perspective on the Effects of Early Internationalization on Firm Survival and Growth. *Academy of Management Review*, *31*(4), 914–933.

Scharl, A., Gebauer, J., & Bauer, C. (2001). Matching Process Requirements with Information Technology to Assess the Efficiency of Web Information Systems. *Information Technology and Management*, *2*(2), 193–210.

Scherer, F. M. (1965). Corporate Inventive Output, Profits, and Growth. *Journal of Political Economy*, *73*(3), 290.

Scherer, F. M. (1982). Inter-Industry Technology Flows and Productivity Growth. *The Review of Economics and Statistics*, *64*(4), 627.

Schmid, H. A., & Swenson, J. R. (1975). On the Semantics of the Relational Data Model. *International Conference on Management of Data - SIGMOD,* 211-223.

Schmoch, U., Laville, F., Patel, P., & Frietsch, R. (2003). Linking Technology Areas to Industrial Sectors. *Final Report to the European Commission*. Disponible en ftp://ftp.cordis.europa.eu/pub/indicators/docs/ind_report_isi_ost_spru.pdf.

Schmookler, J. (1966). *Invention and Economic Growth*. Harvard University Press.

Schoorman, F. D., Mayer, R. C., & Davis, J. H. (2007). An Integrative Model of Organizational Trust: Past, Present, and Future. *Academy of Management Review*, *32*(2), 344–354.

Schwartz, D. G., & Schreiber, Z. (2005). Semantic Information Management. *Encyclopedia of Database Technologies and Applications*, 593-600.

Seol, H., Lee, S., & Kim, C. (2011). Identifying New Business Areas using Patent Information: A DEA and Text Mining Approach. *Expert Systems with Applications*, *38*(4), 2933–2941.

Serva, M. A., & Fuller, M. A. (2004). The Effects of Trustworthiness Perceptions on the Formation of Initial Trust: Implications for MIS Student Teams. *Journal of Information Systems Education*, *15*(4), 383–395.

Sharma, S. (2000). Managerial Interpretations and Organizational Context as Predictors of Corporate Choice of Environmental Strategy. *Academy of Management Journal*, *43*(4), 681–697.

Sharma, S., & Henriques, I. (2005). Stakeholder Influences on Sustainability Practices in the Canadian Forest Products Industry. *Strategic Management Journal*, *26*(2), 159–180.

Sharma, S., & Vredenburg, H. (1998). Proactive Corporate Environmental Strategy and the Development of Competitively Valuable Organizational Capabilities. *Strategic Management Journal*, *19*(8), 729–753.

Shearer, R., Motik, B., & Horrocks, I. (2008). HermiT: A Highly-efficient OWL Reasoner. *Workshop on OWL: Experiences and Directions (OWLED)*, 26-27.

Shih, M. J., Liu, D. R., & Hsu, M. L. (2010). Discovering Competitive Intelligence by Mining Changes in Patent Trends. *Expert Systems with Applications*, *37*(4), 2882–2890.

Shih, M., & Liu, D. (2010). Patent Classification using Ontology-based Patent Network Analysis. *Proceedings Pacific Asian Conference in Information systems*.

Shrivastava, P. (1995). The Role of Corporations in Achieving Ecological Sustainability. *Academy of Management Review*, *20*(4), 936–960.

Sidorova, A., Evangelopoulos, N., Valacich, J. S., & Ramakrishnan, T. (2008). Uncovering the Intellectual Core of the Information Systems Discipline. *MIS Quarterly*, *32*(3), 467–482.

Simons, T. L. (1999). Behavioral Integrity as a Critical Ingredient for Transformational Leadership. *Journal of Organizational Change Management*, *12*(2), 89–104.

Sirin, E., Parsia, B., Grau, B., Kalyanpur, A., & Katz, Y. (2007). Pellet: A Practical OWL-DL Reasoner. *Web Semantics Science Services and Agents on the World Wide Web*, *5*(2), 51–53.

Smith, M. K., Welty, C., & McGuinness, D. L. (2004). *OWL Web Ontology Language Guide*. W3C Recommendation. The World Wide Web Consortium (W3C). Disponible en http://www.w3.org/TR/owl-guide.

Somaya, D. (2012). Patent Strategy and Management: An Integrative Review and Research Agenda. *Journal of Management*, *38*(4), 1084–1114.

Soo, V. W., Lin, S. Y., Yang, S. Y., Lin, S. N., & Cheng, S. L. (2006). A Cooperative Multi-agent Platform for Invention Based on Patent Document Analysis and Ontology. *Expert Systems with Applications*, *31*(4), 766–775.

Soylu, A., De Causmaecker, P., Preuveneers, D. B., & Desmet, P. (2011). Formal Modelling, Knowledge Representation and Reasoning for Design and Development of User-centric Pervasive Software: A Meta-review. *International Journal of Metadata, Semantics and Ontologies*, *6*(2), 96–125.

Stern, N. (2007). *The Economics of Climate Change: The Stern Review*. Cambridge University Press.

Stewart, K. J. (2003). Trust Transfer on the World Wide Web. *Organization Science*, *14*(1), 5–17.

Strassmann, P. A. (1997). *The Squandered Computer: Evaluating the Business Alignment of Information Technologies*. Information Economic Press.

Subramaniam, M., & Venkatraman, N. (2001). Determinants of Transnational new Product Development Capability: Testing the Influence of Transferring and Deploying Tacit Overseas Knowledge. *Strategic Management Journal*, *22*(4), 359–378.

Swan, J., Trawick, I., & Silva, D. (1985). How Industrial Salespeople Gain Customer Trust. *Industrial Marketing Management*, *14*(3), 203–211.

Taduri, S., Lau, G. T., Law, K. H., Yu, H., & Kesan, J. P. (2011a). An Ontology to Integrate Multiple Information Domains in the Patent System. *IEEE International Symposium on Technology and Society ISTAS*, 23–25.

Taduri, S., Lau, G. T., Law, K. H., Yu, H., & Kesan, J. P. (2011b). Developing an Ontology for the U . S . Patent System, *Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times*, 157-166.

Tagliavini, M., Ravarini, A., & Antonelli, A. (2001). An Evaluation Model for Electronic Commerce Activities within SMEs. *Information Technology and Management*, *2*(2), 211–230.

Teece, D. J. (1986). Profiting from Technological Innovation: Implications for Integration, Collaboration, Licensing and Public Policy. *Research Policy*, *15*(6), 285–305.

Thalheim, B. (2000). *Entity-relationship Modeling: Foundations of Database Technology*. Springer.

The Protégé Ontology Editor and Knowledge Acquisition System. (2012). Disponible en http://protege.stanford.edu/

Trappey, A. J. C., Hsu, F. C., Trappey, C. V., & Lin, C. I. (2006). Development of a Patent Document Classification and Search Platform using a Back-propagation Network. *Expert Systems with Applications*, *31*(4), 755–765.

Trappey, A. J. C., Trappey, C. V., Wu, C. Y., & Lin, C. W. (2012). A Patent Quality Analysis for Innovative Technology and Product Development. *Advanced Engineering Informatics*, *26*(1), 26–34.

Trappey, A., Trappey, C., & Wu, C. Y. Automatic Patent Document Summarization for Collaborative Knowledge Systems and Services. *Journal of Systems Science and Systems Engineering, 18(1),* 71–94.

USPTO. (2009). *The U.S. Commerce Department's Patent and Trademark Office (USPTO) will pilot a Program to Accelerate the Examination of Certain Green Technology Patent Applications*. Press Release USPTO. Disponible en http://www.uspto.gov/news/pr/2009/09_33.jsp

Van Heijst, G., Schreiber, A. T., & Wielinga, B. J. (1997). Using Explicit Ontologies in KBS Development. *International Journal of Human-Computer Studies*, *46*(2-3), 183–292.

Veefkind, V., Hurtado-Albir, J., Angelucci, S., Karachalios, K., & Thumm, N. (2012). A New EPO Classification Scheme for Climate Change Mitigation Technologies. *World Patent Information*, *34*(2), 106–111.

Venkatesh, V. (2000). Determinants of Perceived Ease of Use: Integrating Control, Intrinsic Motivation, and Emotion into the Technology Acceptance Model. *Information Systems Research*, *11*(4), 342–365.

Venkatesh, V., & Davis, F. D. (2000). A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. *Management Science*, *46*(2), 186–204.

Venkatraman, N. (1989). The Concept of Fit in Strategy Research: Toward Verbal and Statistical Correspondence. *Academy of Management Review*, *14*(3), 423–444.

Vernon, R. (1992). Transnational Corporations: Where are They Coming From, Where are They Headed? *Transnational Corporations*, *1*(2), 7–35.

Vries, D. (2012). *Leveraging Patents Financially. A Company Prespective*. Springer.

W3C. (1999). *XSL Transformations (XSLT)*. W3C Recommendation . Disponible en http://www.w3.org/TR/xslt.

W3C. (2004). *Web Services Architecture*. W3C Recommendation. Disponible en http://www.w3.org/TR/ws-arch/.

W3C. (2006). *Extensible Markup Language (XML) 1.1 (Second Edition)*. W3C Recommendation. Disponible en http://www.w3.org/TR/xml11/.

W3C. (2009). *OWL - Ontology Web Language*. *W3C*. W3C Recommendation. Disponible en http://www.w3.org/TR/owl-features.

Wanner, L., Brügmann, S., Diallo, B., Giereth, M., Kompatsiaris, Y., Pianta, E., Rao, G., et al. (2009). PATExpert: Semantic Processing of Patent Documentation. Disponible en http://recerca.upf.edu/patexpert.

WCED, (1987). *Our Common Future*. Oxford University Press.

Webb M. et al., (2008). *SMART 2020 : Enabling the Low Carbon Economy in the Information Age*. The Climate Change Group. Disponible en http://www.smart2020.org/_assets/files/02_Smart2020Report.pdf

Weissman, A., Petrov, M., & Gupta, S. K. (2011). A Computational Framework for Authoring and Searching Product Design Specifications. *Journal Advanced Engineering Informatics*, *25*(3), 516–534.

Westley, F., & Vredenburg, H. (1997). Interorganizational Collaboration and the Preservation of Global Biodiversity. *Organization Science*, *8*(4), 381–403.

Wi, H., Oh, S., & Jung, M. (2011). Virtual Organization for Open Innovation: Semantic Web based inter-organizational Team Formation. *Expert Systems with Applications*, *38*(7), 8466–8476.

Wielinga, B. J., & Schreiber, A. T. (1994). Reusable and shareable Knowledge Bases: A European Perspective. *Knowledge Building and Knowledge Sharing*, 110–120. Ohmsha Ltd and IOS Press.

Wiersema, M. F., & Bowen, H. P. (1997). Empirical Methods in Strategy Research: Regression Analysis and the Use of Cross-section versus Pooled Time-series, Cross-section Data. *Statistical Models for Strategic Management*, 201–219.

WIPO. (2004). *WIPO Intellectual Property Handbook: Policy, Law and Use*. World Intellectual Property Organization (WIPO). Disponible en http://www.wipo.int/about-ip/en/iprm.

WIPO. (2011). *Standard st.3 Recommended Standard on Two-letter Codes for the Representation of States, other Entities and Intergovernmental Organizations*. WIPO. Disponible en http://www.wipo.int/standards/en/part_03_standards.html

Witten, I., & Frank, E. (1999). Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations. *Special Interest Group on Management Data, 31*(1), 76-77.

Wolstencroft, K., Stevens, R., & Haarslev, V. (2007). Applying OWL Reasoning to Genomic Data. *Semantic Web,* 225–248.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

Wu, I. L., & Chen, J. L. (2005). An extension of Trust and TAM Model with TPB in the Initial Adoption of on-line Tax: An Empirical Study. *International Journal of Human-Computer Studies*, *62*(6), 784–808.

Yeoh, P. L. (2004). International Learning: Antecedents and Performance Implications Among Newly Internationalizing Companies in an Exporting Context. *International Marketing Review*, *21*(4/5), 511–535.

Zander, I. (1999). How do you Mean "Global"? An Empirical Investigation of Innovation Networks in the Multinational Corporation. *Research Policy*, *28*(2-3), 195–213.