

Semantic learning object repositories

Jesús Soto¹, Elisa García¹ and Salvador Sanchez-Alonso²

¹ Departamento de Inteligencia artificial, Universidad Pontificia de Salamanca.
{jesus.soto, elisa.garcia}@upsam.net

² Departamento de Ciencias de la Computación, Universidad de Alcalá.
{salvador.sanchez@uah.es}

Abstract: Current approaches towards enhancing reusability of learning materials pinpoint the concept of learning object as the key element of a new approach based on the creation of distributed repositories, where learning objects can be accessed, searched and retrieved by using the information in their associated metadata records. The existence of different learning object definitions asks for a new generation of flexible repositories where all the existent conceptualizations fit. In this scenario, ontological representations can play an important role as the support for sound semantic models that fulfil the new requirements. This paper introduces the main ideas about a new generation of flexible repositories, where all the definitions of learning object have their place: the semantic repository.

Keywords: learning object repository, ontology, reusability.

1. Introducción.

Un repositorio de objetos para el aprendizaje es un sistema software que almacena recursos educativos y sus metadatos (o solamente estos últimos), y que proporciona algún tipo de interfaz de búsqueda de los mismos, bien para interacción con humanos o con otros sistemas software. Los repositorios proporcionan acceso a colecciones de recursos educativos generalmente en formato electrónico, si bien la mayoría no almacenan los recursos educativos en sí, sino solamente sus metadatos. Por lo tanto, es posible encontrar el mismo recurso a través de diferentes repositorios. La funcionalidad fundamental de un repositorio de objetos de aprendizaje es la de búsqueda de recursos educativos, observándose dos grandes tipos de repositorios:

1. Interfaces de búsqueda interactivos, para uso de humanos.
2. Interfaces de consulta que puedan ser utilizados por agentes software, por ejemplo, mediante *Servicios Web*.

En ocasiones, la misma forma de búsqueda puede servir para los dos usos. No obstante, hay que tener en cuenta que la búsqueda mediante los habituales mecanismos de recuperación de información (Baeza-Yates & Ribeiro-Neto, 1999) de

propósito general (como los que usan los buscadores de Internet) deben complementarse con una búsqueda guiada por los metadatos. La forma más sencilla de este tipo de búsqueda es permitir buscar por campos de metadatos. No obstante, esas interfaces a veces tampoco resultan satisfactorias, por lo que actualmente se investiga en técnicas avanzadas que permitan hacer uso de conocimiento sobre el dominio de los metadatos, por ejemplo, mediante ontologías (como se describe más adelante).

Hay estudios recientes que intentan unificar las diferentes definiciones de objeto de aprendizaje (McGreal, 2004). Estos estudios muestran que las entidades orientadas al aprendizaje de un repositorio tienen un alto grado de variabilidad en sus caracterizaciones. La inexistencia de un vocabulario común, así como la coexistencia de diferentes definiciones de objeto de aprendizaje, señalan la necesidad de repositorios flexibles en los que quepan todas las conceptualizaciones existentes.

Para permitir comparar distintas representaciones de objetos de aprendizaje de acuerdo con la flexibilidad de su cobertura y términos asumidos de propiedades es necesario aplicar una disciplina basada, por ejemplo, en una ontología formal (Welty & Guarino, 2001). En efecto, las representaciones ontológicas desempeñan un importante papel como soporte de modelos semánticos sólidos que rellenan un número de nuevos requisitos relacionados con los procesos automáticos, tales como la búsqueda, recuperación o composición de nuevos materiales didácticos. La existencia de esquemas basados en ontologías es esencial cuando algunas de las funciones están siendo delegadas a sistemas automáticos o semiautomáticos, siguiendo la visión de la Web Semántica (Berners-Lee et al., 2001).

Este artículo presenta avances sobre una investigación en curso, cuyo objetivo fundamental es el diseño del esquema de una ontología para proporcionar mayor flexibilidad en la descripción de las entidades en un repositorio semántico de objetos de aprendizaje, además de permitir funciones automatizadas o delegación de tareas a agentes. Entre las aportaciones significativas que surgen como resultado del proceso de investigación, es posible destacar las siguientes:

- Descripción semántica de los metadatos en el estándar IEEE LOM¹.
- Descripción de recursos con distintas ontologías.
- Diseño de las funciones básicas del núcleo del repositorio que permiten una variedad de caracterizaciones ontológicas de objetos de aprendizaje.
- Potenciar las funciones de búsqueda y navegación.
- Generalización del concepto *repositorio semántico*.

Estas aportaciones son el punto de partida para otras aplicaciones que utilicen la técnica propuesta. En el resto de este artículo se señalan las carencias de los repositorios de objetos actuales y se muestran algunos de los beneficios de los repositorios semánticos en comparación con la situación existente. Finalmente, se proporcionan conclusiones y se apuntan líneas de trabajo futuro en esta área de trabajo.

¹ <http://ltsc.ieee.org/wg12/par1484-12-1.html>

2. Carencias de los repositorios actuales.

Los repositorios actuales de objetos de aprendizaje, tales como MERLOT² o CAREO³, generalmente describen los distintos recursos didácticos existentes en la Web, almacenando registros de metadatos asociados a los objetos descritos. De este modo se garantiza una búsqueda mucho más estructurada del conocimiento existente. No obstante, la búsqueda no es la única ventaja de la existencia de repositorios, pues por ejemplo éstos permiten revisiones cooperativas de los objetos de aprendizaje, de modo que la calidad del contenido es revisada y cuestionada por los distintos participantes que acceden al repositorio a través de Internet.

Al proporcionar un soporte para albergar los metadatos, estos repositorios desempeñan un importante papel de cara al futuro. No sólo los humanos pueden consultar y buscar información, sino también agentes software o sistemas LMS (*Learning Management Systems*) externos. Para procesar la información existente en los metadatos, es necesaria la presencia de metadatos de calidad, entendiéndose por ello que cumplan unos mínimos de compleción y que los datos aportados se correspondan con un esquema de metadatos preestablecido, uniforme y, a ser posible, universal. El principal problema de estos repositorios es el de carecer de un modelo conceptual que establezca qué es un objeto de aprendizaje y qué descriptores de metadatos hay asociados a cada una de las diferentes conceptualizaciones. Sin acuerdo universal sobre el modelo de metadatos a utilizar, ni certeza sobre la compleción de los mismos por parte del creador del registro de metadatos, nos encontramos ante una grave carencia que dificulta la automatización en estos repositorios. Actualmente, la calidad de los registros de metadatos depende, entre otros, de los siguientes factores:

- La información proporcionada en los metadatos depende de la bondad del creador del registro y del tiempo necesario para añadir dicha información.
- Las capacidades de edición o herramientas proporcionadas por el repositorio.
- El nivel de conocimiento del creador del registro sobre los estándares de metadatos de objetos de aprendizaje.
- El modelo conceptual del repositorio: qué entiende el creador del registro que es un objeto de aprendizaje, y qué estructura de información de metadatos debe tener.

Incluyendo información en registros de metadatos sobre el contenido de los objetos de aprendizaje se facilitan varios procesos, tales como el almacenamiento, la búsqueda y la recuperación desde repositorios distribuidos, así como la composición de nuevos materiales didácticos (entre otros). Aceptando las especificaciones de metadatos y los estándares, los objetos de aprendizaje almacenados son más interoperables y reutilizables, si bien aún existen varios inconvenientes reseñables en los registros de metadatos de los repositorios:

- La información definida por los estándares (tales como IEEE LOM) no está orientada a ser procesada por agentes externos. Este hecho, dificulta la programación de aplicaciones con capacidad de interactuar formalmente con el repositorio.
- Los estándares actuales son de propósito descriptivo. Proporcionan información sobre los contenidos o el formato del objeto de aprendizaje, pero no disponen de

² <http://www.merlot.org>

³ <http://www.careo.org/>

una semántica de ejecución para los LMS ni de un modelo formal que aporte significado dentro del contenido de los registros de metadatos.

Todas las carencias expuestas derivan a un problema de interpretación de los elementos existentes en un repositorio por parte de agentes externos. Por ejemplo, el conocimiento que poseen e intercambian los agentes software se encuentra formalizado en estructuras cognitivas que adoptan la forma de ontologías, y por lo tanto, la única forma de interpretar el conocimiento contenido en los metadatos pasa por el uso de una ontología que dé soporte a un modelo formal de conceptualización.

Otros problemas reseñables tienen que ver con la interoperabilidad y comunicación, pues al no poseer ninguna restricción en el uso de las especificaciones y ningún esquema de intercambio definido, estos repositorios quedan aislados en un sistema cerrado, cuya única ventana de acceso es la interfaz Web ofrecida desde el servidor de aplicaciones. Las funciones de intercambio de conocimiento entre distintos repositorios son imposibles de realizar, ya que no existe ningún acuerdo ni recomendación que regule este tipo de transacciones. Además, ningún agente software puede realizar un tratamiento autónomo con el mismo, como por ejemplo seleccionar, en función de las precondiciones del cliente, entre dos objetos de aprendizaje que consiguen los mismos objetivos didácticos. Otro ejemplo de las limitaciones actuales es la imposibilidad de operar con distintos grados de credibilidad o de anexar nuevos esquemas a los repositorios para procesar registros de metadatos que alberguen planificaciones por coste según los niveles de conocimiento del estudiante (Soto & García, 2005).

3. Hacia una definición flexible de objeto de aprendizaje.

Un estudio reciente de McGreal (2004) sobre la unificación de los conceptos relacionados con los recursos de aprendizaje establece que la terminología actual considera que un objeto de aprendizaje puede ser desde cualquier cosa (digital o no digital) hasta algo concreto que tenga un propósito didáctico. De acuerdo con este estudio, coexisten cinco definiciones según las cuales un objeto de aprendizaje puede ser (de más general a más específico) lo siguiente:

1. Cualquier cosa.
2. Cualquier cosa digital, con o sin propósito educativo.
3. Cualquier cosa con propósito educativo.
4. Objetos digitales que tengan formalmente un propósito educativo.
5. Objetos digitales que tengan información específica (de metadatos) para su utilización con un propósito educativo.

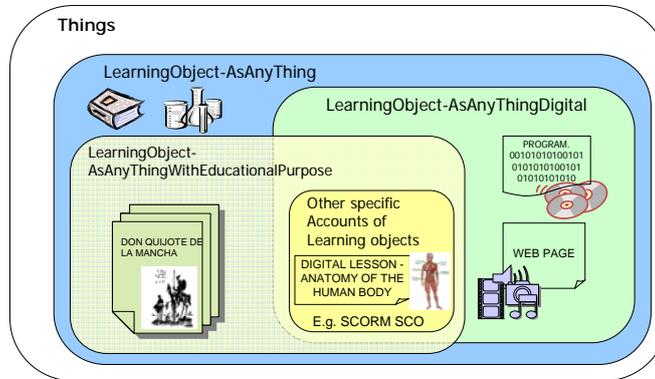


Fig. 9. Estudio de McGreal: Diferentes conceptualizaciones de un Learning Object

De acuerdo con la primera definición (en la figura 1, *LearningObject-AsAnything*), los objetos de aprendizaje son cosas que pueden ser utilizadas en eventos de aprendizaje. Por ejemplo un libro. Pero también una silla, pues puede ser utilizada como ejemplo en un entorno de enseñanza. La segunda definición (*LearningObject-AsAnythingDigital*) introduce el concepto de *objeto digital*, incluyendo el significado de representación digital utilizada en el aprendizaje. La tercera definición (*LearningObject-AsAnythingWithEducationalPurpose*) restringe los objetos de aprendizaje a aquellos que posean un propósito educativo, y cuyo diseño irá por tanto orientado a este propósito. El propósito del aprendizaje puede (y debe) describirse en un registro de metadatos, aunque de esta definición no se deriva la obligación de utilizar una especificación estandarizada como IEEE LOM. Las últimas definiciones definen objeto de aprendizaje como entidad digital con propósito didáctico o que está provista de descripciones que especifican los posibles usos didácticos que posee. La cuarta definición ya sí obliga a seguir una especificación formal en su descripción del propósito didáctico. La última de estas definiciones especifica que debe existir un esquema de metadatos que informe sobre el contenido del objeto de aprendizaje con el objetivo de facilitar su utilización en otros sistemas.

4. Ontología SLOR.

El estudio de McGreal de los conceptos existentes dentro de la definición de objeto de aprendizaje sugiere la coexistencia de todas las definiciones disponibles. Partiendo de este supuesto, se ha diseñado una ontología donde cualquier concepto enlazado a la representación de una actividad de aprendizaje es considerado un objeto de aprendizaje. Este nuevo modelo debe proporcionar una serie de funcionalidades adaptadas a cada conceptualización particular de "objeto de aprendizaje" y no necesariamente restringidas a una única definición. La flexibilidad de este nuevo esquema, nos permitirá almacenar en un repositorio cualquier tipo de información sobre un objeto de aprendizaje, normalizada o sin normalizar, y si se desea (depende de la concepción particular del creador) con la especificación del propósito educativo del mismo.

término 'aprendizaje' es representado por la clase "*oc_Learning*", que simboliza la definición de aprendizaje en OpenCyc. La representación abstracta en la cual derivan todos los términos analizados por McGreal es definida en la clase "*LearningObject-Generic*". Una instancia de esta clase será 'algo' utilizado en el aprendizaje. La relación 'utilizado' entre los dos términos se describe con la propiedad 'used-in'. La clase "*LearningObject-AsAnything*" engloba todos los posibles significados de objeto de aprendizaje. La clase "*LearningObject-AsAnythingDigital*" representa los objetos digitales utilizados en el aprendizaje "*LearningObject-AsAnything*". El término "*ComputerFileCopy*" de OpenCyc puede ser utilizado como una caracterización de cualquier objeto digital, por lo tanto, sirve como clase base de la definición de objeto de aprendizaje digital. La clase "*LearningObject-AsAnythingWithEducationalPurpose*" ha sido definida para representar los objetos que tienen descrito su propósito educativo. Estas dos últimas caracterizaciones están combinadas en las especificaciones actuales de e-learning. Cualquier individuo de la clase "*LearningObject-AsAnythingDigital*" puede tener asociado un registro de metadatos según alguna especificación formal existente, por ello, se ha creado en el modelo formal la propiedad "*hasAssociatedMetadataRecord*" cuyo rango es una clase que representa un registro abstracto de metadatos. Para dar soporte a la cuarta y quinta definición de McGreal, es necesario introducir dentro del modelo formal los esquemas de metadatos de las especificaciones estándar, tales como LOM "*LOM_Record*" o SCORM "*SCORM_SCO_Manifest*". La figura 2 muestra las relaciones entre todos los términos descritos. En esta figura, las clases de la ontología son mostradas en rectángulos con un borde sólido negro, con varios compartimientos separados por líneas horizontales. El nombre de la clase está en el compartimiento superior, el resto almacena las propiedades (aparecen precedidas por un círculo 'O'). Mientras que las restricciones aparecen precedidas por un círculo 'R'. Por ejemplo, la clase "*LearningObject-Generic*" aparece enlazada a la clase "*oc_Learning*" por una flecha que representa la propiedad objeto "*used-in*" en el correspondiente compartimiento de la clase. Esta figura también muestra la jerarquía de relaciones entre las clases de la ontología. Las clases que representan conceptos genéricos aparecen en la parte superior de la figura, en cambio las clases específicas están en la parte inferior. La relación de jerarquía es representada con una flecha acabada en triángulo. Por ejemplo, la clase "*LearningObject-LOM*" está enlazada a la clase "*LearningObject-AsAnythingDigital*" indicando la relación de jerarquía. Otras representaciones de restricciones son mostradas en la figura 2.

5. Un repositorio flexible: SLOR.

Todas las definiciones discutidas en la sección anterior pueden coexistir, derivando del análisis de tal diversidad de conceptualizaciones y de la taxonomía de las mismas, las bases de un modelo conceptual neutro. Este nuevo modelo proporciona una serie de funcionalidades adaptadas a cada conceptualización particular de objeto de aprendizaje y no necesariamente restringidas a una única definición. La flexibilidad de este nuevo esquema permite almacenar en un repositorio cualquier tipo de información (normalizada o no) sobre un objeto de aprendizaje. Si se desea (lo cual

depende de la concepción particular del creador) será posible además incluir una especificación con el propósito educativo de los objetos.

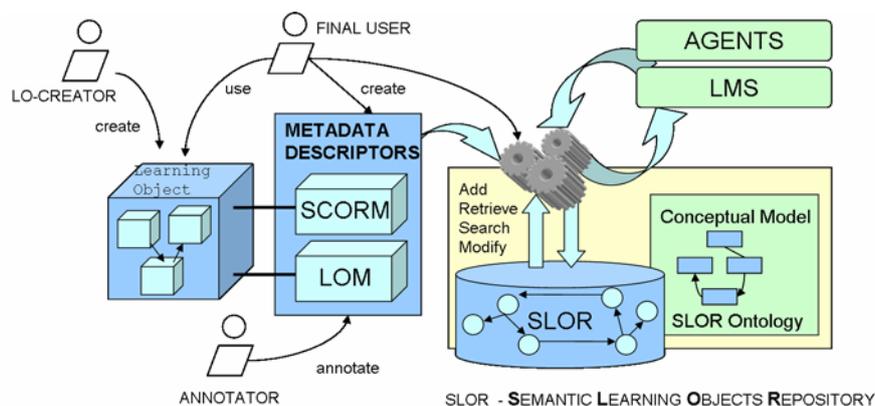


Fig. 11. Agentes externos a SLOR.

Según la figura 3, un objeto de aprendizaje es creado por un autor “LO-CREATOR”, un usuario final utiliza el objeto que puede localizarlo gracias a las funciones de búsqueda del repositorio. También puede describirlo en múltiples formatos. Por ejemplo, en “SCORM” indicando la organización de contenidos, el tipo de recursos utilizados, la secuenciación del mismo, etc. o en “LOM”. Agentes externos pueden buscar y operar con esos objetos gracias al modelo formal propuesto en esta investigación.

Este enfoque puede servir como el pilar de un nuevo modelo de repositorio de objetos de aprendizaje, donde el diseño está basado en un sólido modelo semántico que incluye todas las definiciones del análisis de McGreal para soportar los diferentes tipos de objetos de aprendizaje. Los clientes del repositorio (usuarios finales, agentes software y sistemas de administración de contenidos didácticos), podrían, entre otras funcionalidades, añadir, recuperar, modificar y buscar objetos de aprendizaje sin importar la conceptualización utilizada por el creador. Por ejemplo, un cliente software utilizando un modelo basado en LOM, podría recuperar objetos de tales repositorios, incluso aquellos objetos que estén albergados en otro tipo de sistemas con otras especificaciones de ejecución (como por ejemplo sistemas de administración de contenidos de aprendizaje basados en SCORM).

El prototipo SLOR (*Semantic Learning Object Repository*), basado en la ontología descrita, ha sido específicamente diseñado para la creación y administración de metadatos de los objetos de aprendizaje con propósitos de integración e intercambio con otros sistemas. Esta propuesta aporta mejores y nuevas funcionalidades sobre los repositorios actuales, gracias a la posibilidad que la ontología subyacente ofrece para ejecutar inferencias sobre el conocimiento albergado en los registros del repositorio. Sus funcionalidades están agrupadas en módulos según el principio de escalabilidad:

- La función de *creación* de un objeto de aprendizaje permite incluir un nuevo registro de metadatos según el modelo conceptual previamente establecido por el creador a través de la interfaz. El método de creación obtiene una referencia a un

registro de metadatos, creando una instancia de la clase correspondiente, y permite establecer las propiedades del mismo según la definición del concepto elegido en la ontología subyacente, lo que puede ocasionar variaciones entre modelos.

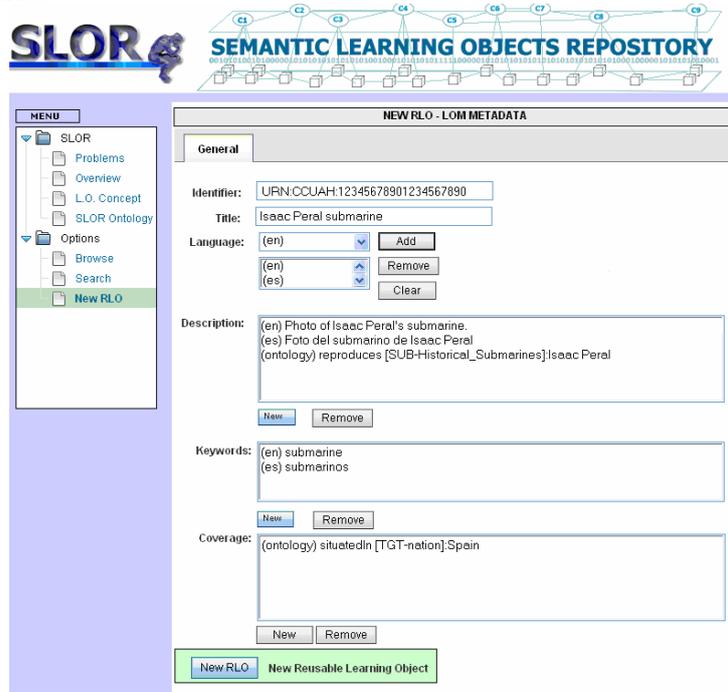


Fig. 12. Nuevo Registro de Metadatos (LOM)

La información puede ser enlazada a conceptos de otras ontologías externas al modelo base de conceptualización, como es el caso del ejemplo mostrado en la figura 5.

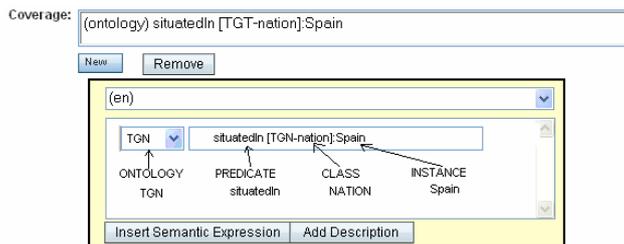


Fig. 13. Enlace a Conceptos de otras ontologías externas.

- La *búsqueda semántica* permite solicitar instancias de las distintas conceptualizaciones del modelo ontológico, por ejemplo, recuperar todos los objetos digitales, o todos los objetos con propósito educacional. Por otro lado, los registros de SLOR almacenan información enlazada a conceptos de otras

ontologías. Este escenario proporciona un profundo nivel de búsqueda que permite construir consultas complejas, como por ejemplo, recuperar todos los elementos del periodo barroco que estén situados en España. Todas estas restricciones son almacenadas en una lista antes de invocar el método de búsqueda.

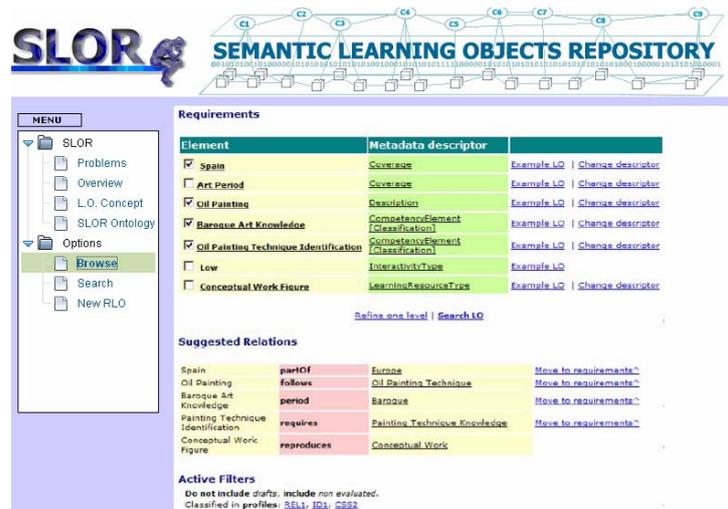


Fig. 14. Búsqueda de Objetos de Aprendizaje

6. Conclusiones.

El modelo actual de repositorios de objetos de aprendizaje cuenta con un buen número de inconvenientes que imposibilitan, o al menos dificultan, su uso generalizado. El modelo de repositorio semántico esbozado en este artículo (SLOR), prototipo de una nueva generación de repositorios de objetos de aprendizaje, proporciona la unificación de las distintas conceptualizaciones existentes de los objetos de aprendizaje. Basado en las especificaciones formales de la Web Semántica, SLOR proporciona mayores oportunidades para el intercambio y reutilización de los contenidos didácticos. Sus mayores aportaciones tienen que ver con el procesamiento autónomo de la meta-información por parte de los agentes software externos, lo cual facilitará el desarrollo de nuevas herramientas de construcción didáctica, con mayores capacidades de búsqueda, integración y combinación de objetos de aprendizaje.

Referencias.

Baeza-Yates, Ricardo; Ribeiro-Neto, Berthier (1999). *Modern Information Retrieval*. New York: Addison-Wesley.

Berners-Lee, T., Hendler, J. y Lassila, O. (2001) The Semantic Web. *Scientific American* 284, pp. 34-43.

McGreal, R. (2004). Learning Objects: A Practical definition. *International Journal of Instructional Technology and Distance Learning* 1(9).

Soto, J. & Garcia, E. (2005). Sistema multiagente inteligente para la planificación organizada del estudio de un alumno. En *actas del III Simposio Internacional de Sistemas de Información e Ingeniería del Software en la Sociedad del Conocimiento* 1(1), 34-51.

Welty, C. & Guarino, N. (2001). Supporting ontological analysis of taxonomic relationships. *Data and Knowledge Engineering* 39(1), pp. 51-74.