

Metodología de creación de contenidos docentes en un Sistema Inteligente de Tutorización Avanzada (SITA)

Carmen PAGES

Ciencias de la computación, Universidad de Alcalá
Alcalá de Henares, Madrid 28801, España

y

José J. MARTINEZ

Ciencias de la computación, Universidad de Alcalá
Alcalá de Henares, Madrid 28801, España

y

Miguel A. SICILIA

Ciencias de la computación, Universidad de Alcalá
Alcalá de Henares, Madrid 28801, España

RESUMEN

En los últimos años el aprendizaje se enfrenta a un cambio de concepción y de método, sobre todo con la entrada en escena de las nuevas tecnologías que permiten creación y divulgación de material didáctico interactivo y metodologías de trabajo fuera de las aulas tradicionales. Como conclusión a los estudios realizados de los sistemas actuales: sistemas comerciales de teleformación y sistemas de aprendizaje basados en la Inteligencia Artificial, hemos encontrado dos limitaciones fundamentales, relativas a la instrucción centrada en el alumno y a la formalización del conocimiento. En este artículo, basándonos en una arquitectura síntesis de los actuales sistemas de aprendizaje basados en internet y de los sistemas basados en conocimientos, se propone una metodología de creación de contenidos docentes soportados por dicha arquitectura. Nuestra propuesta de Sistemas Inteligente de Tutorización Avanzada (SITA) representa una visión diferente de la modelización, almacenamiento y tratamiento de los contenidos a partir del lenguaje natural, además de permitir al alumno guiar su propia instrucción. Desarrollaremos la arquitectura modular del sistema propuesto. Describiremos la metodología utilizada para la creación formalizada de contenidos docentes, basada en la Ingeniería Lingüística del Conocimiento, que incluye reglas sobre el dominio.

Palabras Claves: Teleformación, Tutorización, Contenidos Docentes, Regla Semántica, EPC.

1. ENTORNO Y ARQUITECTURA

La idea de proporcionar inteligencia a los entornos de teleformación no es nueva. Desde el principio de la década de los 90 diversos autores, principalmente Brusilovsky, [15], han desarrollado sistemas ILE (Intelligent Learning Environment) para distintas materias. El propio Brusilovsky [17], ha finales de la década, amplía el concepto para cubrir una necesidad formativa más general. Su objetivo es soportar distintos estilos de aprendizaje y adaptarlos a cada alumno. Estos sistemas han desarrollado la interactividad y adaptabilidad de los sistemas de teleformación mediante técnicas relacionadas con la interacción sistema-alumno, tales como las agrupadas bajo el término AH (Adaptative Hypermedia) [20]. El sistema que nosotros proponemos pretende, también, aunar las ventajas de los ITS y de los actuales sistemas de teleformación, pero enfatizando el desarrollo de las técnicas relacionadas con el modelo del conocimiento, es decir, nosotros proponemos que la adaptabilidad e interoperabilidad del sistema se base en una nueva forma de representación del conocimiento más que en nuevas teorías relacionadas con el modelo del alumno, ampliamente desarrolladas en la línea de investigación de Brusilovsky.

Otras investigaciones relacionadas con nuestro trabajo (dominios de conocimiento orientados a la enseñanza de tareas procedurales), tales como los prototipos TOTS, STEVE, PAT y PACO [22], van dirigidos a entornos de aprendizaje virtual, priorizando el dialogo colaborativo entre el sistema y el estudiante y las animaciones gráficas. Estos sistemas modelizan el conocimiento procedural basándose en la representación declarativa de este. El sistema por nosotros propuesto desarrolla los principios de Anderson [25] utilizando reglas de producción que el sistema ejecuta directamente. En nuestro caso las reglas

son ejecutadas a través de las decisiones del alumno. Recientemente se ha presentado un trabajo [26] que utiliza algunos de los principios básicos antes expuestos, aunque no está orientado al aprendizaje virtual ni a la enseñanza de conocimiento procedural.

El objetivo al definir la arquitectura del SITA es presentar la estructura del sistema en torno a sus componentes, incluyendo: los elementos estructurales con sus propiedades visibles y su funcionalidad y las relaciones y colaboración entre ellos, sin presentar los detalles de implementación. Una premisa inicial ha guiado este trabajo: utilizar estándares para homogeneizar problemas y soluciones en cuanto a nomenclatura, estilo, planificación y procedimientos.

SITA estará formado por un conjunto de módulos que conforman el sistema inteligente de tutorización instalados en un sistema de teleformación comercial elegido como plataforma:

- ✍ Módulo del Dominio. Representación del conocimiento de un dominio determinado. El conocimiento no existirá solamente, como en los sistemas corrientes, en forma de texto, dibujos, videos, etc., sino que relacionara todos estos contenidos de una forma inteligente utilizando modelos auténticos para la representación del conocimiento como son las reglas de producción formalizadas mediante un modelo gráfico.
- ✍ Módulo del Tutor. Representación de los conocimientos de profesores y tutores a cerca de como impartir determinados contenidos. El módulo extrae información del modelo del alumno para determinar aquellos aspectos de la materia a impartir en los que se debe insistir para cada alumno en particular, identificando sus necesidades y evolución individuales. También proporciona modelos de procesos didácticos y genera adecuadas formas de instrucción a presentar a cada alumno. Por ejemplo, se decidirá cuando revisar un concepto ó presentar uno nuevo, que concepto presentar y como hacerlo.
- ✍ Módulo del Alumno. Nivel de conocimientos del alumno. Este módulo procesa la información específica para cada estudiante. Evalúa la comprensión del alumno sobre los temas a aprender, modelizando el nivel de aprendizaje de cada uno mediante la utilización de las herramientas de diagnóstico contenidas en el modelo del tutor. Almacena la situación de aprendizaje de cada alumno en cada proceso de tutorización, de forma tal que el alumno pueda retomar un tema tutorizado donde lo dejo, empezar otra vez desde el principio, o volver a cualquier punto por el que ya pasó, recordando el sistema las actuaciones realizadas por él hasta ese momento.

Estos módulos se conectarán con:

- ✍ El sistema de teleformación. Se encarga de la gestión y administración de alumnos y cursos, alguno de los cuales puede contener información tutorizada.
- ✍ El alumno. Interaccionará con el sistema inteligente de tutorización a través de una interfaz gráfica. El sistema guiará al alumno presentándole las reglas de conocimiento

almacenadas y el alumno decidirá con sus respuestas el camino a seguir.

La arquitectura de componentes está representada en la figura 1.



Figura 1. Arquitectura modular de SITA2.

2. METODOLOGÍA DE CREACIÓN DE CONTENIDOS DOCENTES

Como hemos ya visto, el objetivo de esta metodología es obtener una base de datos formalizada que guíe al alumno de forma inteligente (tutorizada) por el conocimiento almacenado en el Módulo del Dominio del SITA. El conocimiento objeto de esta metodología es de tipo procedural orientado a la tutorización de habilidades y procesos del tipo condición/acción. Para tener una visión global de todas las fases de esta metodología se presentan de forma resumida, que será posteriormente ampliada en detalle, las actividades a realizar en cada una:

- 1) Se parte del texto que queremos tutorizar. Se realiza un análisis lingüístico del texto para traducirlo a un conjunto de reglas semánticas formalizadas.
- 2) Se analizan las reglas semánticas obtenidas convirtiéndolas en diagramas del tipo EPC (Event Driven. Process Chains). El conjunto de diagramas obtenidos forma el modelo de EPC.
- 3) Se obtienen los datos que representan el modelo de EPC: acciones, actividades, secuencia de ejecución y relaciones lógicas, que serán los que guíen el proceso de tutorización. Estos datos cargarán la base de datos de conocimiento del Módulo del Dominio. Para facilitar el laborioso trabajo que necesita esta fase, se utiliza la herramienta ARIS Toolset, que permite representar EPC y generar sus datos automáticamente.

Esquemáticamente la secuencia de las fases que intervienen en el proceso se representa en la figura 2.



Figura 2. Proceso de generación de contenidos docentes

En la descripción anterior de las fases de la metodología podemos observar la utilización de tres herramientas que describiremos brevemente:

- 1) El análisis lingüístico del lenguaje natural para la obtención de reglas semánticas. Nuestro trabajo parte de la transformación de manuales y textos educativos, que representan procesos y habilidades tipo Máquina-Herramienta, en reglas semánticas (de inferencia) que guían el proceso de enseñanza tutorizado, basándonos en la comprensión del lenguaje natural.
- 2) Los diagramas EPC para la representación gráfica de las reglas semánticas. Event-driven process chains (EPC) es una herramienta gráfica muy intuitiva que puede ser usada para describir procesos en un sentido muy general. Esta forma de representación fue introducida por [13] y desde entonces se ha difundida muy rápidamente y es ampliamente utilizada en muchos campos debido a su sencillez de lectura y diseño, su generalidad y facilidad de adaptación a distintos niveles de especificación. La metodología en la que se basa consiste en diseñar diagramas que muestran la estructura del flujo de control de un proceso como un cambio entre eventos (condiciones) y funciones (acciones). Este modelado de procesos como secuencia lógica y/o temporal de funciones nace de los conceptos subyacentes en las redes estocásticas y las redes de Petri. De hecho se vienen realizando estudios

comparativos entre la utilización de modelos de EPC o de redes de Petri para la formalización del conocimiento.

- 3) La herramienta comercial ARIS Toolset para validar el modelo EPC y obtener datos estructurados de los contenidos docentes. ARIS Toolset es un software de IDS-Scheer (www.ids-scheer.com) que soporta los conceptos de la arquitectura ARIS y comprende un conjunto de herramientas integradas. La arquitectura ARIS (Architecture of Integrated Information Systems) nace como consecuencia de un esfuerzo de estandarización de los diferentes métodos de desarrollo y modelización de procesos llevado a cabo por A.Scheer [16]. La utilización, por nuestra parte, de esta herramienta le proporciona un uso hasta ahora desconocida en el ámbito educativo.

3. OBTENCIÓN DE REGLAS SEMÁNTICAS

La primera fase de la metodología consiste en la transformación de manuales y textos educativos, que representan procesos y habilidades tipo Máquina-Herramienta, en reglas semánticas que guían el proceso de enseñanza tutorizado. Esta fase es la más laboriosa ya que es equivalente al diseño instructivo de contenidos docentes elaborado por los autores de contenidos y los diseñadores instructivos en los sistemas de e-learning tradicionales. El discurso que venga reflejado en un libro, rara vez es, ó totalmente explicativos ó completamente ejecutables. La mayoría de las veces el escritor realiza una mezcla de los dos tipos de discurso de acuerdo a sus necesidades. Además el texto puede no contener todos los posibles valores de los parámetros de las reglas (descriptores del dominio). El proceso de construcción de reglas, ya analizado por P.Sturdza [12], abarca más actividades que la mera reformulación de material existente en formato expositivo en reglas ejecutables. En el caso extremo tenemos que cuando la expresión “X es causada por Y” se reformula como una regla (“Si X entonces Y”) no se introducen nuevos elementos. El mismo número de parámetros y de valores de parámetros se mantienen en la reformulación. Sin embargo, esta no será la norma preponderante ya que, en casi todos los casos, el construir reglas incluye el enriquecimiento de la regla reformulada con nuevos elementos: nuevos valores para los descriptores, o incluso nuevos descriptores. Esto implica la obtención de más reglas que aquellas que se pueden producir a partir del texto considerado en exclusividad.

4. REPRESENTACIÓN MEDIANTE UN MODELO DE DIAGRAMAS EPC

Gráficamente un EPC estará formado por combinaciones de los siguientes elementos:

- ≠ Acciones. Una acción es un paso ejecutable, se corresponde con una actividad o una tarea que debe ser ejecutada.

- ☞ **Condiciones.** Una condición describe la situación antes y/o después de que una acción se haya ejecutado. También se pueden denominar como eventos. Las condiciones están unidas a las acciones y pueden reflejar la post-condición de una acción y/o actuar como precondition para otra acción. De esta forma una condición es la ejecución o no de una acción, es decir, una condición solo puede tener dos valores: positivo o negativo. El resultado de una acción se convierte en condición en la siguiente acción guiada por el EPC.
- ☞ **Conectores Lógicos.** Los conectores lógicos guían el flujo de control especificado en el diagrama. Se utilizan para unir condiciones. Los conectores unen una o más condiciones, ejecutan su función lógica y pasan el control a la acción o proceso de salida correspondiente en función del resultado de la función lógica. Hay tres tipos de conectores:
 - **AND.** El conector AND pasa el control a la acción o proceso siguiente si se cumplen todas y cada una de las condiciones unidas por él.
 - **OR.** El conector OR pasa el control a la acción o proceso siguiente si se cumple al menos una de las condiciones unidas por él.
 - **XOR u OR EXCLUSIVO.** El conector XOR pasa el control a la acción o proceso siguiente si se cumple una y solo una de las condiciones unidas por él.
- ☞ **Interconexiones.** Una interconexión representa el paso del flujo de control de un elemento a otro. Puede unir los siguientes tipos de elementos:
 - Condiciones con acciones para representar su post-condición;
 - Condiciones con conectores para actuar como precondition;
 - Conectores con acciones para indicar la dirección del flujo de control.

La figura 3 muestra los principales componentes que forman un diagrama EPC (condiciones, acciones y conectores) y un ejemplo sencillo de diagrama.

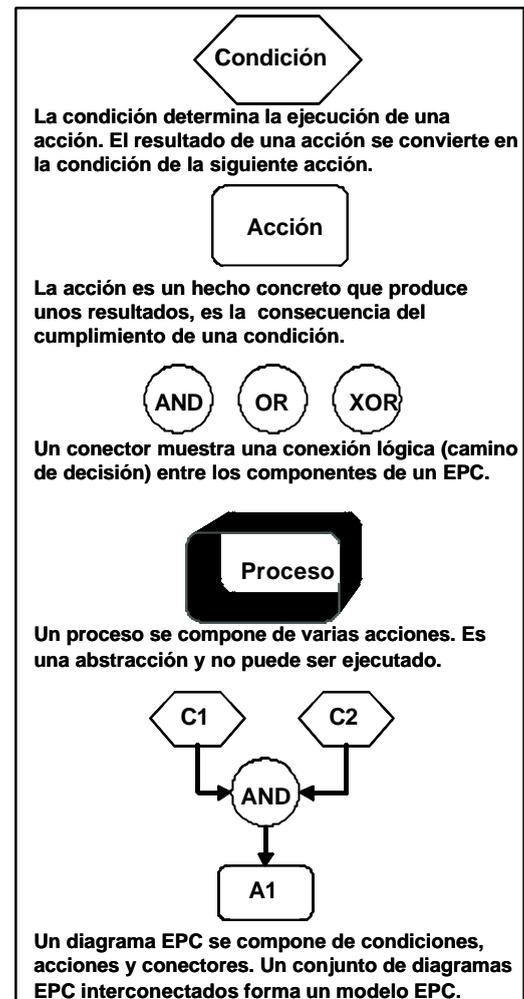


Figura 3. Componentes fundamentales del modelo de diagramas EPC

Además de las normas generales de representación de un EPC hemos añadido las siguientes indicaciones particulares para mejor comprensión del diseño y para facilitar su paso a reglas semánticas:

- 1) Simplificación del diseño, utilizando acciones complejas que agrupan acciones simples, siempre que se crea conveniente para facilitar la comprensión de la actividad y siempre que aparezcan más de diez acciones relacionadas en el mismo EPC. Estas acciones complejas las llamaremos Procesos y, de esta forma, añadimos un elemento posible más al EPC. Un proceso no es ejecutable por sí mismo y debe descomponerse en otros procesos, acciones, condiciones y conectores para poder ser ejecutado. Este conjunto de elementos forma un nuevo EPC. Todas las funciones que realizan las condiciones, los conectores lógicos y las interconexiones con acciones son igualmente aplicables a los procesos y se realizarán de la misma forma.
- 2) Utilización de un conector negativo para indicar el no cumplimiento de una condición. De esta forma las dos

salidas de la condición (positiva y negativa) pueden quedar representadas en el mismo EPC si es necesario. El control de flujo habitual en un EPC consiste en el paso a la siguiente actividad unida al conector cuando el resultado de la función lógica aplicada a las condiciones del conector sea afirmativo. Puede ocurrir que si el resultado es negativo se deba realizar una actividad alternativa y volver después a la siguiente actividad del EPC.

- 3) Utilización de interconexiones entre conectores. Tiene dos utilidades:
 - ⊗ Entre dos conectores de tipo AND, OR y XOR. Sirve para establecer varios niveles de agrupación entre precondiciones de una acción o proceso. Por ejemplo una precondición puede estar definido como (A OR B) AND C, donde A, B y C son post-condiciones de otras acciones o procesos. En este caso A y B entrarán en el conector OR y la salida del conector OR y C entrarán en el conector AND.
 - ⊗ Entre un conector negativo y cualquier otro tipo de conector. Es una ampliación debida a la presencia de conectores negativos y se usará entre un conector negativo y cualquier otro tipo de conector cuando se quiera indicar una precondición negativa para una acción, es decir, desviar el flujo de control hacia una actividad alternativa como se explico al definir los conectores negativos.
- 4) Adición de una nueva información al contenido de un EPC relativa al orden de ejecución de las actividades. Hasta ahora una acción se llevaba a cabo cuando el resultado del conector lógico que unía sus precondiciones era afirmativo (o negativo en el caso de acciones alternativas). A veces es necesario definir el orden de validación de las precondiciones antes de ejecutar una acción, es decir, las acciones cuyas post-condiciones constituyen las precondiciones de otra acción deben ser realizadas en un cierto orden. Esto es importante en la definición de procedimientos tutorizados, donde se debe guiar al alumno en el entrenamiento de realización de un trabajo o actividad. El orden o secuencia de ejecución en un EPC vendrá definido por la posición relativa (horizontal/vertical) de las actividades entre sí dentro del mismo EPC.

5. VALIDACIÓN Y GENERACIÓN DEL CONTENIDO DOCENTE FORMALIZADO

El modelo EPC muestra gráficamente el conocimiento de un sistema y constituye, tanto una descripción intuitiva, como una explicación del dominio del conocimiento. El modelo EPC es comprensible para los humanos y ejecutable mediante un sistema informático. Para que esto sea posible se debe convertir en un conjunto de reglas almacenadas que aseguren la interactividad con el estudiante. Para la implementación del

modelo EPC, su validación y su posterior conversión a reglas almacenadas se ha elegido la herramienta ARIS Toolset. Ha sido necesario un proceso de adaptación a realizar en el modelo de EPC obtenido anteriormente para cubrir todas las exigencias y criterios de ARIS.

Antes de utilizar un EPC se debe comprobar que cumple las siguientes características: está lógicamente estructurado, es consistente y tiene significado completo.

La importancia de la validación de los EPC y la definición de las condiciones que tiene que cumplir un EPC para verificar las características anteriores han sido analizadas por diversos autores. En ambos trabajos se parte de una definición formal de los EPC y de la definición de reglas de necesario cumplimiento basadas en la definición formal. La propia herramienta ARIS Toolset contiene un módulo específico para realizar el control y validación de los EPC, que se compone de un conjunto de reglas definidas y un proceso de ejecución de estas reglas contra un diagrama o modelo. También es posible modificar, eliminar o incluir reglas durante el proceso de verificación. Los puntos a tener en cuenta para la validación de la metodología de construcción de modelos EPC vista en los pasos anteriores son los siguientes:

- ⊗ Ninguna función puede estar aislada o ser innecesaria.
- ⊗ No deben crearse bucles en el flujo de control entre funciones sin salida alternativa.
- ⊗ Para cada paso del flujo de control por una función debe haber una única condición o evento que asegure su cumplimiento.
- ⊗ Todos los procesos deben ser explotados hasta el nivel de función.
- ⊗ Los operadores lógicos pueden tener varias entradas y una única salida o una única entrada y varias salidas.
- ⊗ Las funciones deben estar unidas mediante eventos, nunca directamente.
- ⊗ Debe haber una única función final (función tutorizada).
- ⊗ El flujo de procesos debe conducir a la función final.
- ⊗ No hay una única descomposición válida de un conjunto de reglas semánticas y, además, puede haber muchas representaciones incorrectas. Es muy importante la comprensión de las reglas para su representación.

Dentro de las funciones de Evaluación de ARIS Toolset utilizaremos la generación de ficheros de informes para diagramas de procesos, que nos proporcionarán automáticamente tablas de datos estructuradas que describen los objetos y relaciones que forman el modelo de EPC. Una vez obtenidas estas tablas se formalizarán para cargar la Base de Datos que formará parte del modelo de los conocimientos del dominio.

6. CONCLUSIONES

Durante el desarrollo de la arquitectura y metodología del SITA hemos obtenido una serie de conclusiones sobre el aprendizaje y la teleformación que han retroalimentado el trabajo desarrollado:

- 1) Para que se produzca un aprendizaje significativo es esencial la tutorización guiada que permita la interactividad. La actitud del alumno debe ser activa, sintiéndose responsable de su propio entrenamiento y de los avances realizados. Para ello nada mejor que la libertad del alumno para profundizar en ciertos contenidos o para elegir un camino que no se domina prefiriéndolo a otro ya conocido. Estas ventajas las proporciona la tutorización guiada propuesta en SITA.
- 2) Es fundamental para el desarrollo de la teleformación tener una metodología que permita generar contenidos de una manera eficaz y utilizando un método riguroso de trabajo basado en una descomposición del proceso de creación de contenidos en etapas de diseño, construcción y utilización, en las que se aplique un conjunto de procedimientos, técnicas y herramientas. Además de una visión meramente estructural, que es muy importante como medio de organizar y gestionar los contenidos, la metodología de SITA da un mayor grado de detalle a la hora de generarlos físicamente, sin limitarse a realizar propuestas para el análisis y diseño de los mismos.
- 3) Es preciso tener una manera de representar los flujos de aprendizaje mediante modelos de fácil y rápida creación. Gracias a la utilización de los modelos de EPC es posible obtener un buen nivel de rendimiento a la hora de capturar e implementar reglas que representen el conocimiento procedimental.

Como continuación de este trabajo de investigación y como propuesta de trabajos futuros podríamos indicar algunas posibilidades de investigación en la materia:

- 1) Herramientas que adapten los contenidos a los estándares. El hecho de que las herramientas de preparación y creación de contenidos sean innumerables y, en este caso, estamos hablando sobre todo de contenidos multimedia que acompañan de manera permanente a las lecciones tutorizadas, hace muy interesante la creación de utilidades que, centrándose en aquellos estándares que se perfilan como realmente utilizados, es decir, que lo sean "de facto", permitan la adaptación de los contenidos a dichos estándares.
- 2) Utilidades de investigación y valoración de Repositorios de Objetos y Contenidos Docentes. El propósito original de dichos Repositorios era el de mantener lugares con contenidos docentes susceptibles de ser reutilizados por parte de quien tuviera una necesidad específica. El hecho de que la gran mayoría de ellos estén accesibles a través de internet ha propiciado que algunos se transformen en amalgamas ingentes de contenidos multimedia difícilmente catalogables [3]. Un buen camino de investigación sería el del estudio de la composición y estructura de tales

Repositorios, algunos de los cuales siguen incluso ciertos estándares de teleformación.

- 3) Ampliación del marco metodológico y las técnicas de modelado a otros aspectos del proceso de aprendizaje. Una idea, que por otro lado está ya presente en la propuesta arquitectural, es la posibilidad de utilizar el formato y la estructura definido para las reglas de tutorización, que en el sistema presentado sirve exclusivamente para los contenidos propiamente dichos, en la configuración de, por ejemplo, las pruebas de test y evaluaciones [2].

7. REFERENCIAS

- [1] H.D.Jorgensen. "Formal Process Modelling". Trondheim: Norwegian University of Science and Technology, 2003.
- [2] P.Brusilovsky. "Intelligent Tutor, Environment and Manual for introductory Programming". Educational and Training Technology Internacional, 29(1), 26-34. 1992
- [3] G.Olimpo et alt. "On the concept of database of multimedia learning material". In Proceedings of the World Conference on Computers and Education". North Holland, Ámsterdam. 431-436. 1990.
- [4] P.Brusilowsky. "Adaptive Navigation Support in Educational Hypermedia: the Role of Student Knowledge Level and the Case for Meta-Adaptation". British Journal of Educational Technology. Volume 34 Issue 4 Page 487. September 2003.
- [5] J.Rickel et alt. "Using a Model of Collaborative Dialogue to Teach Procedural Tasks". Submitted to 10th Internacional Conference on Artificial Intelligence in Education, San Antonio, Texas, Mayo 2001.
- [6] E.Sacerdoti. "A Structure for Plans and Behavior". New York: Elsevier/North-Holland. 1977.
- [7] G.Keller, M.Nüttgens y A.W.Scheer. "Semantische Prozessmodellierung auf der Grundlage Ereignisgesteuerter Prozessketten (EPK)". Ver öffentlichungen des Instituts für Wirtschaftsinformatik, Heft 89, University of Saarland (ed.) Saarbrücken: 1992.
- [8] A.Scheer. "Architecture of integrated information systems systems - bases for company modeling" Berlin: 1992.
- [9] P.Sturdza, M.L.Conde y L.López. "An intelligent tutoring e-learning for training archivist". En: Actas del DLM-Forum 2002 (Barcelona, 6-8 mayo 2002).
- [10] C.Pagés, J.J.Martinez, E.Garcia y M.A.Sicilia. "On the evaluation of completeness of learning object metadata in open repositories: a case study on MERLOT". En: Actas

de la Segunda Conferencia Internacional de Multimedia y Tecnologías de la Información y la Comunicación en la Educación (Badajoz, España 3-6 de diciembre de 2003) (en prensa).

- [11] J.J.Martinez, R.Barchino y J.M.Gutierrez. "Scoring algorithms for evaluation tests in advanced e_learning systems". En: Actas del e_Society 2003 IADIS International Conference. (Lisboa Portugal. 3-6 Junio 2003). A.P.dos Reis y P.Isaias (ed.). ISBN:972-98947-0-1.