

Hacia la compartición de recursos en el ámbito de la Formación del Profesorado mediante un repositorio de metadatos federado con base semántica

Guillermo Vega-Gorgojo, Juan Ignacio Asensio-Pérez, María Jesús Rodríguez-Triana y Yannis Dimitriadis
Escuela Técnica Superior de Ingenieros de Telecomunicación, Universidad de Valladolid
Camino Viejo del Cementerio s/n, 47011 Valladolid, España
 {juaase@tel, guiveg@tel, chus@gsic, yannis@tel}.uva.es

Resumen—A pesar del esfuerzo actual de investigación en repositorios de contenidos educativos, la compartición y reutilización de recursos para la Formación de Profesorado sigue siendo escasa. Una de las razones es que las instituciones educativas suelen emplear sus propios recursos y repositorios de manera aislada, por lo que se reduce en gran medida el uso de recursos en contextos diferentes. Además, las especificaciones de metadatos educativos como LOM o Dublin Core (DC) no permiten etiquetar las características pedagógicas de los recursos, tal como demandan los formadores de profesores. Para superar estas limitaciones, el proyecto Share.TEC pretende poner en marcha un repositorio federado de metadatos sobre recursos para la Formación del Profesorado que facilite su compartición y reutilización a escala europea. En este artículo se presenta la solución técnica desarrollada en el proyecto Share.TEC de la que cabe destacar la definición de una ontología y un modelo de metadatos asociado para el ámbito de la Formación de Profesorado. El artículo también describe la arquitectura del repositorio federado resultante y el uso que se hace de estándares educativos y herramientas *software* asociadas. Se está completando un primer prototipo ya que el lanzamiento para los usuarios finales se llevará a cabo en el primer trimestre de 2010. La realimentación de los mismos será vital para comprobar si se ha podido dar respuesta a los retos del proyecto tales como la sostenibilidad o el grado de adopción de la solución propuesta.

I. INTRODUCCIÓN

La reutilización de recursos digitales educativos se ha visto altamente favorecida en los últimos años por la amplia aceptación de estándares de metadatos tales como *Dublin Core* (DC) [1] e *IEEE Learning Object Metadata* (LOM) [2]. Este hecho se une a la proliferación de repositorios de recursos de aprendizaje (como por ejemplo *iLumina* [3]) cuyas capacidades de anotación, catalogación y búsqueda de recursos educativos promueven su compartición a escala Internet.

El siguiente paso en el fomento de la reutilización y compartición de recursos educativos tiene que ver con la interoperabilidad y federación de repositorios. En ese sentido, diversos estándares como *Open Archive Initiative Protocol for Metadata Harvesting* (OAI-PMH) [4] y *Simple Publishing Interface* (SPI) [5] se están proponiendo como la infraestructura tecnológica sobre la que basar dicho objetivo de interoperabilidad y federación (véase, por ejemplo, la arquitectura propuesta en el proyecto *Ariadne* [6])

No obstante, los estándares anteriormente citados tienen un carácter general e intentan ser válidos para todo tipo de recursos educativos (de ahí probablemente su amplia aceptación). Ahora bien, ese hecho hace que su uso no permita explotar características específicas de determinadas comunidades del ámbito educativo. Un ejemplo significativo de tales comunidades es el constituido por los docentes en el ámbito de la Formación del Profesorado (TE – *Teacher Education*) cuyo soporte y fomento es un aspecto clave de la Estrategia de Lisboa [7] para la consecución de una Sociedad Europea del Conocimiento. A pesar de su relevancia, la compartición y reutilización de recursos digitales en el ámbito de la Formación del Profesorado, especialmente a una escala europea, sigue siendo un objetivo insatisfecho. A modo de ejemplo, en el ámbito español, es posible encontrar múltiples repositorios de recursos destinados a Formación del Profesorado que, sin embargo, destacan por su carácter aislado y el uso esporádico de los estándares anteriormente citados¹. Una causa probable para esa situación hay que buscarla en las limitaciones de los estándares actuales a la hora de dar soporte a los requisitos específicos que, en anotación y búsqueda de recursos, impone la comunidad de Formación de Profesorado [8]. A modo de ejemplo, no es posible que con LOM un Formador de Profesores puede anotar un recurso (por ejemplo un documento de texto detallando una unidad didáctica) indicando que es “un ejemplo para profesores de educación musical en el ámbito de la educación infantil”. Este hecho se agrava aún más, si cabe, cuando se desea compartir el mencionado recurso con Formadores de Profesores de otros países en los que, incluso, los niveles educativos (educación infantil) y las disciplinas de Formación de Profesorado (educación musical) puede no coincidir.

A este tipo de limitaciones se enfrenta el proyecto Europeo Share.TEC [9], el cual tiene como principal objetivo diseñar y poner en marcha un repositorio federado de metadatos sobre recursos para Formación de Profesorado que favorezca su compartición y reutilización a escala europea. El repositorio de Share.TEC trata de agregar metadatos de recursos almacena-

¹Véase, por ejemplo, http://www.rute.edu.es/index.php?option=com_weblinks&catid=13&Itemid=23 o <http://www.cfievalladolid2.org/webquest/common/index.php>

dos en repositorios existentes de terceros, permitir su anotación con metadatos adicionales relevantes para la comunidad de Formación de Profesorado, y proporcionar soporte a funcionalidades avanzadas de búsqueda e intermediación que tengan en cuenta las particularidades del dominio. Para conseguir el mencionado objetivo, el proyecto Share.TEC adopta una aproximación incremental que intenta reutilizar los estándares existentes, y sus desarrollos asociados, de anotación de recursos educativos (LOM y DC) y federación de repositorios (OAI-PMH y SPI). Dichos estándares son complementados, para satisfacer los requisitos específicos del dominio, mediante el uso de las tecnologías de la denominada *Web Semántica* [10].

Más concretamente, este artículo presenta la aproximación seguida por el proyecto Share.TEC consistente en la creación (y soporte a la posterior evolución) de una ontología específica del dominio de la Formación del Profesorado. Dicha ontología [11], entendida como una conceptualización compartida del ámbito de la Formación del Profesorado, es el punto de partida para la propuesta de un modelo de metadatos que extiende el estándar LOM y que permite a los Formadores de Profesorado anotar los recursos disponibles en repositorios existentes con información específica de dicho ámbito. El hecho de basar el nuevo modelo de metadatos en un estándar aceptado como es LOM, facilita que al repositorio de Share.TEC se puedan unir nuevos repositorios basados en ese formato (o en otros como DC para los que es posible llevar a cabo traducciones) facilitando de esa manera la perdurabilidad de la propuesta.

Este artículo de estructura de la siguiente manera: la sección II describe el contexto en el que se enmarca el proyecto Share.TEC indicando los problemas que se tratan de solventar, así como las funcionalidades principales que debe ofrecer el sistema propuesto. Sigue la sección III en la que se describe la ontología TEO que describe el dominio de la Formación del Profesorado, así como el modelo de metadatos propuesto en el proyecto para el etiquetado de recursos. En la sección IV se describe la arquitectura del sistema Share.TEC. Finalmente, en la sección V se presentan las principales conclusiones, así como las líneas futuras.

II. EL CONTEXTO DEL PROYECTO SHARE.TEC

Actualmente, gran parte de las instituciones y organismos dedicados a la Formación del Profesorado cuentan con sus propios repositorios de recursos de aprendizaje. La adopción de estándares de metadatos como LOM y DC para el etiquetado recursos ha permitido facilitar la reutilización de los mismos. Sin embargo, estos repositorios siguen estando aislados entre sí por lo que se reducen considerablemente las posibilidades de compartición. Más aún, los formadores de profesorado no pueden emplear las abstracciones propias de su dominio para la búsqueda de recursos al no estar contempladas en los mencionados formatos de metadatos. Estas limitaciones intentan responderse a escala europea en el proyecto Share.TEC para la comunidad de Formación del Profesorado.



Figura 1. Visión general del sistema Share.TEC

Debido a las inversiones ya realizadas en repositorios institucionales, un aspecto clave para la sostenibilidad de la propuesta es evitar que los participantes tengan que adoptar nuevos repositorios que sustituyan a los actuales. Además, los derechos de protección de los recursos no deberían comprometerse por formar parte del consorcio. Por estas razones, la solución que se ha propuesto en Share.TEC es la de un repositorio federado de metadatos de recursos de Formación del Profesorado, como se muestra en la figura 1.

Así, las instituciones educativas y editoriales participantes en el proyecto ponen a disposición del consorcio sus repositorios institucionales, permitiendo el acceso a sus recursos para la Formación del Profesorado. Los metadatos de dichos recursos se almacenan en el repositorio central de Share.TEC con el propósito de ofrecer una vista unificada y coherente de los mismos. Para ello, es necesario recoger de cada repositorio institucional los metadatos de sus recursos (en formato DC o LOM) y realizar las conversiones necesarias al formato de metadatos común desarrollado en el proyecto (descrito en la sección III).

Además de la migración de metadatos al repositorio central, el sistema Share.TEC incluye una capa semántica basada en una ontología que describe el campo de Formación del Profesorado. Los formadores de profesorado pueden emplear los vocabularios definidos en dicha ontología para etiquetar los recursos empleando abstracciones propias del dominio. De esta manera, los propios usuarios enriquecen las descripciones de los recursos ofrecidos mediante sus propias anotaciones referidas al contexto de uso de los mismos.

Por último, una de las principales funcionalidades ofrecidas por el sistema es la búsqueda de recursos. Los formadores de profesorado pueden utilizar en la búsqueda conceptos propios de su ámbito, empleando para ello los vocabularios

definidos en la ontología desarrollada. Con esta aproximación se pretende ajustar más las consultas a las necesidades e intereses de los formadores, obteniendo así resultados más relevantes que empleando únicamente los campos definidos en LOM o DC.

III. ONTOLOGÍA Y MODELO DE METADATOS PARA FORMACIÓN DEL PROFESORADO

Gran parte del esfuerzo inicial del proyecto Share.TEC se ha dedicado al desarrollo de la ontología de Formación del Profesorado denominada TEO (*Teacher Education Ontology*) [11]. El propósito de esta ontología es el de modelar las abstracciones más relevantes del dominio de Formación de Profesorado para facilitar la catalogación y compartición de recursos. De esta manera, se pretende que TEO proporcione un vocabulario común con el que describir convenientemente recursos para la Formación de Profesorado.

Es de destacar que en el desarrollo de esta ontología se ha intentado limitar la complejidad de la misma, con el objetivo de facilitar su adopción por los usuarios finales; por esta razón se ha recogido la realimentación de los propios educadores desde las etapas iniciales de concepción de TEO. Debido a este pragmatismo en el enfoque el modelo conceptual es relativamente sencillo, centrado en el recurso digital (comúnmente llamado objeto de aprendizaje en la literatura) al ser la unidad de compartición del repositorio de Share.TEC. De esta manera, un recurso en TEO puede ser caracterizado mediante una serie de elementos pedagógicos como la estrategia didáctica con la que se diseñó. También puede indicarse a qué área de conocimiento se asocia un recurso, empleando como base la lista de tópicos definida en EUROSTAT [12]. Además, pueden indicarse las competencias que intentan fomentarse empleando el modelo [13], así como el contexto en el que se ha empleado un recurso. Cabe destacar que los aspectos mencionados fueron considerados especialmente relevantes por los educadores para la compartición de recursos de Formación de Profesorado, a pesar de que no pueden describirse ni con DC ni con LOM.

Tras estabilizar una primera versión de TEO, a partir de esta ontología se ha desarrollado el Modelo de Metadatos Común (CMM - *Common Metadata Model*) para el etiquetado de recursos de Formación de Profesorado en el proyecto Share.TEC. Una decisión de diseño clave en la propuesta de CMM ha sido la de adoptar LOM como base con la intención de facilitar la transición de repositorios existentes basados en LOM hacia el repositorio federado de Share.TEC que utiliza CMM como formato común de metadatos. Así, los elementos definidos en CMM están organizados en 10 categorías, cuyas 8 primeras se corresponden directamente con las categorías 1 a 8 de LOM. Las dos categorías restantes son las que incluyen los elementos específicos de Formación de Profesorado, utilizando para ello los vocabularios definidos en TEO.

Adicionalmente, se persigue en el consorcio de Share.TEC el tratar las diferencias lingüísticas y culturales entre usuarios. Debido a esto, a partir de CMM se ha derivado una serie de Modelos de Metadatos Multiculturales correspondientes a los

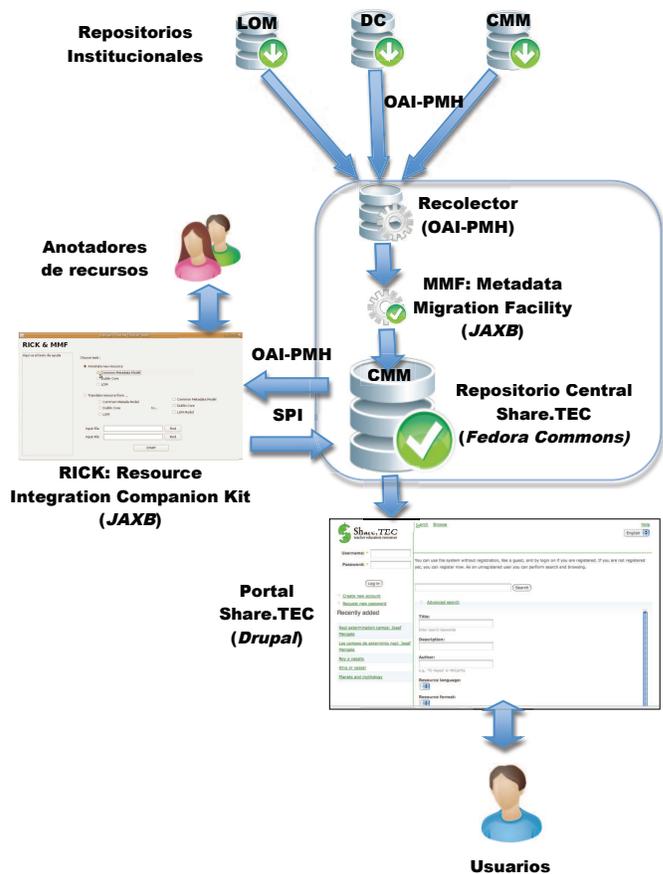


Figura 2. Arquitectura lógica del sistema Share.TEC

diferentes países representados en el proyecto. Así, los modelos MMM desarrollados pueden considerarse una colección de adaptaciones/traducciones del modelo CMM a diferentes culturas y lenguajes. Más en concreto, cada modelo MMM adquiere una dimensión multicultural mediante la traducción al lenguaje local de los elementos de las categorías 1 a 8 de CMM y con la adaptación de los vocabularios de Formación de Profesorado definidos en TEO al contexto propio (ej. disciplina, nivel educativo, institución...).

IV. EL SISTEMA SHARE.TEC

La figura 2 ilustra la arquitectura del sistema Share.TEC, así como las tecnologías involucradas en el desarrollo de cada una de sus partes. En primer lugar, es importante destacar que para que un repositorio institucional pueda formar parte del repositorio federado de Share.TEC, es necesario que cumpla con una serie de requisitos técnicos. Dichos requisitos intentan alcanzar un compromiso entre la sostenibilidad del sistema (repositorios institucionales actuales y futuros podrán cumplir los requisitos y, por lo tanto, unirse a Share.TEC) y la viabilidad técnica del mismo (el sistema no tiene que dar soporte a las particularidades de cada uno de los potenciales repositorios institucionales que deseen federarse). Por todo ello, se ha decidido imponer a los repositorios institucionales el soporte al estándar OAI-PMH (*Open Archive Initiative Protocol for*

Metadata Harvesting) [4] para poder extraer los metadatos sobre sus recursos que, obligatoriamente, deberán seguir los formatos estándar LOM, DC (ver sección II) o, incluso, el propio CMM.

OAI-PMH es un protocolo entre aplicaciones clientes (denominados “recolectores” o, en inglés, *harvesters*) y aplicaciones servidoras (repositorios) que permite el intercambio de registros de metadatos, con formatos diversos, empaquetados en documentos XML y transportados por mensajes del protocolo HTTP. Usando OAI-PMH, el “recolector” de Share.TEC recoge los metadatos de los distintos repositorios institucionales y los entrega al denominado “Sistema de Migración de Metadatos” (en inglés *Metadata Migration Facility* o MMF). El MMF se encarga de traducir sintácticamente los registros recolectados al formato común de metadatos CMM (ver sección III). Para llevar a cabo las traducciones, se ha recurrido al soporte de la biblioteca JAXB (*Java Architecture for XML Binding* o Arquitectura Java para Enlazado XML) [14] que proporciona una flexibilidad mayor a la hora de extender o readaptar las traducciones que otra tecnología probablemente más comúnmente utilizada como es XSLT (*eXtensible Stylesheet Language Transformations* o Transformaciones de Lenguajes de Hojas de Estilo Extensibles) [15]. Esa mayor flexibilidad se necesita para poder alinear más fácilmente las traducciones al formato CMM con los vocabularios definidos en la ontología TEO (ver sección III).

Los registros de metadatos en el formato CMM se almacenan en el repositorio central de Share.TEC. Los registros CMM traducidos desde los formatos LOM o DC no dispondrán de la información específica del dominio de Formación de Profesorado en los campos dispuestos al efecto (ver sección III). Para permitir el enriquecimiento de esos registros de metadatos por parte de formadores de profesorado, el sistema Share.TEC incluye la denominada herramienta RICK (*Resource Integration Companion Kit* o Sistema Complementario de Integración de Recursos). RICK es una herramienta interactiva cuyo caso de uso principal se puede resumir en los siguientes puntos:

1. A petición de su usuario, recoge un registro de metadatos CMM incompleto del repositorio central de Share.TEC.
2. Muestra una interfaz basada en formularios mediante la que se solicita al usuario que complete los campos del formato CMM específicamente relacionados con la Formación del Profesorado. En los casos necesarios, la herramienta RICK restringe los posibles valores de dichos campos a los definidos por los vocabularios derivados de la ontología TEO.
3. Una vez que el usuario está satisfecho con los metadatos introducidos, la herramienta RICK devuelve el registro CMM completo al repositorio central de Share.TEC. En este caso, se recurre al soporte del protocolo SPI (*Simple Publishing Interface*) [5].

Gracias al soporte de la herramienta RICK se espera que, de manera paulatina, los proveedores de recursos para Formación de Profesorado puedan ir enriqueciendo la descripción de sus

recursos con información específica del dominio (formalizada y compartida gracias a la ontología TEO) y, de esa manera, favorecer su compartición y, por ende, su reutilización.

Para finalizar la descripción del sistema Share.TEC, cabe mencionar que, en su versión actual, la implementación del repositorio central hace uso del *software* de repositorio “*Fedora Commons*”². Por su parte, el usuario final del sistema Share.TEC (comúnmente formadores de profesorado buscando recursos potencialmente reutilizables según sus necesidades) interactuarán con él a través del denominado “Portal Share.TEC”. Dicho portal, que integra las funcionalidad de búsqueda de navegación a través de las descripciones de recursos disponibles, se basa en el sistema de gestión de contenidos *Drupal*³.

V. CONCLUSIONES

La compartición y reutilización de recursos de aprendizaje es un objetivo largamente deseado, especialmente en el ámbito de la Formación del Profesorado. La federación de los repositorios existentes se prevé que sea un paso de gran importancia hacia la consecución de este objetivo. Con este propósito, en el proyecto Share.TEC se propone un repositorio federado de metadatos a escala europea de recursos para la Formación del Profesorado. De esta manera se pretende ofrecer una vista unificada y homogénea de los recursos ofrecidos por los socios en el consorcio, actualmente compartimentados en “silos” correspondientes a los respectivos repositorios institucionales.

El sistema Share.TEC incluye una capa semántica basada en la ontología TEO del dominio de la Formación del Profesorado. A partir de esta ontología se ha derivado el formato CMM al que se adhiere el repositorio central del sistema Share.TEC y que permite etiquetar recursos con información del campo de Formación del Profesorado (por ejemplo empleando la herramienta RICK). Las diferencias lingüísticas y multiculturales propias de los diferentes contextos de uso también han sido consideradas, desarrollando múltiples adaptaciones del modelo común CMM.

Actualmente, se está completado el primer prototipo del sistema Share.TEC y se está llevando a cabo el proceso de migración de los metadatos de los repositorios institucionales al central. En paralelo, las interfaces de usuario están siendo evaluadas en múltiples casos de estudio para mejorar la usabilidad de las mismas. En el primer trimestre de 2010 se realizará el lanzamiento del sistema, para que los usuarios finales puedan emplearlo en escenarios de uso reales. El seguimiento y la realimentación de los mismos serán vitales para evaluar el grado de éxito del sistema Share.TEC. Entre los múltiples retos planteados por el proyecto, se pueden destacar los referentes a la sostenibilidad del sistema debido a la evolución de estándares o la volatilidad de los repositorios existentes, la introducción de técnicas propias de Web 2.0 para la anotación basada en experiencia, o los problemas relacionados con la implicación de los usuarios en la definición y evaluación del sistema.

²<http://www.fedora-commons.org/>

³<http://drupal.org>

AGRADECIMIENTOS

El trabajo de investigación presentado en este artículo está financiado por el proyecto de la Unión Europea (programa *eContentplus*) Share.TEC (ECP 2007 EDU 427015). Los autores desean dar las gracias a los socios del proyecto Share.TEC y a los miembros del grupo multidisciplinar GSIC/EMIC⁴ de la Universidad de Valladolid.

REFERENCIAS

- [1] DCMI Usage Board, "DCMI Dublin Core Metadata Element Set, Version 1.1.1," DCMI, Recommendation, jan 2008, URL: <http://dublincore.org/documents/dces/>, visitado por última vez en Octubre de 2009.
- [2] IEEE Learning Technology Standards Committee, "IEEE standard for learning object metadata," Computer Society/Learning Technology Standards Committee, Specification 1484.12.1-2002, nov 2002.
- [3] B. P. Heath, D. J. McArthur, M. K. McClelland, and R. J. Vetter, "Metadata lessons from the iLumina digital library," *Communications of the ACM*, vol. 48, no. 7, pp. 68–74, 2005.
- [4] C. Lagoze, H. Van de Sompel, M. Nelson, and S. Warner, "Open Archives Initiative - Protocol for Metadata Harvesting," 2008, <http://www.openarchives.org/OAI/openarchivesprotocol.html>, visitado por última vez en Octubre de 2009.
- [5] S. Ternier, "Simple Publishing Interface Specification," 2008, http://ariadne.cs.kuleuven.be/lomi/images/b/ba/CEN_SPI_interim_report.pdf, visitado por última vez en Octubre de 2009.
- [6] ARIADNE Foundation, "ARIADNE Website - Foundation for the Knowledge Pool," 2006, <http://www.ariadne-eu.org/>, visitado por última vez en Octubre de 2009.
- [7] W. Kok, "Facing the challenge: The Lisbon strategy for growth and employment," High Level Group, Tech. Rep., November 2004, URL: http://ec.europa.eu/growthandjobs/pdf/kok_report_en.pdf, visitado por última vez en Octubre de 2009.
- [8] S. Alvino, P. Forcheri, M. G. Ierardi, and L. Sarti, "A general and flexible model for the pedagogical description of learning objects," in *Proceedings of the 20th World Computer Congress (WCC 2008)*, Milan, Italy, 2008, pp. 55–62.
- [9] Share.TEC Consortium, "Share.TEC Project Website - SHaring digital RESources in the Teaching Education Community (ECP 2007 EDU 427015)," 2008, <http://www.share-tec.eu/>, visitado por última vez en Octubre de 2009.
- [10] T. Berners-Lee, J. Hendler, and O. Lassila, "The semantic web," *Scientific American*, vol. 284, no. 5, pp. 34–43, 2001.
- [11] S. Alvino, S. Bocconi, P. Boytchev, J. Earp, and L. Sarti, "An ontology-based approach for sharing digital resources in Teacher Education," in *Proceedings of the Seventh International Workshop on Ontologies and Semantic Web for E-Learning (SWEL)*, Brighton, UK, 2009.
- [12] R. Andersson and A. Olsson, "Fields of education and training. telearc," 1999.
- [13] G. Paquette, "An ontology and a software framework for competency modeling and management," *Journal of Educational Technology and Society*, vol. 10, no. 3, p. 1, 2007.
- [14] GlassFish Community, "Java Architecture for XML Binding (JAXB)," 2008, <https://jaxb.dev.java.net/>, visitado por última vez en Octubre de 2009.
- [15] World Wide Web Consortium (W3C), "XSL Transformations (XSLT)," W3C, Recommendation, Nov. 1999, uRL: <http://www.w3.org/TR/xslt>, visitado por última vez en Octubre de 2009.

⁴<http://gsic.tel.uva.es>