

ILSA (*Ingeniería de Lenguajes Software y Aplicaciones*), UCM: Investigación en Repositorios de Objetos Educativos en Dominios Especializados

Iván Arias Rodríguez, Joaquín Gayoso Cabada, Ana M^a Fernández-Pampillón Cesteros, Manuel Márquez Cruz, Elena del Olmo Suárez, Antonio Sarasa Cabezuelo, José Luis Sierra Rodríguez
Facultades de Informática y de Filología
Universidad Complutense de Madrid
28040 Madrid, España
{iarias01, jgayoso, apampi, manmarqu, elenadelolmo, asarasa, jlsierra}@ucm.es

Resumen: En este artículo se presentan algunas de las líneas de investigación llevadas a cabo actualmente en el grupo de investigación ILSA (Ingeniería de Lenguajes Software y Aplicaciones) en torno a la producción de repositorios de objetos educativos en dominios especializados. Más concretamente, después de presentar al grupo, el artículo resume las siguientes líneas de investigación: (i) desarrollo de plataformas de producción de repositorios reconfigurables; (ii) desarrollo de herramientas de soporte semi-automatizado al proceso de producción; (iii) formulación de mecanismos avanzados para la navegación por los repositorios; y (iv) transformación de repositorios dirigida por gramáticas formales.

Palabras clave: Grupo de Investigación, Repositorios de Objetos Educativos, Desarrollo de Software dirigido por Lenguajes

Abstract: This paper describes some of the research lines concerning the production of learning object repositories in specialized domains that are being carried out in the ILSA Research Group (Implementation of Language-Driven Software and Applications). After introducing the research group, the paper summarizes the following research lines: (i) development of platforms for the production of reconfigurable repositories; (ii) development of tools for the semi-automatic support of the production process; (iii) formulation of advanced navigation mechanisms; and (iv) grammar-driven transformation of learning object repositories.

Key words: Research Group, Learning Object Repositories, Language-Driven Software Development

1. Introducción

El grupo ILSA (Grupo de Investigación en Ingeniería de Lenguajes Software y Aplicaciones) es un grupo de investigación oficial de la Universidad Complutense de Madrid (UCM), creado en el año 2010 y reconocido oficialmente por la Universidad en el año 2011 (grupo no. 962022). En (Sierra y Sarasa, 2012) puede encontrarse una descripción del grupo en el año 2012.

Actualmente, el grupo está formado por investigadores de tres facultades de la UCM:

- Joaquín Gayoso, Antonio Pareja, Antonio Sarasa, José Luis Sierra (coordinador del Grupo), en la Facultad de Informática.
- Iván Arias, Ángel Encinas, Elena del Olmo, Manuel Márquez y Ana M^a Fernández-Pampillón en la Facultad de Filología.
- José María Ruiz en la Facultad de Educación.

Asimismo, cuenta con Alfredo Fernández-Valmayor como miembro honorario.

Los orígenes de ILSA se encuentran en los trabajos sobre el desarrollo de sistemas de gestión de objetos

educativos promovidos por el Profesor Alfredo Fernández-Valmayor durante la década pasada a través de varios proyectos del Plan Nacional de I+D+i: TIC2000-0737-C03-01 “Sistema de Información Multimedia Basado en Lenguajes de Marcado y Agentes (SIMBA)”, TIC2002-04067-C03-02 “Recursos Educativos e Informativos basados en componentes distribuidos: Metodología, Lenguajes y Herramientas (REI-MLH)”, y TIN2005-08788-C04-01 “Objetos de Aprendizaje en el Campus Virtual (OdA-Virtual)” (Navarro, Fernández-Valmayor, Fernández-Manjón y Sierra, 2004; Navarro, Sierra, Fernández-Valmayor y Hernanz, 2005; Sierra, Fernández-Valmayor, Guinea y Hernanz, 2006). Estos proyectos dieron lugar a la formulación de una arquitectura para la autoría distribuida de repositorios educativos, así como para la navegación por dichos repositorios, y a la implementación de una plataforma genérica de soporte a dicha arquitectura, denominada OdA 1.0. Esta plataforma permitía la creación colaborativa de repositorios de objetos educativos, soportando no únicamente la creación de los objetos en sí, sino (y más importante) la creación colaborativa de los esquemas de catalogación de dichos objetos.

Tras la jubilación del Profesor Fernández-Valmayor, su equipo en la UCM decidió formar el grupo para continuar con estas líneas de investigación, líneas que, durante la década actual, han sido continuadas a través de los siguientes proyectos del Plan Nacional:

- Proyecto TIN2010-21288-C02-01 “Un Enfoque Generativo para la Construcción de Herramientas de Producción y Despliegue de Objetos Educativos en el Campus Virtual (GENHOE-VIRTUAL)”, en el que se hizo evolucionar OdA 1.0 a la plataforma OdA 2.0 (Fernández-Valmayor et al., 2013)
- Proyecto TIN2014-52010-R “Repositorios Educativos Dinámicamente Reconfigurables en Humanidades (RedR+Human)”, en el que se evolucionó la arquitectura de OdA 2.0 hacia una arquitectura de repositorio educativo que: (i) permite manejar adecuadamente la reconfiguración dinámica de los esquemas de catalogación, contemplando la creación inductiva de esquemas cuya estructura cambia conforme se añaden nuevos objetos al repositorio y se modifican o eliminan los ya existentes, (ii) facilita la integración de fuentes externas de información, y (iii) facilita la

integración de herramientas de autoría externas. Esta arquitectura ha cristalizado en una nueva plataforma de creación y gestión de repositorios educativos, denominada *Clavy* (Gayoso, Rodríguez y Sierra, 2017)

- Proyecto TIN2017-88092-R “Creación, Exploración y Transformación de Repositorios de Objetos Educativos en Dominios Especializados (CetrO+Spec)”, actualmente en período de ejecución, en el que se está trabajando en: (i) la formulación de enfoques para la creación semi-automática de repositorios que permitan reducir el cuello de botella que supone la catalogación manual de los contenidos, (ii) la formulación de mecanismos avanzados de navegación, y (iii) la formulación de un enfoque ETL (Extract-Transform-Load) declarativo para potenciar la interoperabilidad con otros sistemas (Buendía, Gayoso y Sierra, 2019)

Asimismo, los resultados obtenidos se han aplicado también a:

- El proyecto H2015/HUM-3426 “Edición Literaria Electrónica (eLITE-CM)” financiado por la Comunidad de Madrid. A lo largo del proyecto se han desarrollado dos herramientas que implementan algunas técnicas de ficción interactiva tales como la navegación por capas, la personalización, la bifurcación de caminos de lectura o el planteamiento de preguntas o retos (Sarasa et al., 2018; Temprado et al., 2019).
- El proyecto “Modelo Unificado de Gestión de Colecciones Digitales con Estructuras Reconfigurables: Aplicación a la Creación de Bibliotecas Digitales Especializadas para Investigación y Docencia (MUGECODER)”, financiado por la Fundación BBVA en el programa 2014 de ayudas a equipos de investigación científica en el área de las Humanidades Digitales. En este proyecto se aplicó la plataforma *Clavy* a la consolidación de *Mnemosine*, una biblioteca digital sobre textos raros y olvidados de la Edad de Plata española, período literario comprendido entre 1868 y 1939 (Romero y Bueren, 2018)
- El proyecto “Collaborative Annotation of Digitalized Literary Texts”, financiado por Google en las ediciones 2010 y 2011 del “Digital Humanities Award Program”. En este

proyecto se desarrolló una herramienta educativa para la anotación colaborativa de textos literarios, @note (Cigarrán et al., 2014; Gayoso et al., 2019)

- La acción COST IS0704 “An Interoperable Supranational Infrastructure for Digital Editions (Interedition)” de la Unión Europea, en la que se contribuyó a la interoperabilidad de las herramientas y las metodologías que se utilizan en el campo de la edición digital académica y de investigación.

El grupo mantiene colaboraciones con distintos grupos e instituciones nacionales e internacionales, entre los que se encuentran: Fundación El Caño en Panamá; los grupos LOEP, LEETHI y GAIA en la UCM; el departamento de Informática de Sistemas y Computadores de la Escuela Técnica Superior de Ingeniería Informática de la Universitat Politècnica de València; los grupos NUPILL y LAPESD de la Universidad Federal de Santa Catarina (Brasil); el Arborology Group de la Czech Technical University en la República Checa; el grupo gPel de la Universidade do Minho, en Portugal; el Departamento de Informática Aplicada de la Universidad de Novi Sad (Serbia); el Grupo sobre Sistemas de Visualización Médica Avanzada de la Universidad de Salamanca; o el Medical Physics Laboratory (MedPhys) de la Universidad de Thessaloniki, Grecia.

Varios miembros del grupo colaboran activamente en diversos comités de estandarización y normalización de UNE (antes AENOR). En particular, cabe destacar la participación en el grupo de trabajo CNT 71/SC 36/GR 12 que, liderado por la Profesora Ana M^a Fernández-Pampillón, ha elaborado la norma UNE 71362 para la evaluación de la calidad de los materiales educativos digitales de los repositorios educativos (Fernández-Pampillón, 2014, 2017).

Para finalizar, desde su creación varios investigadores del grupo han participado en la organización de diversos eventos científicos y edición de distintos números especiales en revistas de prestigio. A este respecto, cabe destacar: (i) la serie de talleres ISELEAR sobre Ingeniería del Software en E-Learning, que, actualmente, ha alcanzado ya su décima edición (Conde, Sarasa y Sierra, 2019); (ii) la organización de la 8th International Conference on Language and Automata Theory and Applications (Dediu, Martín-Vide, Sierra y Truthe, 2014); (iii) la organización del 4th Symposium on Languages,

Applications and Technologies (Sierra, Leal, Simões, 2015); o (iv) la edición de los números especiales sobre “Software Development Concerns in the e-Learning Domain” (Sierra y Sarasa, 2014) y “Software e-Learning Architectures” en la revista Science of Computer Programming (Sierra y Sarasa, 2016), y “e-Vocabularies and e-Learning” en la revista Education Science (Fernández-Pampillón y Pareja, 2017).

El resto del artículo resume alguna de las líneas que actualmente se llevan a cabo en el seno del grupo sobre producción de repositorios de objetos de aprendizaje en dominios especializados. En la sección 2 se describen los trabajos relativos a la creación de plataformas de gestión. La sección 3 resume los trabajos relativos al soporte semi-automático para la catalogación de objetos educativos. La sección 4 se centra en los aspectos relativos a la navegación por los repositorios. La sección 5 resume los trabajos relativos a la transformación de repositorios mediante gramáticas formales. Por último, la sección 6 concluye el artículo.

2. Plataformas de Gestión de Repositorios de Objetos Educativos Reconfigurables

En dominios altamente especializados (p.e., arqueología, literatura, ciencias de la salud...) es frecuente objetivar la existencia de material con alto valor didáctico, pero que se encuentra disponible en formas no estandarizadas (p.e., bases de datos específicas para una determinada plataforma), o bien organizado de acuerdo con diferentes estándares o, incluso, disperso en recursos digitales de naturaleza dispar (p.e., hojas Excel, documentos Word, bases de datos Access o Filemaker, etc.). De esta forma, la heterogeneidad de todos estos objetos es una característica inherente de los mismos, y constituye uno de los principales escollos a salvar de cara a su integración y reutilización, aspectos ambos que resultan críticos para la preservación y puesta en valor de los mismos.

En estos dominios altamente especializados, también es frecuente que los especialistas necesiten crear, a partir de fuentes de información generales, repositorios altamente especializados que faciliten la difusión y explotación -investigadora o didáctica- de sus contenidos. En estos dominios tan especializados es frecuente, asimismo, que no existan acuerdos claros sobre los esquemas de estructuración más convenientes de los contenidos, por lo que los

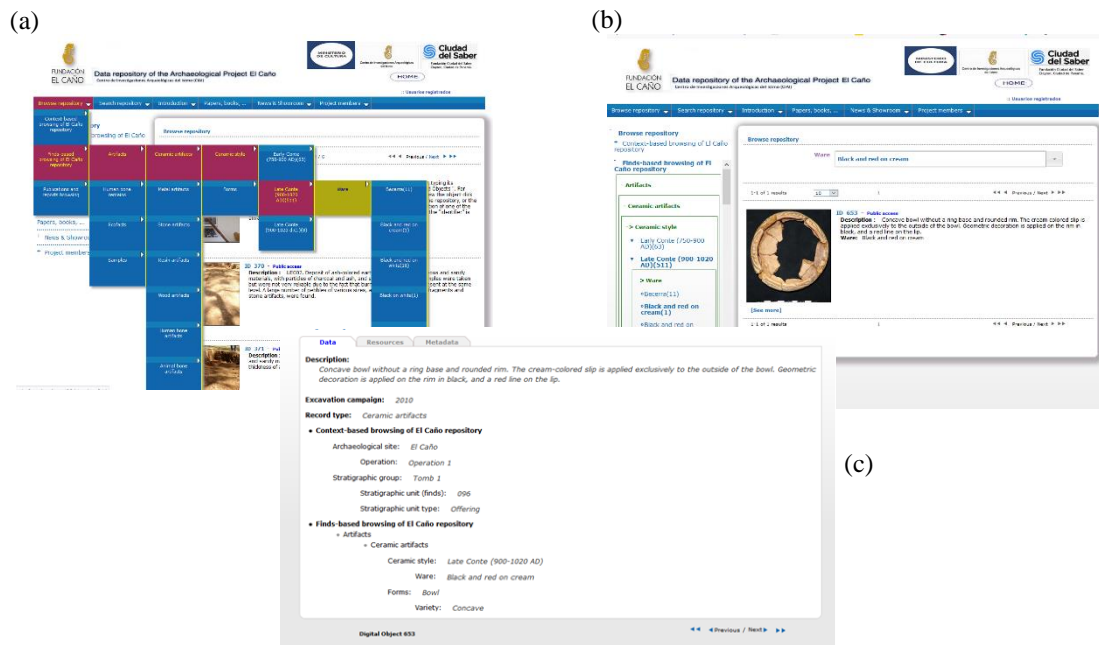


Figura 1. Algunas capturas de pantallas de un repositorio Oda: (a) Taxonomías de navegación; (b) visión general de un objeto; (iii) visión de detalle mostrando algunos de sus datos.

esquemas se crean, revisan y modifican con frecuencia durante la creación de los repositorios (Fernández-Valmayor et al., 2018). Por tanto, en estos dominios es necesario abordar la creación inductiva e incremental, no únicamente de los repositorios, sino también de sus esquemas de catalogación. Dichos esquemas se formulan, revisan, modifican y reconfiguran dinámicamente, a lo largo de todo el ciclo de vida de los repositorios.

En este contexto, una de las principales líneas de investigación desarrollada en el grupo ILSA se centra en la formulación de modelos de construcción de repositorios reconfigurables, que permitan definir inductivamente, y reconfigurar dinámicamente, repositorios de objetos educativos altamente especializados. En las siguientes subsecciones se describen brevemente dos de estas plataformas: Oda (subsección 2.1) y Clavy (subsección 2.2).

2.1 Oda

El sistema Oda (Fernández-Valmayor et al., 2013) permite crear repositorios de objetos digitales dinámicamente reconfigurables a usuarios no expertos en informática. Oda fue diseñado por el profesor Alfredo Fernández-Valmayor a partir de la experiencia del proyecto Chasqui, un repositorio de

objetos digitales arqueológicos procedentes de las piezas de un museo arqueológico precolombino de la UCM (Navarro et al., 2005). Entre sus principales características, la plataforma permite:

- La creación de objetos educativos de cualquier tipo y complejidad estructural. Así, por ejemplo, los objetos educativos pueden compartir archivos de otros objetos de un mismo repositorio, o crearse mediante la agregación de otros objetos más simples. En última instancia Oda permite crear redes de objetos y archivos.
- Un mecanismo de creación de repositorios sencillo y dinámico. Mediante una interfaz gráfica, cualquier especialista (no necesariamente informático) puede diseñar la estructura del repositorio y de los objetos educativos. Los objetos educativos se crean, después, rellenando una plantilla. La estructura del repositorio y la de los objetos puede modificarse en cualquier momento. Los objetos educativos creados adaptarán automáticamente su información a la nueva estructura.
- La creación inductiva de repositorios de objetos educativos (consecuencia de la anterior característica). De esta forma los especialistas pueden ir ajustando y adaptando el repositorio

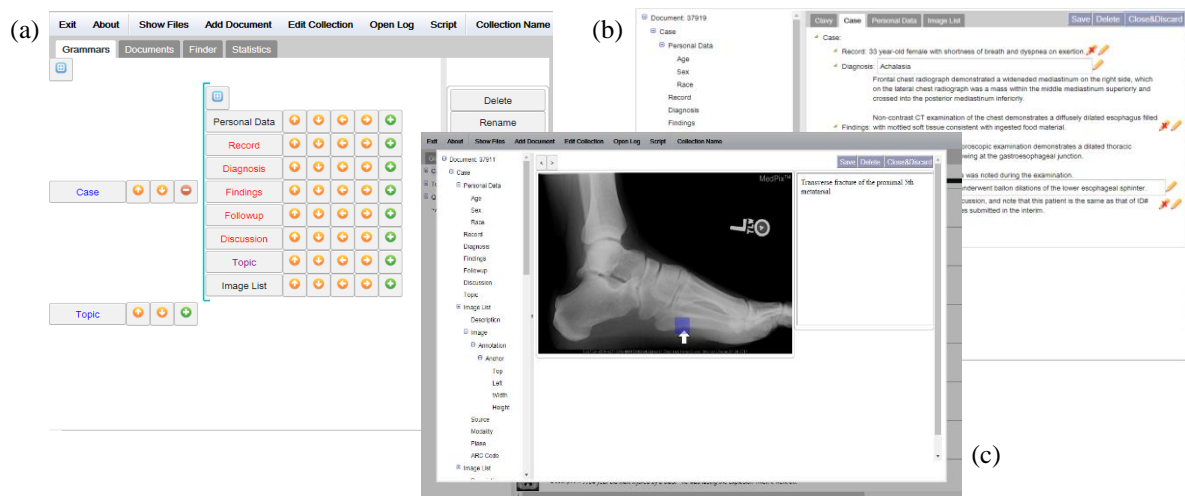


Figura 2. Algunas capturas de pantallas de la plataforma Clavy: (a) editor de esquemas; (b) editor de objetos de aprendizaje; (c) anotación de una imagen a través del correspondiente conector de edición.

(estructura y contenido) y sus objetos conforme este crece o según cambian las necesidades de los usuarios.

- El mecanismo de navegación y búsqueda es totalmente redefinible, al igual que los repositorios y se crea a partir de la estructura de los objetos.
- La creación colaborativa de los repositorios de objetos educativos. En la construcción de los repositorios pueden participar, con diferentes roles, todos los usuarios que se desee.
- Tres niveles de privacidad: para todo el repositorio, para los objetos educativos o para partes de los objetos educativos.
- Por sus características, OdA es útil, entre otros casos, para la definición de repositorios altamente especializados y de naturaleza heterogénea. También para crear repositorios complejos cuya estructura de datos e información no está claramente definida al comienzo del proyecto o necesita ir modificándose durante la vida del repositorio. Por esta razón, los repositorios creados con OdA están en evolución continua, especialmente en la fase inicial del repositorio, en la que el esquema de catalogación está creándose. La Figura 1 muestra algunas capturas de pantallas relativas a un repositorio OdA sobre materiales arqueológicos mantenidos por la Fundación El Caño, en Panamá. OdA cuenta con un registro de Propiedad Intelectual M005952/2012 (Fernández-Valmayor et al., 2013), y está

disponible en <https://github.com/ILSA-UCM/OdA>.

2.2 Clavy

El sistema Clavy es una evolución de OdA. En este sentido, tiene las mismas capacidades de reconfiguración y creación incremental de repositorios, pero además (Gayoso et al, 2017; Buendía et al., 2019):

- Aporta un modelo más expresivo de definición de esquemas de catalogación, que permite definir, por ejemplo, elementos de información multivaluados (i.e., que pueden ocurrir en los objetos un número arbitrario de veces), así como asociar atributos con los elementos (lo que resulta esencial para automatizar distintos procesos que operan sobre los repositorios).
- Aporta mecanismos para la importación y exportación de objetos digitales desde y hacia otros repositorios, así como para llevar a cabo transformaciones entre repositorios.
- Aporta mecanismos que permiten especializar el flujo de autoría / reproducción de los objetos educativos, permitiendo la incorporación de mecanismos de edición y reproducción de partes concretas de los objetos en términos específicos de cada tarea de edición / reproducción o de cada dominio. En particular, la versión actual de Clavy integra: (i) una herramienta de anotado de contenidos digitales basada en la anteriormente citada herramienta

@note; y (ii) una herramienta para la edición / visualización de información geolocalizada.

De esta forma, Clavy incorpora:

- Arquitecturas flexibles de conectores que permiten: (i) la importación y exportación de contenidos; (ii) la transformación de repositorios; y (iii) la especialización del flujo de autoría.
- Una interfaz gráfica para que los especialistas en las disciplinas de los repositorios puedan modificar esquemas de catalogación o hacer compatibles entre sí esquemas de catalogación de repositorios diferentes. A este respecto, Clavy aporta un mecanismo mucho más expresivo que el de ODA para la definición de esquemas, así como una mayor eficiencia a la hora de soportar la reconfiguración de los esquemas definidos.

La Figura 2 muestra algunas capturas de pantalla relativas a un repositorio Clavy en el dominio de las Ciencias de la Salud. Se trata de un repositorio que se ha obtenido exportando y reconfigurando contenidos obtenidos de MedPix (<https://MedPix.nlm.nih.gov/>), una colección de casos clínicos sobre radiología (Buendía et al., 2019).

3. Soporte a la Creación de Repositorios

Los repositorios de objetos educativos permiten agregar contenidos digitales interrelacionados para formar objetos dotados de un objetivo pedagógico bien definido (Polsani, 2003). A fin de permitir la adecuada clasificación de los objetos y su posterior búsqueda y recuperación, facilitando funcionalidades básicas para la reutilización de los mismos en distintos escenarios y/o aplicaciones educativas, dichos objetos deben catalogarse adecuadamente, bien de manera descriptiva, mediante metadatos (Miller, 2011), bien de manera temática, mediante vocabularios controlados (Fernández-Pampillón, 2017), o bien, como resulta lo más habitual, mediante una combinación de ambas técnicas. El proceso de catalogación conlleva las siguientes actividades:

- Por una parte, es necesario establecer los esquemas de catalogación utilizados para organizar los contenidos. A este respecto, una línea de trabajo que ha resultado particularmente fecunda ha sido la creación de estándares para la catalogación (p.e., LOM 2002). No obstante, tal y como ya se ha

discutido en la sección anterior, en dominios especializados dichos estándares pueden resultar, para algunos aspectos, innecesariamente detallados, e insuficientes para otros más específicos. Asimismo, tales estándares difícilmente cubren los aspectos relativos a la catalogación temática, al ser dicha catalogación también muy específica (Harpring, 2013). Por tanto, en la creación de repositorios especializados, la formulación de esquemas de catalogación específicos es un aspecto esencial. Sin embargo, la creación de este tipo de esquemas, que son, en última instancia, modelos de dominio muy especializados, puede ser un proceso complejo.

- Por otra parte, es necesario catalogar los objetos en sí de acuerdo con los esquemas de catalogación propuestos. Este proceso, cuando se lleva a cabo de manera manual, es un proceso tedioso, costoso y propenso a errores y omisiones.

De esta forma, actualmente estamos trabajando en la formulación e implementación de herramientas que permitan paliar estas dificultades. Tales propuestas, en lugar de concebirse como sustitutos de los catalogadores humanos, se conciben como herramientas de autoría avanzadas (integrables, p.e., en la plataforma Clavy) para dar lugar a sistemas semiautomatizados. A continuación, resumimos uno

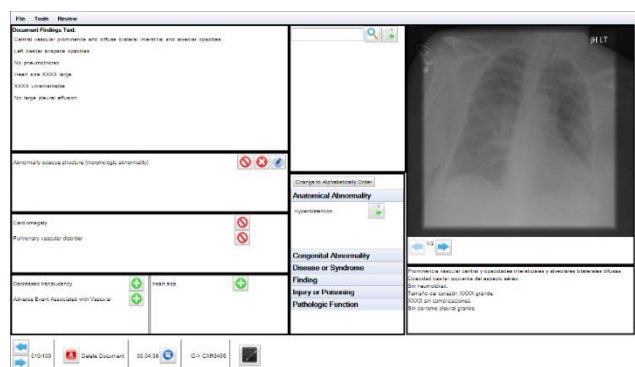


Figura 3. Captura de pantalla del prototipo de ayuda a la catalogación de informes médicos con términos UMLS mediante MetaMap.

de los trabajos ya desarrollados en esta línea.

3.1 Soporte a la catalogación mediante el refinamiento de ontologías genéricas

Con el fin de dar soporte a los expertos que crean los repositorios, hemos desarrollado un mecanismo inicial para automatizar parte de las tareas de creación de dichos repositorios. Dicho modelo parte de una ontología genérica sobre el dominio, así como de un anotador automático de los contenidos textuales de los objetos educativos con conceptos de dicha ontología. Como resultado, genera un repositorio inicial, que, posteriormente, puede reconfigurarse para adaptarlo mejor a las necesidades específicas de los expertos. Más concretamente, el modelo parte de:

- Una ontología genérica sobre el dominio en el que se formula el repositorio.
- Un anotador automático de textos basado en dicha ontología, que reconoce y anota las ocurrencias de conceptos en los textos.

Estos componentes permiten que el creador del objeto educativo añada los contenidos de los objetos, y que el sistema proponga conceptos relevantes para catalogar dichos objetos. Como resultado, es posible generar un repositorio inicial, con un esquema de catalogación inicial. Dicho esquema puede, posteriormente, ser revisado y reconfigurado por los expertos, utilizando para ello las técnicas de reconfiguración proporcionadas por la plataforma Clavy.

Para realizar una validación inicial de este modelo, hemos implementado un prototipo para la catalogación de objetos basados en informes clínicos en el dominio de ciencias de la salud (Figura 3) (Buendía, Gayoso y Sierra, 2019). Dicho prototipo utiliza UMLS (Unified Medical Language System) (Bodenreider, 2004) como la ontología genérica en base a la que clasificar los informes, así como MetaMap (Aronson, 2001) como anotador automático de textos médicos con conceptos UMLS. Con el fin de ir acotando los conceptos relevantes, el prototipo permite al anotador crear incrementalmente una lista de parada con conceptos no relevantes propuestos por MetaMap, así como deshabilitar conceptos irrelevantes para cada documento concreto. Asimismo, para aumentar el rango de aplicabilidad del enfoque, hemos incorporado la posibilidad de aplicar sistemas de traducción automática a los informes (en particular, Google Translate), como paso previo (Buendía, Gayoso, Juanes, Martín y Sierra, 2019). El prototipo se ha integrado en Clavy, utilizando para ello la infraestructura para la

especialización del flujo de autoría de la plataforma. De esta forma:

- El creador del repositorio puede ir incorporando incrementalmente informes y catalogándolos con la ayuda del componente de autoría basado en el prototipo citado.
- En cualquier momento, el creador puede utilizar el editor de esquemas de catalogación de Clavy para reconfigurar el esquema resultante.

4. Navegación en repositorios

Otro aspecto crítico para fomentar la utilidad práctica de los repositorios radica en la localización efectiva de los recursos. En concreto, los mecanismos de navegación por los materiales educativos juegan un papel fundamental. A este respecto, actualmente investigamos en distintos mecanismos para organizar los sistemas de navegación en los repositorios: organización de sistemas de navegación guiada robustos a las reconfiguraciones de los repositorios (subsección 4.1), navegación basada en similitud semántica (subsección 4.2), y, mapas de progresión temática (subsección 4.3). Finalmente, hemos empezado a investigar la posibilidad de utilizar los sistemas de navegación como recursos didácticos en sí mismos que ayudan a los estudiantes a entender y aprender los conceptos del dominio de conocimiento del repositorio (subsección 4.4).

4.1. Sistemas robustos de navegación guiada

La mayor parte de las plataformas de repositorios existentes incluyen mecanismos de navegación guiada por los esquemas de catalogación. No obstante, dichos sistemas son especialmente sensibles a modificaciones en los esquemas (ya que, en el peor de los casos, dichas modificaciones supondrán reindexar completamente los contenidos). A este respecto, estamos trabajando en modelos de navegación guiada que, bien no se vean afectados por tales modificaciones, bien minimicen el impacto de las mismas, y que, al mismo tiempo, proporcionen rendimientos óptimos.

En (Gayoso, Cerezo y Sierra 2016, Gayoso et al. 2017) describimos uno de estos modelos, basado en autómatas finitos, especialmente concebido para abordar el tipo de reconfiguraciones soportadas en la plataforma Clavy. En dicho modelo:

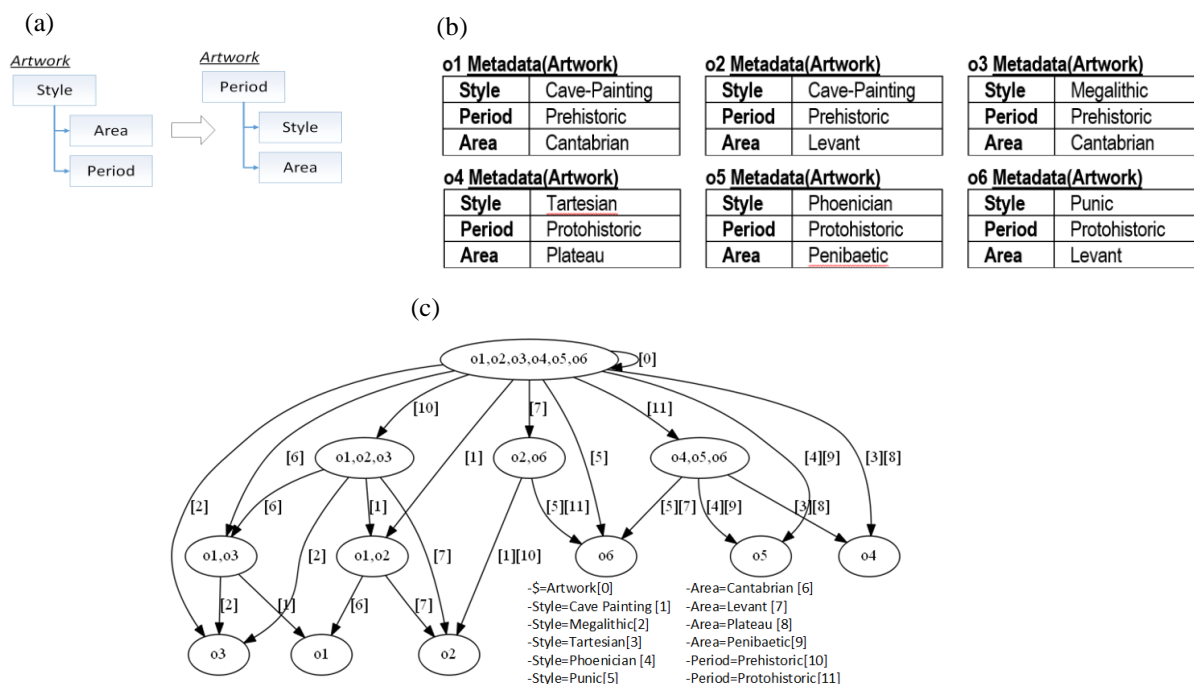


Figura 4. (a) Dos esquemas alternativos para un repositorio (el segundo es una reconfiguración del primero); (b) objetos en el repositorio; (c) autómata de navegación (permanece invariante ante la reconfiguración del esquema).

- La navegación en sí se soporta internamente mediante un sistema de navegación basado en etiquetas. En este tipo de sistemas, el usuario elige etiquetas para restringir el conjunto de objetos seleccionados. Para ello, los repositorios se abstraen como *sistemas basados en etiquetas* (Eck y Schaefer, 2011). A este respecto, los diferentes elementos de información de los objetos se mapean en etiquetas atómicas. Asimismo, dichos objetos se conciben como *bolsas* formadas por dichas etiquetas.
- El sistema resultante se modela como un autómata finito determinista, el *autómata de navegación*, en el que: (i) los estados representan conjuntos de objetos seleccionados; y (ii) las transiciones tienen asociadas etiquetas utilizadas para filtrar dichos objetos.

La Figura 4 ilustra el modelo mediante un repositorio muy simple en el dominio del arte prehistórico y protohistórico en España. Tal y como hace patente dicha figura, el autómata de navegación es invariante a las distintas reconfiguraciones de los esquemas (entendidas estas como reordenaciones en la jerarquía de los elementos de información).

Aunque para repositorios de tamaño moderado (algunos miles de objetos) hemos comprobado la factibilidad de mantener explícitamente el autómata de navegación, conforme aumenta la complejidad de los repositorios puede aumentar la complejidad de dichos autómatas (en el peor de los casos, su número de estados puede crecer exponencialmente con el número de objetos del repositorio -Gayoso et al., 2017-). Para paliar este inconveniente hemos propuesto alternativas de indexado basadas en autómatas finitos no deterministas equivalentes a los autómatas de navegación (Gayoso et al., 2016, 2017), así como en el uso de cachés semánticas que permitan recrear eficientemente, y de manera incremental, los autómatas durante la navegación (Gayoso, Gómez y Sierra, 2018a, 2018b, 2019).

4.2. Sistemas de navegación basados en mapas semánticos léxicos

Los sistemas de navegación basados en mapas semánticos léxicos permiten la navegación, siguiendo algún tipo de representación topológica del léxico que representa los conceptos del dominio, basada en la similitud semántica entre palabras-conceptos de las descripciones textuales de los objetos de un

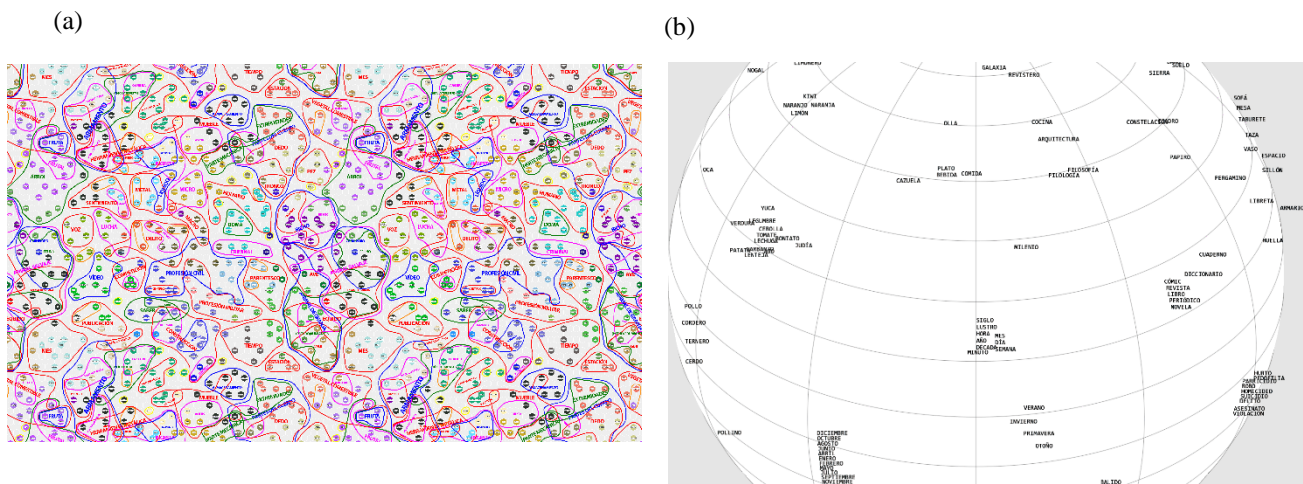


Figura 5. (a) Mapa de similitud semántica de 354 palabras (las líneas de colores se han creado manualmente para mostrar las diferentes categorías semánticas -clústers); (b) Mapa de Kohonen de similitud semántica tridimensional del léxico en (a).

repositorio digital. Para lograr de forma automática estos mapas semánticos (que son mapas conceptuales), actualmente, estamos trabajando en el estudio y aplicación de un modelo de aprendizaje profundo para analizar y representar los dominios de conocimiento: la teoría HTM (Hierarchical Temporal Memory). HTM es una teoría sobre el funcionamiento del córtex del cerebro aplicada exitosamente a muchos campos de organización y procesamiento de la información (Hawkins, Ahmad y Dubinsky, 2011). Esta teoría, a su vez, presenta una importante similitud con el modelo de memoria dinámica propuesto en (Fernández-Valmayor y Fernández, 1992).

La teoría HTM permite modelar una memoria predictivo-asociativa muy robusta al ruido, explicada en términos de un algoritmo de aprendizaje cortical, que reconoce de forma predictiva o aprende patrones de las señales de entrada formando memorias, y genera como salida determinadas respuestas (p.e., señales de aviso si se detectan anomalías en las entradas). De esta forma, utilizando esta teoría estamos trabajando en el diseño de un sistema de navegación que, dado un determinado contexto léxico que represente a los objetos de un repositorio, permita crear mapas léxico-semánticos para navegar en términos de la similitud conceptual de los objetos. En el punto actual de la investigación, se ha logrado:

- Definir un nuevo algoritmo de aprendizaje no supervisado de etiquetado morfosintáctico y lematización de textos que mejora

sustancialmente los algoritmos anteriores del mismo tipo y que tiene niveles de precisión semejantes a los algoritmos supervisados. El algoritmo se ha implementado como un etiquetador-lematizador para el español (<https://pypi.org/project/syntactic-tagger/>). Se ha aplicado sobre dos corpus del español (uno formado a partir del contenido íntegro de Wikipedia, y un segundo corpus de 560 novelas de 63 escritores españoles aún vivos o muertos en el siglo XXI). El resultado ha sido un lexicón computacional del español, que contiene 3.936.036 formas y 4.229.729 etiquetas.

- Definición de un nuevo sistema de representación léxica basado en semántica distribuida, apta para servir de entrada a la HTM, y que incorpora información morfosintáctica. Según las primeras pruebas de evaluación de la corrección de la representación, parece que mejora sustancialmente los sistemas actuales de representación semántico-léxica distribuidos.
- Creación de una representación gráfico-topológica mediante mapas autoorganizados de Kohonen, basada en cercanía semántica, en dos y tres dimensiones (Figura 5), del contenido de cualquier léxico que permite distinguir visualmente las distintas familias semánticas (clústers semánticos) y que constituirá la base

para la creación de los mapas semánticos de navegación.

4.3. Mapas de progresión temática

Para ayudar a localizar los objetos en un repositorio,

Forma	Lemma	Morph	Ro	Dep	Thematic Prog.	Text	Rate	n
El	el	ds0m0	2	det		El Ayuntamiento de Barcelona		1
Ayuntamiento	ayuntamiento	ap0000o	5	ncobj	A	instalará, finalmente, ocho cámaras para controlar la velocidad en las rondas de Barcelona, y no las seis		2
de	de	sp00	2	prepn				3
Barcelona	barcelona	ap0000i	3	pobj				4
instalará	instalar	vmaí3z0	21	coord				5
.	.	fc	5	punct				6
finalmente	finalmente	rg	5	sdmo				7
.	.	fc	5	punct				8
ocho	ocho	ds0cp0	10	det				9
cámaras	cámaras	scfp000	5	dobj	B			10
para	para	sp00	5	prepn				11
controlar	controlar	vma0000	11	pobj				12
la	el	ds0e0	14	det				13
velocidad	velocidad	scf0000	12	dobj				14
de	de	sp00	12	prepn				15
las	el	ds0fp0	17	det				16
rondas	rondas	scfp000	15	pobj				17
de	de	sp00	17	prepn				18
Barcelona	barcelona	ap00000	18	pobj				19
.	.	fc	21	punct				20
y	y	cc	0	root				21
no	no	rn	23	neg				22
las	el	ds0fp0	24	det				23
seis	seis	pa0sp00	21	coord				24
previstas	previsto	sq0fpp	24	smod				25
inicialmente	inicialmente	rg	25	sdmo				26
.	.	fp	21	punct				27
.	.							28
Los	el	ds0mp0	2	det		Los radares se podrán sitar, indirectamente, en alguno de los 26		29
radares	radar	scmp000	4	ncobj	B	repartidos entre la Ronda de Dalt y la del Litoral, además que el director de Circulación del		30
se	se	p000000	5	reflex		de		31
podría	poder	vmaí3p0	30	dobj		de		32
sitar	sitar	vma0000	4	cobj		de		33
.	.	fc	4	punct		de		34
indirectamente	indirectament	rg	5	sdmo		de		35
.	.	fc	4	punct		de		36
de	de	sp00	5	prepn		de		37

Figura 6. Análisis de la progresión temática (columnas coloreadas 1ª y 6ª) de dos resúmenes (en la 7ª columna).

los sistemas de búsqueda y navegación presentan breves descripciones textuales de los objetos que, a modo de resumen, sirven para que el usuario pueda seleccionar el objeto u objetos que busca sin necesidad de visualizar todo el contenido. Estas descripciones deben incorporarse, normalmente, de forma manual en los metadatos de cada objeto cuando se almacenan en el repositorio.

Para automatizar la creación de estas descripciones a partir de los contenidos textuales de los documentos que contienen los objetos estamos aplicando una propuesta basada en la teoría lingüística de la progresión temática propuesta por Daneš (1974). De forma resumida, la teoría de la progresión temática propone que la información que contiene un texto sigue un patrón de combinaciones de tema

(información conocida por el lector) y rema (nueva información) y que dichos temas y remas pueden identificarse automáticamente mediante marcas sintáctico-morfológicas (Figura 6). El sentido de un texto, por lo tanto, puede representarse como una progresión, estructurada, de conceptos “sabidos”, temas, y conceptos “nuevos” o remas. Esta estructura constituye una descripción conceptual del contenido de los objetos que puede utilizarse para: (i) generar las descripciones textuales de los objetos, y, (ii) extraer los conceptos claves y sus relaciones que



Figura 7. Sistema de navegación del diccionario didáctico digital de latín.

formarán los mapas de navegación léxicos. Actualmente, se está creando una herramienta de anotado de la progresión temática que facilitará el análisis lingüístico y la extracción automática de los patrones informativos de progresión temática implícitos en los textos descriptivos de los objetos de aprendizaje de un repositorio educativo.

4.4. Los sistemas de navegación como recurso didáctico

El uso de las aplicaciones OdA o Clavy para construir repositorios permiten la creación inductiva e incremental del "modelo de datos" y sistema de navegación de la colección de objetos digitales del repositorio de forma que represente el modelo del dominio de conocimiento de dicha colección. En este sentido, OdA y Clavy permiten obtener un modelo de datos y sistema de navegación aplicando una estrategia de construcción por etapas para obtener resultados similares a los diseños de modelos facetados basados en el conocimiento propuestos por Broughton (2013), o a las “faceted lightweight ontologies” propuestas por Giunchiglia et al. (2009) (Fernández-Valmayor et al., 2018). Además, el tipo de modelo a construir utilizando OdA o Clavy es una decisión de los especialistas, profesores o

investigadores, creadores del repositorio. Esto supone que los profesores pueden crear los modelos de datos de los repositorios educativos no solo para gestionar el material didáctico que contienen sino también como material didáctico en sí mismo. Esta idea ha sido aplicada a la elaboración de algunos diccionarios didácticos (en particular, diccionarios de latín y de alemán; véase, por ejemplo, Figura 7).

Los diccionarios son repositorios de objetos léxicos, en el que cada objeto léxico es una entrada léxica del diccionario. La novedad de estos diccionarios radica en que el sistema de navegación del diccionario sirve para explicar al alumno que se inicia en el aprendizaje de la nueva lengua, cómo entender y construir frases simples en esta nueva lengua (Márquez, Fernández-Pampillón y Sánchez, 2019). Los resultados académicos obtenidos hasta el momento (en tres grupos de iniciación al latín y uno de alemán) parecen indicar que existe una mejora significativa en el aprendizaje y en la motivación (concretamente el aspecto de actitud hacia el aprendizaje) respecto a los métodos actualmente utilizados en España en Enseñanza Secundaria en las dos lenguas de aplicación (Márquez Cruz, y Fernández-Pampillón, *en prensa*).

5. Transformación de repositorios mediante gramáticas formales

Desde la creación del grupo ILSA en 2010, hemos estado trabajando en la aplicación de métodos, técnicas y herramientas utilizadas en el diseño e implementación de lenguajes informáticos al desarrollo de aplicaciones educativas y a la gestión de información semiestructurada. A este respecto:

- Hemos integrado herramientas de generación de procesadores de lenguaje (en concreto, JavaCC y CUP) con marcos de procesamiento de documentos XML (Sarasa, Temprado, Cerezo y Sierra, 2012). Dicha integración permite la especificación declarativa de tareas de procesamiento de documentos XML muy complejas, como las que surgen, por ejemplo, en el desarrollo de generadores de aplicaciones (Sierra, Fernández-Valmayor y Fernández-Manjón, 2008).
- Hemos desarrollado metalenguajes de más alto nivel, basados en el formalismo de las *gramáticas de atributos* (Sarasa y Sierra, 2013a), para especificar tales tareas de procesamiento. El resultado es una herramienta

denominada XLOP (XML Language Oriented Processing) que, aceptando como entrada una especificación declarativa basada en gramáticas de atributos, permite generar automáticamente un procesador para una clase particular de documentos XML. La Figura 8 ilustra el

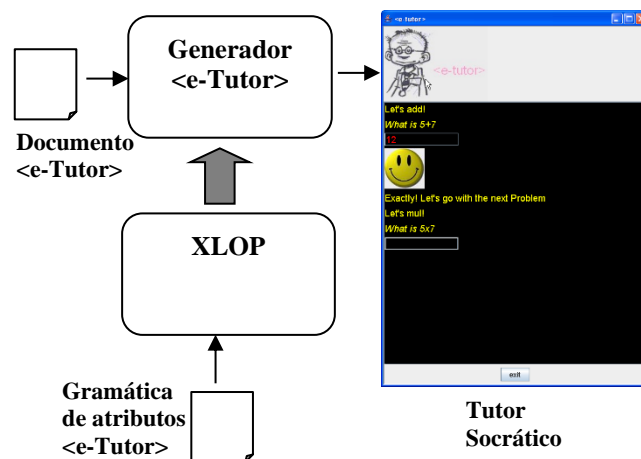


Figura 8. Desarrollo dirigido por gramáticas de un generador de tutoriales.

enfoque a la construcción de un sistema para la generación de tutoriales descritos como documentos XML. Dicho sistema, al cual denominamos <e-Tutor> (Sierra et al., 2008), puede generarse a su vez automáticamente con la herramienta XLOP a partir de una especificación declarativa mediante una gramática de atributos.

- Hemos aplicado el enfoque a otros tipos de información semiestructurada (p.e., documentos JSON -Sarasa y Sierra, 2013b-, o redes de objetos Java -Sarasa y Sierra, 2015-).
- Hemos desarrollado varios sistemas educativos para fomentar la enseñanza y el aprendizaje de las gramáticas de atributos (Sierra, Fernández-Pampillón y Fernández-Valmayor, 2008; Rodríguez, Gómez y Sierra, 2013; Rodríguez, Sarasa, Gómez y Sierra, 2014).

Mediante las facilidades de importación / exportación de Clavy hemos comprobado la aplicabilidad práctica de estos enfoques dirigidos por gramáticas a la transformación de repositorios, exportando los repositorios como documentos XML, transformando dichos documentos, e importando los documentos resultantes nuevamente en Clavy. Actualmente estamos trabajando en el desarrollo de un sistema

integrado en Clavy que permite especificar conectores de transformación mediante especificaciones basadas en gramáticas de atributos.

6. Conclusiones

En este artículo hemos introducido el grupo ILSA (Ingeniería de Lenguajes Software y Aplicaciones), y hemos descrito algunas de las líneas de investigación que actualmente estamos desarrollando en este grupo: plataformas de gestión de repositorios reconfigurables, herramientas de ayuda a la creación de los repositorios, navegación en repositorios, y transformación de repositorios dirigida por especificaciones declarativas basadas en gramáticas formales. Aunque estas líneas no agotan los intereses de investigación del grupo, sí constituyen un buen ejemplo del tipo de trabajos que actualmente se están desarrollando en el mismo.

6. Agradecimientos

Trabajo soportado por el proyecto TIN2017-88092-R.

7. Referencias

- Aronson A.R (2001). Effective mapping of biomedical text to the UMLS Metathesaurus: the metamap program. Proc. of the AMIA Symposium, 2001, pp. 17-21
- Bodenreider, O (2004). The unified medical language system (UMLS): integrating biomedical terminology. *Nucleic acids research* 32, 267-270
- Broughton, V. (2013). Faceted Classification as a General Theory for Knowledge Organization. *SRELS Journal of Inf. Management*, 50(6)
- Buendía, F., Gayoso, J., Juanes, J.A., Martín, M., Sierra, J.L (2019). Cataloguing Spanish Medical Reports with UMLS Terms. Proc. of the Seventh Int. Conf. on Technological Ecosystems for Enhancing Multiculturality, León, Spain, 16-18 Octubre
- Buendía, F., Gayoso, J., Juanes, J.A., Sierra, J.L (2019). Transforming Unstructured Clinical Free-Text Corpora into Reconfigurable Medical Digital Collections. Proc. of the 32nd IEEE CBMS Int. Symp. on Computer-Based Medical Systems (CBMS'19), Cordoba, Spain, June 5-7, 519-522
- Buendía, F., Gayoso, J., Sierra, J.L (2019). Generation of Standardized E-Learning Content from Digital Medical Collections. *J. of Medical Syst.* 43(188), 8pp. 2019
- Cigarrán, J., Gayoso, J., Rodríguez, M., Romero, D., Sarasa, A., Sierra, J.L (2014). Assessing semantic annotation activities with formal concept analysis. *Expert Syst. with App.*, 41(11), 5495-5508
- Conde, M.A., Sarasa, A., Sierra, J.L (2019). 10th International Workshop on Software Engineering for E-learning (ISELEAR'19). Proc. of the Seventh Int. Conference on Technological Ecosystems for Enhancing Multiculturality, León, Spain, 16-18 Octubre
- Daneš, F. (1974): Functional sentence perspective and the organisation of the text. En F. Daneš (ed.) *Papers on functional sentence perspective*. The Hague: Mouton. 106-128
- Dediu, A.H., Martín-Vide, C., Sierra, J.L., Truthe, B - Eds. (2014). *Language and Automata Theory and Applications - 8th International Conference, LATA 2014, Madrid, Spain, March 10-14, 2014*. Proceedings. Lecture Notes in Computer Science 8370, Springer
- Eck, O., Schaefer, D (2011). A semantic file system for integrated product data management. *Advanced Engineering Informatics*, 25(2), 177-184
- Fernández-Pampillón, A (2014). Development of a Spanish Standard for Quality Assessment of Digital Educational Material. *IEEE-RITA*, 9(4), 151-158
- Fernández-Pampillón, A (2017). The Role of E-Vocabularies in the Description and Retrieval of Digital Educational Resources. *Education Sciences* 7(33), 38 pp
- Fernández-Pampillón, A (2017). UNE 71362: calidad de los materiales educativos digitales. *Revista AENOR* 329, 44-47.
- Fernández-Pampillón, A., Pareja, A. (2017). E-Vocabulary and e-Learning (Editorial). *Education Science*, 7(1), 3pp.
- Fernández-Valmayor, A., Fernández, C (1992). Educational and research utilization of a dynamic knowledge base. *Comp. & Ed.*, 18(1-3), 51-61
- Fernández-Valmayor, A., Fernández, C., Navarro, A., Sierra, J.L., Sarasa, A., Chenlo, J.B., López, J.A., Segura, V., Fernández-Pampillón, A (2013). *Contenedor de Objetos Digitales OdA 2.0*. Patente de Propiedad Intelectual, referencia M-5952-12, nº de asiento registral 16/2013/1094
- Fernández-Valmayor, A., Fernández-Pampillón, A., Mayo, J., Guinea, M., Mayo, C., Hervás, M.A., Herrero, J (2018). *La Gestión de la Información en*

Iván Arias Rodríguez, Joaquín Gayoso Cabada, Ana M^a Fernández-Pampillón Cesteros, Manuel Márquez Cruz, Elena del Olmo Suárez, Antonio Sarasa Cabezuelo, José Luis Sierra Rodríguez

- el Proyecto Arqueológico El Caño (PAEC). En *Arqueología-ICA'18, Memoria del 56º Congreso Internacional de Americanistas*, Ediciones Univ. Salamanca, 157-171
- Gayoso, J., Cerezo, D., Sierra, J.L (2016). Multilevel Browsing of Folksonomy-based Digital Collections. En *Web Information Systems Engineering – WISE 2016 - 17th International Conference*, Shanghai, China, November 8-10, 2016, Proceedings, Part II. Lect. Notes in Comp. Sc. 10042, pp. 993-996. Springer
- Gayoso, J., Goicoechea, M., Gómez, M., Sanz, A., Sarasa, A., Sierra, J.L (2019). Ontology-enhanced educational annotation activities. *Sustainability* 11(6), 22pp
- Gayoso, J., Gómez, M., Sierra, J.L (2018a). Tag-Based Browsing of Digital Collections with Inverted Indexes and Browsing Cache. Proc. of the Sixth Int. Conf. on Technological Ecosystems for Enhancing Multiculturality, Salamanca, Spain, October 24-26
- Gayoso, J., Gómez, M., Sierra, J.L (2018b). Query-Based Versus Resource-Based Cache Strategies in Tag-Based Browsing Systems. En *Maturity and Innovation in Digital Libraries - 20th International Conference on Asia-Pacific Digital Libraries, ICADL 2018, Lecture Notes in Computer Science* 11279, 41-54. Springer
- Gayoso, J., Gómez, M., Sierra, J.L (2019). A Smart Cache Strategy for Tag-Based Browsing of Digital Collections. En *New Knowledge in Information Systems and Technologies - Volume 1, World Conference on Information Systems and Technologies, WorldCIST 2019, Galicia, Spain, 16-19 April, 2019. Advances in Intell. Systems and Computing* 930, 546-555, Springer
- Gayoso, J., Rodríguez, C., Sierra, J.L (2017). Browsing Digital Collections with Reconfigurable Faceted Thesauri. En J. Gołuchowski et al. (eds.), *Complexity in Information Systems Development, Lecture Notes in Inf. Systems and Organisation* 22, Springer, 69-86
- Giunchiglia, F., Dutta, B., Maltese, V. (2009): "Faceted lightweight ontologies". En: Borgida, A., Chaudhri, V.K., Giorgini, P., Yu, E.S.K., (eds.): *Conceptual Modeling: Foundations and Applications*. Volume 5600 of *Lect. Notes in Comp. Sc.*, Springer (2009) 36–51
- Harpring, P (2013). Introduction to Controlled Vocabularies: Terminology for Art, Architecture, and Other Cultural Works. Getty Research Institute.
- Hawkins, J., Ahmad, S., Dubinsky, D (2011). Hierarchical temporal memory including HTM cortical learning algorithms. Technical Report. Numenta, Inc. Palo Alto
- LOM (2002). IEEE Standard 1484.12.1-2002. IEEE Standard for Learning Object Metadata.
- Márquez Cruz, M., Fernández-Pampillón Cesteros, A. (en prensa) Motivación en el aprendizaje del latín: evaluación de una nueva metodología didáctica, *ReiDoCrea*, ISSN 2254-5883
- Márquez, M., Fernández-Pampillón, A., Sánchez, P. (2019). A novel Cognitive Model of Digital Didactic Dictionary for Learning Foreign Languages. Application to Latin and German. DEMO. Sixth biennial Conf. on Elect. Lexic., eLex 2019, Sintra, Portugal. 1-3 October
- Miller, S.J (2011). *Metadata for Digital Collections*. Neal-Schuman Publishers
- Navarro, A., Fernández-Valmayor, A., Fernández-Manjón, B., Sierra, J.L (2004). Conceptualization, Prototyping and Process Of Hypermedia Applications, *International J. Soft. Eng. and Know. Eng.*, 14(6), 565-602. 2004.
- Navarro, A., Sierra, J.L., Fernández-Valmayor, A., Hernanz, H (2005). From Chasqui to Chasqui II: an Evolution in the Conceptualization of Virtual Objects. *Journal of Universal Comp. Science* 11(9),1518-1529.
- Polsani, P (2003). Use and Abuse of Reusable Learning Objects. *Journal of Digital Information*, 3(4)
- Rodríguez, D., Gómez, M., Sierra, J.L (2013). Interactive educational simulations for promoting the comprehension of basic compiler construction concepts. *Innovation and Technology in Computer Science Education conference 2013, ITiCSE '13, Canterbury, United Kingdom - July 01 – 03*, 28-33
- Rodríguez, D., Sarasa, A., Gómez, M., Sierra, J.L (2014). Serious games in tertiary education: A case study concerning the comprehension of basic concepts in computer language implementation courses. *Computers in Human Behavior* 31, 558-570
- Romero, D., Bueren, J.L (2018). Networking Women Translators in Spain (1868-1936) and their Presence in the Mnemosyne Digital Library. *Electronic Library*, 36(2), 305-318

- Sarasa, A., Sierra J. L. S., Diez, Covadonga. (2018). A tool for the digital edition of interactive fiction using Stretchtext. En Proceedings of the 22nd International Conference Information Visualisation (IV), 356-361
- Sarasa, A., Sierra J.L (2013b). Grammar-Driven Development of JSON Processing Applications. Proc. of the 2013 Fed. Conf. on Comp. Sc. and Inf. Systems, Kraków, Poland, September 8-11
- Sarasa, A., Sierra, J.L (2013a). The grammatical approach: A syntax-directed declarative specification method for XML processing tasks. *Comp. Stand. & Int.* 35(1), 114-131
- Sarasa, A., Sierra, J.L (2015). A Syntax-Directed Model Transformation Framework Based on Attribute Grammars. Languages, Applications and Technologies - 4th International Symposium, SLATE 2015, Madrid, Spain, June 18-19, 2015, Revised Selected Papers. *Comm. in Comp. and Inf. Sc.* 563, 145-152, Springer
- Sarasa, A., Temprado, B., Rodríguez, D., Sierra, J.L (2012). Building XML-driven application generators with compiler construction tools. *Computer Science and Information Systems*, 9(2), 485-504
- Sierra, J.L (2016). Special issue on eLearning Software Architectures. *Science of Comp. Programming* 129, 1-2
- Sierra, J.L, y Sarasa, A (2012). ILSA (Ingeniería de Lenguajes Software y Aplicaciones), UCM. *IE Comunicaciones* 15, 95-102.
- Sierra, J.L., Fernández-Pampillón, A., Fernández-Valmayor, A. An environment for supporting active learning in courses on language processing. Proc. of the 13th Ann. SIGCSE Conf. on Inn. and Tech. in Comp. Sc. Ed., ITiCSE 2008, Madrid, June 30 - July 2, 128-132
- Sierra, J.L., Fernández-Valmayor, A., Fernández-Manjón, B (2008). From Documents to Applications Using Markup Languages. *IEEE Software* 25(2), 68-76
- Sierra, J.L., Fernández-Valmayor, A., Guinea, M., Hernanz, H (2006). From Research Resources to Learning Objects: Process Model and Virtualization Experiences. *Educ. Tech. & Society* 9(3), 56-68.
- Sierra, J.L., Leal, J.P., Simões, A. Languages, Applications and Technologies - 4th International Symposium, SLATE 2015, Madrid, Spain, June 18-19, 2015, Revised Selected Papers. *Communications in Computer and Information Science* 563, Springer
- Sierra, J.L., Sarasa, A (2014). Preface for the special issue on Software Development Concerns in the e-Learning Domain. *Science of Computer Programming* 88, 1-2
- Temprado-Battad, B., Sierra-Rodriguez, J. L., Sarasa-Cabezuelo, A. (2019). An Online Authoring Tool for Interactive Fiction. En Proc. of the 23rd International Conference Information Visualisation (IV), 339-344