

Resultados más notables de las investigaciones llevadas a cabo por el grupo ILSA (Ingeniería del Lenguaje, Software y Aplicaciones)

Investigadores del grupo ILSA:

- Alfredo Fernández-Valmayor Crespo (miembro honorario). Fundación Arqueológica El Caño
- Ana M^a Fernández-Pampillón Cesteros. Facultad de Filología, Universidad Complutense de Madrid
- Ángel Luis Encinas Moral. Facultad de Filología, Universidad Complutense de Madrid
- Antonio Pareja Lora. Facultad de Filosofía y Letras, Universidad de Alcalá de Henares.
- Antonio Sarasa Cabezuelo. Fac. Informática, Universidad Complutense de Madrid
- Aurora Santa Olalla. Facultad de Filología, Universidad Complutense de Madrid
- Elena del Olmo Suárez. Facultad de Filología, Universidad Complutense de Madrid
- Félix Buendía García. Escuela Técnica Superior de Ingeniería Informática. Universitat Politècnica de València
- Iván Arias Rodríguez. Facultad de Filología, Universidad Complutense de Madrid
- Joaquín Gayoso Cabada. Facultad de Informática, Universidad Complutense de Madrid
- José Luis Sierra Rodríguez (director). Facultad Informática, Universidad Complutense de Madrid
- José María Ruiz Ruiz. Facultad de Educación, Universidad Complutense de Madrid
- Manuel Márquez Cruz. Fac. Filología, Universidad Complutense de Madrid



Plataforma experimental de gestión de repositorios digitales reconfigurables *Clavy*


Clavy es un sistema que permite crear repositorios digitales con estructura dinámicamente reconfigurables. Al contrario que otras herramientas, que establecen esquemas de catalogación prefijados, *Clavy* permite definir el esquema de catalogación que mejor se adecua a cada repositorio, así como modificarlo para que se ajuste a las necesidades cambiantes que surgen durante todo el proceso de creación y preservación de dicho repositorio. Asimismo, *Clavy* está dotado de mecanismos para poblar repositorios desde fuentes externas de información, transformar repositorios para ajustarlos a necesidades específicas, y exportar información en formatos estándar para facilitar su inclusión en otras plataformas. Por último, *Clavy* permite también especializar los flujos de visualización y autoría mediante la inclusión de conectores específicos de edición y reproducción.

Clavy ha sido desarrollada a través de distintos proyectos del Plan Nacional de I+D+i (proyectos TIN2014-52010-R RedR+Human “Repositorios Educativos Dinámicamente Reconfigurables en Humanidades”, y TIN2017-88092-R CetrO+Spec “Creación, Exploración y Transformación de Repositorios Educativos en Dominios Especializados”), así como a través del proyecto MUGECODER “Modelo Unificado de Gestión de Colecciones Digitales con Estructuras Reconfigurables: Aplicación a la Creación de Bibliotecas Especializadas para Investigación y Docencia”, uno de los 5 proyectos financiados en el área de las Humanidades Digitales en la Convocatoria 2014 de ayudas a equipos de investigación de la Fundación BBVA.

Clavy está disponible en <http://clavy.fdi.ucm.es>.

Algunos trabajos sobre *Clavy*:

- Buendía, F., Gayoso-Cabada, J. & Sierra, JL. Generation of Standardized E-Learning Content from Digital Medical Collections. *J Med Syst* **43**, 188 (2019). <https://doi.org/10.1007/s10916-019-1330-5>
- Gayoso-Cabada J., Gómez-Albarrán M., Sierra JL. (2019) A Smart Cache Strategy for Tag-Based Browsing of Digital Collections. In: Rocha Á., Adeli H., Reis L., Costanzo S. (eds) *New Knowledge in Information Systems and Technologies. WorldCIST'19 2019. Advances in Intelligent Systems and Computing*, vol 930. Springer, Cham. https://doi.org/10.1007/978-3-030-16181-1_52
- Gayoso-Cabada J., Gómez-Albarrán M., Sierra JL. (2018) Query-Based Versus Resource-Based Cache Strategies in Tag-Based Browsing Systems. In: Dobrevá M., Hinze A., Žumer M. (eds) *Maturity and Innovation in Digital Libraries. ICADL 2018. Lecture Notes in Computer Science*, vol 11279. Springer, Cham. https://doi.org/10.1007/978-3-030-04257-8_4
- Gayoso-Cabada J., Rodríguez-Cerezo D., Sierra JL. (2017) Browsing Digital Collections with Reconfigurable Faceted Thesauri. In: Goluchowski J., Pankowska M.,

	<p>Linger H., Barry C., Lang M., Schneider C. (eds) Complexity in Information Systems Development. Lecture Notes in Information Systems and Organisation, vol 22. Springer, Cham. https://doi.org/10.1007/978-3-319-52593-8_5</p>
	<p>Herramienta de anotación colaborativa de textos literarios digitalizados @note.</p> <p>@note es una herramienta de anotación colaborativa de textos literarios que, aparte de permitir anotar colaborativamente dichos textos, también permite crear colaborativamente los esquemas de anotación utilizados para catalogar las anotaciones. @note es especialmente útil para orquestar actividades de anotación con propósito educativo. Para este fin, @note permite a los instructores configurar dichas actividades proporcionando (i) el texto a anotar; (ii) el grupo de estudiantes que colaborarán en la anotación del texto; y (iii) el esquema de anotación a utilizar durante la actividad. Dicho esquema se concibe como jerarquías conceptuales, que pueden ser creadas, reconfiguradas y reutilizadas por los instructores. Asimismo, @note soporta anotaciones ricas, que pueden integrar tanto texto, como imágenes, audio, video, o enlaces a recursos externos en la web, así como foros de discusión asociados con cada anotación. Por último, @note permite exportar borradores de análisis críticos sobre los textos anotados, mediante la selección e inclusión en los mismos de las distintas anotaciones creadas durante el proceso colaborativo de anotación.</p> <p>@note comenzó a desarrollarse a través de una de las 12 ayudas concedidas por Google a nivel mundial en la edición 2010 del “Digital Humanities Award Program”, ayuda que fue renovada en 2011 en base a la excelencia de los resultados obtenidos. Asimismo, la herramienta ha continuado perfeccionándose a lo largo de distintos proyectos tanto europeos (COST Action IS0704 - An Interoperable Supranational Infrastructure for Digital Editions, Interedition) como del Plan Nacional de I+D+i (proyecto TIN2010-21288-C02 GENHOE-VIRTUAL “Un enfoque generativo para la construcción de herramientas de producción y despliegue de objetos educativos en el campus virtual”, y los ya citados TIN2014-52010-R RedR+Human y TIN2017-88092-R CetrO+Spec)</p> <p>@note está disponible en http://a-note.fdi.ucm.es/</p> <p>Algunos trabajos sobre @note:</p> <ul style="list-style-type: none"> • Gayoso-Cabada, J.; Goicoechea-de-Jorge, M.; Gómez-Albarrán, M.; Sanz-Cabrerizo, A.; Sarasa-Cabezuelo, A.; Sierra, J.-L. Ontology-Enhanced Educational Annotation Activities. <i>Sustainability</i> 2019, 11, 4455. https://doi.org/10.3390/su11164455 • Cigarrán-Recuero, J.; Gayoso-Cabada, J.; Rodríguez-Artacho, M.; Romero-López, M-D.; Sarasa-Cabezuelo, A.; Sierra, J.-L. Assessing semantic annotation activities with formal concept analysis, <i>Expert Systems with Applications</i>, 41, 11, 2014, https://doi.org/10.1016/j.eswa.2014.02.036. • Gayoso-Cabada, J.; Sanz-Cabrerizo, A.; Sierra, J.-L. 2013. @note: an electronic tool for academic readings. In <i>Proceedings of the 1st International Workshop on</i>

	<p>Collaborative Annotations in Shared Environment: metadata, vocabularies and techniques in the Digital Humanities (DH-CASE '13). Association for Computing Machinery, New York, NY, USA, Article 17, 1–4. https://doi.org/10.1145/2517978.2517996</p> <ul style="list-style-type: none"> • Gayoso-Cabada, J.; Ruiz, C.; Pablo-Nuñez, L.; Sarasa-Cabezuelo, A.; Goicoechea-de-Jorge, M.; Sanz-Cabrerizo, A.; Sierra, J-L. A flexible model for the collaborative annotation of digitized literary works. DH 2012: 195-197. http://www.dh2012.uni-hamburg.de/conference/programme/abstracts/a-flexible-model-for-the-collaborative-annotation-of-digitized-literary-works.1.html
<p>XLOP</p>	<p>Herramienta XLOP para la especificación declarativa de procesamiento de documentos XML mediante gramáticas de atributos</p> <p>XLOP (<i>XML Language-Oriented Processing</i>) es una herramienta que facilita el desarrollo de tareas de procesamiento complejas de documentos XML. Para ello proporciona un lenguaje de especificación declarativo de tareas de procesamiento de documentos basado en el formalismo de las gramáticas de atributos. A partir de este tipo de especificaciones declarativas, XLOP genera automáticamente implementaciones eficientes de los procesamientos especificados.</p> <p>XLOP comenzó a desarrollarse en el marco del proyecto PR34/07 – 15865 “Tecnologías de Mercado Descriptivo -XML- como base a un proceso de desarrollo de software guiado por lenguajes” financiado por el Banco Santander y la UCM, y se consolidó en el marco del ya citado proyecto TIN2010-21288-C02 GENHOE-VIRTUAL.</p> <p>Algunos trabajos sobre XLOP:</p> <ul style="list-style-type: none"> • Sarasa-Cabezuelo, A.; Sierra, J-L. The grammatical approach: A syntax-directed declarative specification method for XML processing tasks, <i>Computer Standards & Interfaces</i>, Volume 35, Issue 1, 2013, Pages 114-131, https://doi.org/10.1016/j.csi.2012.06.006. • Temprado-Battad, B.; Sarasa-Cabezuelo, A.; Sierra, J-L. Managing the Production and Evolution of e-learning Tools with Attribute Grammars, <i>10th IEEE International Conference on Advanced Learning Technologies</i>, Sousse, 2010, pp. 427-431, https://doi.org/10.1109/ICALT.2010.124 • Sarasa-Cabezuelo, A.; Sierra, J-L.; Fernandez-Valmayor, A Processing Learning Objects with Attribute Grammars, <i>Ninth IEEE International Conference on Advanced Learning Technologies</i>, Riga, 2009, pp. 527-531, https://doi.org/10.1109/ICALT.2009.211 • Sarasa-Cabezuelo, A.; Temprado-Battad, B.; Sierra, J-L.; Fernández-Valmayor, A., XML Language-Oriented Processing with XLOP, <i>2009 International Conference on Advanced Information Networking and Applications Workshops</i>, Bradford, 2009, pp. 322-327, https://doi.org/10.1109/WAINA.2009.17.

<p>Evaluators</p>	<p>Sistema <i>Evaluators</i> para la generación de juegos y simulaciones interactivas educativos a partir de gramáticas de atributos</p> <p>Evaluators es una herramienta orientada a facilitar la enseñanza-aprendizaje de los mecanismos básicos subyacentes al formalismo de las gramáticas de atributos. La herramienta proporciona un lenguaje específico que permite especificar ejercicios de evaluación semántica relativos al formalismo de las gramáticas de atributos. Utilizando este lenguaje, los instructores especifican tareas de procesamiento dirigido por la sintaxis como gramáticas de atributos y sentencias a procesar. El sistema, entonces, permite generar tanto juegos serios como simulaciones interactivas, que permite a los estudiantes emular el proceso de evaluación semántica inducido por las gramáticas de atributos sobre los árboles de análisis sintáctico asociados con las sentencias proporcionadas.</p> <p><i>Evaluators</i> se desarrolló en el marco del proyecto TIN2010-21288-C02 GENHOE-VIRTUAL.</p> <p>Algunos trabajos sobre <i>Evaluators</i>:</p> <ul style="list-style-type: none"> • Rodríguez-Cerezo, D.; Sarasa-Cabezuelo, A.; Gómez-Albarrán, M.; Sierra, J-L. User-Centered Development of Generative Educational Systems for Computer Engineering: The <i>Evaluators</i> Case Study International Journal of Engineering Education, 31(3), 751-763. 2015. http://www.ijee.ie/latestissues/Vol31-3/08_ijee3034ns.pdf • Rodríguez-Cerezo, D.; Sarasa-Cabezuelo, A.; Gómez-Albarrán, M.; Sierra, J-L. Serious games in tertiary education: A case study concerning the comprehension of basic concepts in computer language implementation courses, Computers in Human Behavior, Volume 31, 2014, https://doi.org/10.1016/j.chb.2013.06.009 • Rodríguez-Cerezo, D.; Gómez-Albarrán, M.; Sierra, J-L. Interactive educational simulations for promoting the comprehension of basic compiler construction concepts. In Proceedings of the 18th ACM conference on Innovation and technology in computer science education (ITICSE '13). Association for Computing Machinery, New York, NY, USA, 28–33, 2013. https://doi.org/10.1145/2462476.2462498
<p>PAG</p>	<p>Sistema PAG para el prototipado de procesadores de lenguaje mediante gramáticas de atributos</p> <p>PAG (Prototyping with Attribute Grammars) es una herramienta que permite generar automáticamente prototipos en Prolog de procesadores de lenguaje a partir de especificaciones declarativas basadas en gramáticas de atributos. Los prototipos permiten experimentar con la ejecución de los procesadores especificados a través de una interfaz visual que, aceptando como entrada sentencias a procesar, permite inspeccionar los árboles de análisis sintáctico atribuidos, y navegar por dichos árboles tomando como guía las ecuaciones semánticas de la gramática. Esta característica confiere a la herramienta un claro valor didáctico.</p>


	<p>PAG se desarrolló en el contexto del proyecto del Plan Nacional de I+D+i TIN2005-08788-C04-01 OdA-Virtual “Objetos de Aprendizaje en el Campus Virtual”</p> <p>Algunos trabajos sobre PAG:</p> <ul style="list-style-type: none"> • Sierra, J-L, Fernández-Pampillon, A-M., Fernández-Valmayor, A. An environment for supporting active learning in courses on language processing. In Proceedings of the 13th annual conference on Innovation and technology in computer science education (ITICSE '08). Association for Computing Machinery, New York, NY, USA, 128–132, 2008. https://doi.org/10.1145/1384271.1384307 • Sierra, J-L., Fernández-Valmayor, A. A Prolog Framework for the Rapid Prototyping of Language Processors with Attribute Grammars, Electronic Notes in Theoretical Computer Science, Volume 164, Issue 2, 2006, Pages 19-36, https://doi.org/10.1016/j.entcs.2006.10.002.
<p>EvDebugger</p>	<p>Sistema EvDebugger para la depuración de especificaciones basadas en gramáticas de atributos</p> <p>EvDebugger es un sistema similar a PAG, pero específicamente orientado a la depuración de especificaciones basadas en gramáticas de atributos formuladas sobre gramáticas incontextuales LALR(1). EvDebugger permite transformar este tipo de gramáticas en implementaciones basadas en JLex + CUP, así como utilizar una interfaz visual que facilita la comprensión y depuración de dichas especificaciones.</p> <p>EvDebugger se ha desarrollado en el marco de los proyectos TIN2010-21288-C02 GENHOE-VIRTUAL “Un enfoque generativo para la construcción de herramientas de producción y despliegue de objetos educativos en el campus virtual” y TIN2014-52010-R RedR+Human “Repositorios Educativos Dinámicamente Reconfigurables en Humanidades”.</p> <p>Algunos trabajos relacionados con EvDebugger:</p> <ul style="list-style-type: none"> • Rodríguez-Cerezo, D.; Henriques, P.R: Sierra, J-L, Attribute grammars made easier: EvDebugger a visual debugger for attribute grammars, 2014 International Symposium on Computers in Education (SIIE), Logroño, 2014, pp. 23-28, IEEE, doi: https://doi.org/10.1109/SIIE.2014.7017699 • Rodríguez-Cerezo, D., Sarasa, A., Sierra, J-L: A systematic approach to the implementation of attribute grammars with conventional compiler construction tools. Comput. Sci. Inf. Syst. 9(3): 983-1017 (2012). https://doi.org/10.2298/CSIS111223022R
<p>ILSAEditor</p>	<p>Editor de ficción interactiva ILSAEditor</p> <p>ILSAEditor es un editor de libros interactivos para ficción interactiva que utiliza el editor HTML tinyMCE. Se caracteriza porque que un libro es una única página en la que se pueden añadir determinados recursos de ficción interactiva tales como texto extensible (stretchtext), o anotaciones. Así mismo, permite exportar/importar la página como html o libro epub, permite añadir metadatos necesarios para la generación del epub, permite dividir</p>


	<p>lógicamente el documento en secciones del libro, y permite visualizar un grafo que representa los vínculos en la página.</p> <p>ILSAEditor se ha desarrollado en el contexto del proyecto de la Comunidad de Madrid H2015/HUM-3426 eLITE-CM, Edición literaria electrónica</p> <p>El siguiente trabajo describe ILSAEditor:</p> <ul style="list-style-type: none"> • Sarasa-Cabezuelo, A, Sierra, J. L., Diez-Sanmartín, Covadonga. A tool for the digital edition of interactive fiction using Stretchtext. In 2018 22nd International Conference Information Visualisation (IV) (pp. 356-361). IEEE, 2018. https://doi.org/10.1109/iV.2018.00067
<p>IFDBMaker</p>	<p>Sistema para la creación de ficción interactiva y la lectura colaborativa IFDBMaker</p> <p>IFDBMaker es un sistema para la creación de ficción interactiva que mejora a ILSAEditor en aspectos tales como el desarrollo colaborativo (varios usuarios pueden modificar el documento) para lo cual se trabaja con libros estructurados en secciones independientes (que vienen representadas por diferentes páginas). Asimismo, soporta más recursos descriptivos, como personalización, lectura condicional, etc. También añade aspectos de lectura colaborativa y de red social.</p> <p>IFDBMaker se ha desarrollado en el contexto del proyecto de la Comunidad de Madrid H2015/HUM-3426 eLITE-CM, Edición literaria electrónica.</p> <p>El siguiente trabajo describe ILSAEditor:</p> <ul style="list-style-type: none"> • Temprado-Battad, B., Sierra, J. L., Sarasa-Cabezuelo, A. An Online Authoring Tool for Interactive Fiction. In 2019 23rd International Conference Information Visualisation (IV) (pp. 339-344). IEEE, 2019. https://doi.org/10.1109/IV.2019.00064
<p>OdA</p>	<p>El sistema OdA “Contenedor de Objetos de Aprendizaje”</p> <p>OdA permite crear repositorios de objetos digitales reconfigurables a usuarios no expertos en informática. Esta plataforma es el predecesor del sistema <i>Clavy</i>. Tiene un registro de Propiedad Intelectual con número de asiento registral 16/2013/1094.</p> <p>OdA se ha desarrollado en el contexto de los proyectos del Plan Nacional de I+D+i TIN2005-08788-C04-01 OdA-Virtual “Objetos de Aprendizaje en el Campus Virtual”, y TIN2010-21288-C02 GENHOE-VIRTUAL “Un enfoque generativo para la construcción de herramientas de producción y despliegue de objetos educativos en el campus virtual”.</p> <p>Algunos repositorios creados con OdA:</p>

- Repositorio Arqueológico del Proyecto El Caño (Panama), <http://oda-fec.org/nata>
- Mito e Imagen en la Antigüedad Clásica, <http://repositorios.fdi.ucm.es/Mythos/>
- Museo de Instrumentos Científicos de la Facultad de Ciencias Físicas (UCM), <http://museofis.ucm.es/museo/>
- Museo de Materiales de Laboratorios de Lenguas de la Facultad de Filología (UCM), <http://museos.hst.ucm.es/odafilol/>
- Diccionario Didáctico Digital de Latín, <http://repositorios.fdi.ucm.es/DiccionarioDidacticoLatin>
- Diccionario Didáctico Digital de Alemán, <http://repositorios.fdi.ucm.es/DiccionarioDidacticoAleman>
- Repositorio Digital de Hábitats y Especies Animales y Vegetales del Parque Regional de la Cuenca Media del Río Guadarrama (en construcción), <http://repositorios.fdi.ucm.es/especiesParqueGuadarrama>
- Corpus de aprendices del español (en construcción), http://repositorios.fdi.ucm.es/corpus_aprendices_español/

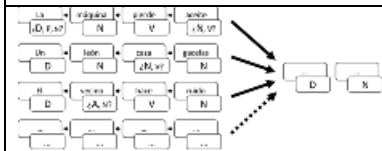
Algunos trabajos sobre OdA:

- Fernández-Valmayor, A. Fernández-Pampillón, A. Mayo, J.; Guinea, M.; Mayo, C.; Hervás, M.A.; Herrerín, J. "La Gestión de la Información en el Proyecto Arqueológico El Caño (PAEC)", [Actas 56º Congreso Internacional de Americanistas](#) (páginas 157-171). 15-20 Julio. Salamanca. España. Disponible en [Google Books](#) y en Researchate: https://www.researchgate.net/publication/329781730_GI-PAEC-ICA18-FINAL.
- Navarro, A; Fernández-Pampillón, A., Fernández-Chamizo, C. Fernández-Valmayor, A. (2013). A Meta-Relational Approach for the Definition and Management of Hybrid Learning Objects. En Educational Technology & Society Journal Volume 16, Issue 4, 2013. 258–274. https://www.j-ets.net/ETS/journals/16_4/20.pdf
- Sierra, J.L.; Fernández-Valmayor, A.; Guinea, M.; Hernanz, H. From Research Resources to Learning Objects: Process Model and Virtualization Experiences. Journal of Educational Technology & Society. Vol:9/3 pp: 56-68. 2006. <https://drive.google.com/open?id=1d6zhmNZ8DBQdKVbMWORpl-us9-H08jgM>

	<ul style="list-style-type: none"> • Fernández-Pampillón, A. Domínguez, E., Lahoz, J.M., Romero, D., De Armas, I. Palmaz, S., Arús, J. A Strategy for the Inductive Generation of Learning Objects in low-Tech Contexts. En Proceedings of the 10th European Conference on E-Learning (ECEL 2011) Brighton. UK. 10-11 November 2011. Disponible en: http://eprints.ucm.es/14024/
 <p>UNE 71362 Calidad de Materiales Educativos Digitales</p>	<p>Norma UNE 71362 Calidad de Materiales Educativos Digitales</p> <p>La norma UNE 71362 constituye un estándar español para la definición y la evaluación cuantitativa y cualitativa de los criterios básicos de calidad de los materiales educativos digitales, creando un marco común consensuado y validado -respecto a su eficacia, usabilidad y la fiabilidad- para los profesionales de las Tecnologías de la Educación en todos los sectores de actividad involucrados en la creación, selección y uso de contenidos educativos digitales, fundamentalmente: académico, empresarial y de las administraciones públicas.</p> <p>Esta norma ha sido desarrollada por el Grupo de Trabajo 12 (GT12) del subcomité CTN 71/SC 36 “Calidad de Materiales Educativos Digitales”. El GT12 integra tres miembros de ILSA uno de ellos dirige el GT12. La norma está oficialmente disponible en https://www.aenor.com/normas-y-libros/buscador-de-normas/une/?c=N0063263.</p> <p>Este proyecto se desarrolló a partir del modelo COdA formulado por ILSA en el contexto del proyecto TIN2014-52010-R RedR+Human “Repositorios Educativos Dinámicamente Reconfigurables en Humanidades”.</p> <p>Algunos trabajos relacionados con la norma y con CODA:</p> <ul style="list-style-type: none"> • Moreno, L.; Fernández-Pampillón, A.; Sarasa, A.; Rodrigo C.; García R.; García-Villalobos, J. (2019) How to combine accessibility with the didactic and technological quality of Digital Educational Materials. Journal of Accessibility and Design for All (ISSN 213-7087), Vol. 9. Issue. 2. Nov. 2019: http://www.jaccs.org/index.php/jaccs/article/view/193 • Fernández-Pampillón, A. (2017). UNE 71362 Calidad de los materiales educativos digitales. Revista AENOR. Septiembre 2017. Disponible en: http://www.aenor.com/revista/329/une-71362.html • Domínguez Romero, E.; De Armas Ranero, I.; Fernández-Pampillón Cesteros A. (2016). The COdA scoring rubric. En Technology-Enhanced Language

	<p>Learning for Specialized Domains: Practical Applications and Mobility. Editorial Routledge. Págs: 86-94. Disponible en Google books</p> <ul style="list-style-type: none"> • Fernández-Pampillón, A.M. (2014). Development of a Spanish Standard for Assessment of Digital Educational Material. IEEE Revista Iberoamericana de Tecnologías del Aprendizaje, Issue 4, pp: 151-158. IEEE Xplore Digital Library. Nov 2014.
 <p>DICCIONARIO DIDÁCTICO DIGITAL DE LATÍN</p>	<p>Diccionarios Didácticos Digitales</p> <p>En este proyecto se ha desarrollado un nuevo modelo cognitivo de diccionario y método didáctico que facilita las fases iniciales de comprensión y uso de una nueva lengua. El problema que aborda es la desmotivación inicial del alumno que encuentra dificultades en el aprendizaje de una nueva lengua, entre otros factores, por la falta de conocimientos lingüísticos suficientes en los que apoyarse. Esta situación se produce, fundamentalmente, en la enseñanza secundaria y en los inicios de los estudios superiores y, en muchos casos, desemboca en el abandono del aprendizaje en su etapa inicial.</p> <p>La idea básica del proyecto es que el alumno, con el apoyo del diccionario, entienda el funcionamiento global de una lengua antes de profundizar en cada uno de sus aspectos. Para ello, se utiliza un método cognitivo basado en la gramática de valencias, la gramática de casos de Fillmore y las ontologías de rasgos de Lyon. El método propone la comprensión de las frases a partir de los verbos principales y sus argumentos como si fueran piezas de un rompecabezas que deben encajarse para entender el significado de la frase.</p> <p>La eficacia didáctica del diccionario de latín está siendo probada empíricamente. Se ha encontrado un aumento de la motivación y una disminución de los errores en los ejercicios de comprensión y producción de frases en latín. Además, el modelo del diccionario se ha exportado a la lengua alemana con resultados satisfactorios.</p> <p>El modelo del diccionario ha sido registro en el Registro de Propiedad Intelectual con número de asiento registral 16/2020/9</p> <p>Las URL de dos diccionarios creados con este modelo son:</p> <ul style="list-style-type: none"> • Diccionario de Latín: http://repositorios.fdi.ucm.es/DiccionarioDidacticoLatín/ • Diccionario de Alemán: http://repositorios.fdi.ucm.es/DiccionarioDidacticoAleman/ <p>Algunos trabajos relacionados con este proyecto son:</p>

- del Olmo, Elena & Fernández-Pampillón, Ana. (2020). A new approach for extracting the conceptual schema of texts based on the linguistic Thematic Progression theory. Actas del Workshop on Hybrid Intelligence for Natural Language Processing Tasks, 24th European Conference on Artificial Intelligence (ECAI 2020). Disponible en: [ResearchGate](#) y [arXiv.org](#)

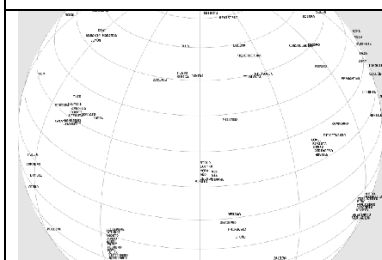


Etiquetador sintáctico con entrenamiento con texto plano. Versión completa para el español.

En este proyecto se ha creado un nuevo algoritmo de aprendizaje no supervisado para el etiquetado y lematización de textos. Se entrena con corpus de texto plano. Debe utilizar además un lexicón de formas con información lexicográfica básica: a cada forma se le asignan los lemas y las etiquetas morfosintácticas que puedan llevar. El algoritmo se ha implementado y evaluado para el español, una lengua fusionante con una morfología mucho más rica que la del inglés. Para su construcción se han utilizado dos corpus planos 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

Los resultados que se han obtenido respecto a la precisión son satisfactorios (>92% de precisión) e indican que se podría aplicar con resultados similares a otras lenguas que carecen de corpus etiquetados. Esto abre la posibilidad de poder empezar a aplicar técnicas de PLN a las lenguas que, por escasos recursos económicos, no han podido desarrollar todavía corpus etiquetados de tamaño suficiente para entrenar convenientemente a un etiquetador tradicional. El código está disponible en abierto en: <https://pypi.org/project/syntactic-tagger/>

Este trabajo se ha desarrollado en el contexto del proyecto TIN2017-88092-R CetrO+Spec “Creación, Exploración y Transformación de Repositorios Educativos en Dominios Especializados”



Sistema vectorial de representación de conocimiento léxico

En este proyecto en curso, se ha desarrollado un nuevo sistema de representación léxica basado en semántica distribuida que incorpora información morfológica, sintáctica y semántica. Se ha implementado un diccionario para el español 201.000 unidades léxicas (incluye nombres propios) con vectores de 48.000 dimensiones que, mediante SVD, se pueden reducir sin apenas pérdida de información entre 500 a 2000 dimensiones.

Las primeras pruebas de evaluación de la corrección de la representación, basadas en la cercanía entre unidades léxicas, muestran una mejora respecto a los sistemas actuales de representación semántico-léxica distribuidos.

Este trabajo se está desarrollando en el contexto del proyecto TIN2017-88092-R CetrO+Spec "Creación, Exploración y Transformación de Repositorios Educativos en Dominios Especializados"