

FECIES 2012



Loreto Del Río Bermúdez e Inmaculada Teva Álvarez (Comps.)

ISBN-13: 978-84-695-6734-0

FECIES 2012

Autor: IX FORO INTERNACIONAL SOBRE LA EVALUACIÓN DE LA CALIDAD DE LA INVESTIGACIÓN Y DE LA EDUCACIÓN SUPERIOR (FECIES). Santiago de Compostela (España), 12-15 de Junio de 2012.

Compiladores: Loreto Del Río Bermúdez e Inmaculada Teva Álvarez

Edita: Asociación Española de Psicología Conductual (AEPC).

CIF: G-23220056

Facultad de Psicología.

Universidad de Granada.

18011 Granada (España).

Tel: +34 958 273460.

Fax: +34 958 296053.

Correo electrónico: info@aepe.es. Web: <http://www.aepe.es>.

ISBN-13: 978-84-695-6734-0

NOTA EDITORIAL: Las opiniones y contenidos de los capítulos publicados en el libro “FECIES 2012”, son de responsabilidad exclusiva de los autores; asimismo, éstos se responsabilizarán de obtener el permiso correspondiente para incluir material publicado en otro lugar.

UNA EXPERIENCIA DE INNOVACIÓN DOCENTE EN EL GRADO DE COMUNICACIÓN AUDIOVISUAL: LA CREACIÓN DE UN DOCUMENTAL COMO UNA PRÁCTICA COMPARTIDA	
Natalia Quintas-Froufe, Sandra Martínez-Costa y Victoria de León-Sanjuán.....	811
EL PROYECTO REDES DE LA UNIVERSIDAD DE ALICANTE. UNA APUESTA POR LA INNOVACIÓN DOCENTE EN LA ENSEÑANZA SUPERIOR	
María Teresa Tortosa, Salvador Grau y José Daniel Álvarez	816
USO DIDÁCTICO DE LATEX EN LA ENSEÑANZA UNIVERSITARIA	
Ángel José Almeida-Rodríguez y Belén López-Brito.....	822
DIAGNOSTICO NEUROTOXICOLÓGICO: APLICACIÓN DE UN PROGRAMA INFORMATIZADO PARA SEMINARIOS Y CLASES PRÁCTICAS	
Mª Teresa Frejo, Margarita Lobo, Mª Jesús Díaz, Miguel Capó y Mª José Anadón	827
ANÁLISIS DE HERRAMIENTAS PARA LA DOCENCIA PRÁCTICA EN RECUPERACIÓN DE INFORMACIÓN	
Javier Parapar y Álvaro Barreiro.....	833
METODOLOGÍA Y ANÁLISIS DE LA EXPERIENCIA DE TRES EDICIONES DEL CONGRESO DE ACTIVIDADES ACADÉMICAMENTE DIRIGIDAS EN LA EPS DE LA UNIVERSIDAD DE CÓRDOBA	
Rafael E. Hidalgo-Fernández, Roberto Espejo-Mohedano, Mª Antonia Cejas-Molina, Antonio Blanca-Pancorbo y Ezequiel Herruzo-Gómez	840
INNOVACIÓN DOCENTE EN LA PLANIFICACIÓN ESTRATÉGICAS DE LOS SERVICIOS DE ENFERMERÍA	
Olga García-Martínez, Francisco Javier Manzano-Moreno, Elvira De Luna-Bertos, Javier Ramos-Torrecillas y Concepción Ruiz-Rodríguez.....	845
UTILIZACIÓN DE GRABACIONES EN VÍDEO PARA AUMENTAR LA EFICACIA DE LAS CLASES	
Luis Miguel Bravo-González, Beatriz Álvarez-Rodríguez, Manuel Mahamud-López y Juan María Menéndez-Aguado.....	849
LA FORMACIÓN DEL PROFESORADO EN SU ROL COMO TUTOR: UNA APROXIMACIÓN A LAS NECESIDADES PERCIBIDAS POR LOS/AS DOCENTES DE LA USC	
Beatriz García Antelo y Cristina Abeal Pereira	854
VENTAJAS DE LAS TUTORÍAS VIRTUALES SÍNCRONAS: UN ESTUDIO PILOTO EN LA UNIVERSIDAD DE VALENCIA	
Mª Dolores Sancerni y Fernando Cantalapiedra	860

ANÁLISIS DE HERRAMIENTAS PARA LA DOCENCIA PRÁCTICA EN RECUPERACIÓN DE INFORMACIÓN

Javier Parapar y Álvaro Barreiro

Universidad de A Coruña

Introducción

La adaptación de las titulaciones de Ingeniería Informática al Espacio Europeo de Educación Superior (EEES) ha supuesto tanto la renovación de la oferta de materias como el cambio en el paradigma docente establecido. En particular, la Facultad de Informática de la Universidad de A Coruña ha introducido en su *curricula* la asignatura de Recuperación de Información. La Recuperación de Información es ya, a día de hoy, una materia madura y establecida en el ámbito de las Ciencias de la Computación. La Universidad de A Coruña, que ha sido desde su constitución un referente en la comunidad autónoma de Galicia en el ámbito de la informática, la ha incluido como materia fundamental en sus nuevos planes. Concretamente, en el Grado de Ingeniería Informática por la Universidad de A Coruña, la asignatura de Recuperación de Información está asociada al itinerario de Computación y cuenta con 6 créditos. En el plan de Máster en Ingeniería Informática, recientemente propuesto, la asignatura de Recuperación de Información y Web Semántica cuenta también con 6 créditos ECTS. Gran parte de la docencia asociada a estas nuevas materias será de carácter práctico al tratarse de titulaciones en el ámbito de la Ingeniería.

En este escenario, existe pues una necesidad fundamental de contar con herramientas adecuadas que se adapten al nuevo paradigma educativo donde, de acuerdo al espíritu del EEES, aumenta el trabajo autónomo del alumno y se reducen las horas presenciales guiadas por un docente. Es pues nuestra intención a la luz de la nueva situación docente y metodológica revisar las herramientas existentes para la enseñanza práctica de Recuperación de Información, haciendo especial hincapié en los factores introducidos por las restricciones asociadas a la adaptación al EEES. En concreto en este trabajo analizaremos herramientas *software* considerando distintos factores importantes para la docencia, sin ánimo de ser exhaustivos: lenguaje de programación, licencia, comunidad, documentación, soporte, modelos disponibles, facilidad de evaluación, etc.

A pesar de la existencia de algunas comparativas de herramientas *software* desde el punto de vista de uso comercial o en investigación (Baeza-Yates, R., Ribeiro-Neto B. y

Middleton C. 2011, Cleger-Tamayo, S., Figuerola, C. G. y Rodríguez-Cano J. C. 2012), en este documento consideramos importante analizar las herramientas desde un punto de vista de su idoneidad para la docencia y el aprendizaje. Este trabajo se encuadrará en la línea metodológica y de recursos docentes en el marco de la adaptación al EEES y daremos respuesta a algunas preguntas importantes como: ¿qué herramientas son más adecuadas para el trabajo autónomo del alumnado?, ¿qué herramientas son más adecuadas dado el bagaje adquirido por el alumnado en el contexto de los planes de estudio de la Universidad de A Coruña?, ¿qué herramientas permitirán al docente poner en la práctica el temario explicado en las clases magistrales?, ¿qué herramientas facilitarán la evaluación continua del alumnado?

Método

Para dar respuesta a las preguntas formuladas en el apartado anterior procedimos a analizar las diferentes alternativas de herramientas *software* disponibles en el mercado. En particular dada su popularidad y la extensión de su uso centramos el estudio en cuatro herramientas ampliamente respaldadas por la comunidad de Recuperación de Información:

- Terrier (University of Glasgow 2012). TERabyte RetriEveR es un plataforma de indexación y búsqueda desarrollada en Java de manera modular que permite el desarrollo rápido de aplicaciones de búsqueda. Provee de herramientas *off-the-shelf* para indexación, búsqueda y evaluación acorde con la metodología de referencia establecida en la iniciativa TREC.
- Lemur (Indri) (University of Massachusetts-Amherst y Carnegie-Mellon University 2012). Una herramienta diseñada inicialmente para la investigación en Modelos de Lenguaje para Recuperación de Información y desarrollado en C++.
- Zettair (Lucy) (RMIT University 2012). Un rápido motor de búsqueda sobre texto que provee en un solo binario de las capacidades de indexación y búsqueda. Fue desarrollado en C y provee de las facilidades habituales para su uso sobre las colecciones TREC.
- Lucene (Apache Foundation 2012) . Su desarrollo por parte de la Fundación Apache independiente de cualquier universidad le provee de un carácter más industrial. Está desarrollado en Java y tiene un amplio soporte en forma de

material de aprendizaje y contribuciones por parte de la comunidad.

Sobre las distintas herramientas realizamos un estudio pormenorizado de sus características atendiendo a tres macro-factores importantes de cara a su elección como herramienta para la docencia (i) naturaleza de la tecnología y su uso en el mundo académico y profesional (ii) funcionalidades y características (iii) restricciones de las herramientas y de los planes de estudio. A continuación enumeramos los distintos factores considerados en el análisis de características y el porqué de su importancia en el proceso de toma de decisión.

- Lenguaje de programación: es un aspecto crítico a la hora de valorar la idoneidad de la herramienta puesto que la familiaridad del alumno con el lenguaje determinará si es posible su uso en el contexto de los planes de estudio.
- Licencia: de cara al uso didáctico de la herramienta tener acceso al código fuente de la misma puede ser muy interesante como material de apoyo a la explicación de su funcionamiento, que la licencia nos lo permita será un factor determinante.
- Comunidad de usuarios: la cantidad de usuarios de la herramienta y su naturaleza facilitará la posibilidad de que los alumnos se nutran de su conocimiento y se integren en la misma.
- Paradigma de uso: aquellas herramientas cuyo uso no requiera conocer toda la estructura de clases interna y la recompilación de la misma para el uso por parte del alumno parecen más adecuadas para la docencia práctica. Sin duda de cara a la evaluación continua del alumnado será más adecuado trabajar contra librerías y que el trabajo del alumno se encuentre aislado y desacoplado.
- Tipos de documentos soportados: el soporte nativo de las librerías para múltiples formatos de archivos facilitará que el alumno experimente con la utilidad. También le será más sencillo imaginar las posibilidades que le puede brindar el uso de lo aprendido en su futuro profesional, obteniendo con esto una motivación extra en el proceso de aprendizaje.
- Modelos de indexación y recuperación: las posibilidades de poner en práctica lo aprendido en las clases magistrales de docencia expositiva reforzará el proceso de aprendizaje. Desde este punto de vista poder ver con ejemplos reales las ventajas y superioridades de unos modelos sobre otros será vital para la

compresión de conceptos y la motivación del aprendizaje.

- Formulación de consultas: la expresividad del lenguaje de consulta de las librerías es fundamental para introducir nuevos objetivos de aprendizaje prácticos en un nivel superior a través del estudio de las técnicas de procesado de las mismas.
- Integración con *crawlers*: estas utilidades que permiten el rastreo de la web y su descarga a local para su posterior indexación y búsqueda son cruciales en los buscadores web modernos. Es fácil imaginar que para la adquisición de un conocimiento integrado es deseable tener la posibilidad de experimentar con estas herramientas por parte del alumno.
- Facilidad de indexación y procesamiento distribuido: hoy en día y cada vez más, los ingentes volúmenes de información con los que se trata requieren de su procesamiento en más de una máquina. La facilidad que la herramienta ofrezca de cara al aprendizaje por parte del alumno de este tipo de paradigmas será determinante sobre todo en los niveles de máster.
- Fiabilidad: aunque una herramienta sea óptima para alcanzar los objetivos de aprendizaje necesitaremos pensar también en la utilidad del aprendizaje práctico en el futuro profesional del alumno. Desde este punto de vista preferiremos enseñar con herramientas fiables a las que el alumno les pueda sacar provecho una vez integrado en el tejido productivo.
- Madurez de la herramienta y uso en proyectos: preferiremos herramientas maduras con interfaces estables evitando así el constante cambio en los mismos que invaliden el conocimiento aprendido. Un factor determinante del grado de madurez de una herramienta será el uso de la misma por parte de proyectos reales.
- Soporte al aprendizaje: el disponer de material de aprendizaje abundante y de calidad para ayudar al alumno en su trabajo autónomo tal como libros, cursos o video-lecturas será determinante a la hora de decantarnos por una u otra herramienta de aprendizaje.

Resultados

Tabla 1. Comparación de las distintas características de las herramientas analizadas

	<i>Terrier</i>	<i>Lemur</i>	<i>Zettair</i>	<i>Lucene</i>
<i>Lenguaje</i>	Java	C++	C	Java
<i>Licencia</i>	<i>Mozilla Public License</i>	<i>BSD Like</i>	<i>BSD</i>	<i>Apache</i>
<i>Comunidad</i>	Comunidad pequeña	Comunidad pequeña	Comunidad pequeña	muyComunidad extensa
<i>Paradigma</i>	Fuentes, scripts, API	binarios,Fuentes, scripts, API	binarios,Fuentes, scripts	binarios,Fuentes, binarios, API
<i>Ficheros</i>	HTML, Texto plano, XML, PDF, PS	HTML, Texto plano, XML, PDF	HTML, Texto plano	HTML, Texto plano, PDF, DOC, ODT, etc.
<i>Modelos</i>	Repertorio amplio	Repertorio amplio	Repertorio amplio	Repertorio limitado
<i>Consultas</i>	Expresividad baja	Expresividad media	Expresividad baja	Expresividad alta
<i>Integración con crawler</i>	Integración baja	Integración baja	Integración baja	Integración alta
<i>Distribución</i>	Facilidad media	Facilidad media	Poca facilidad	Gran facilidad
<i>Fiabilidad</i>	Alta	Alta	Alta	Alta
<i>Madurez</i>	Alta	Alta	Alta	Muy Alta
<i>Soporte</i>	Bastante Material	Bastante Material	Poco material	Mucho material

Los resultados del estudio de los factores del análisis de características están resumidos en la Tabla 1. En la toma de decisiones sobre la elección tecnológica para la docencia práctica de Recuperación de Información debemos considerar adicionalmente ciertas consideraciones restricciones:

- Lucene es una plataforma que predomina en el ámbito industrial, siendo las otras tres usadas a nivel investigador pero marginalmente usadas en proyectos de producción.

- Los planes de estudio nos determinan la necesidad de adquirir las siguientes competencias

1. Grado: *Capacidad para conocer y desarrollar técnicas de aprendizaje computacional y diseñar e implementar aplicaciones y sistemas que las utilicen, incluyendo las dedicadas a extracción automática de información y conocimiento a partir de grandes volúmenes de datos.*
2. Máster: *Capacidad para aplicar métodos matemáticos, estadísticos y de inteligencia artificial para modelar, diseñar y desarrollar aplicaciones, servicios, sistemas inteligentes y sistemas basados en el conocimiento. Capacidad de comprender y saber aplicar el funcionamiento y organización*

de Internet, las tecnologías y protocolos de redes de nueva generación, los modelos de componentes, software intermediario y servicios

Discusión/Conclusiones

Como conclusión de los estudios realizados, considerando las restricciones y factores adicionales hemos concluido que las mejores herramientas para facilitar el ejercicio autónomo práctico del alumno del bagaje adquirido en las clases magistrales facilitando la evaluación y el alcance de las competencias serán:

- ^ Lucene para el Grado y Máster en Ingeniería Informática debido sobre todo a la necesidad de alcanzar unas competencias muy sesgadas hacia la aplicación industrial del conocimiento y también debido al gran soporte de material y comunidad para el trabajo del alumno.
- ^ Terrier para la docencia a los alumnos que se incorporen al programa de Doctorado en Computación, dado que su repertorio de modelos nos permitirá profundizar más en el conocimiento avanzado y dispone también de buenas fuentes de información y recursos para el trabajo autónomo del alumno.

Referencias

Baeza-Yates, R., Ribeiro-Neto, B. y Middleton, C. (2011). *Open Source Search Engines*. En Baeza-Yates, R., Ribeiro-Neto, B. (Eds.), *Modern Information Retrieval: the concepts and technology behind search* (pp. 737-754). Segunda Edición. Edimburgo: Pearson

Cleger-Tamayo, S., Figuerola, C.G. y Rodríguez-Cano, J.C. (2012). *Motores de Búsqueda de Código Abierto*. En Cacheda, F., Fernández-Luna, J. y Huete, J. (Eds.), *Recuperación de Información: Un enfoque práctico y multidisciplinar* (pp. 233-259). Madrid: Ra-Ma.

University of Glasgow (2012). *Terrier IR Platform*. Recuperado el 1 de mayo de 2012 de <http://www.terrier.org>

Apache Foundation (2012). *Lucene Project*. Recuperado el 1 de mayo de 2012 de <http://lucene.apache.org/>

University of Massachusetts-Amherst y Carnegie-Mellon Univerisy (2012). *Lemur Project*. Recuperado el 1 de mayo de 2012 de <http://www.lemurproject.org/>

RMIT University (2012). *Zettair an open source search engine*. Recuperado el 1 de mayo de 2012 de <http://www.seg.rmit.edu.au/zettair/index.html>