

Propuesta para la Creación de una Ontología sobre Departamentos Universitarios de Computación en Chile

Ernesto Krsulovic-Morales, Claudio Gutiérrez
{ekrsulov, cguetierr}@dcc.uchile.cl
Departamento de Ciencias de la Computación, Universidad de Chile
Blanco Encalada 2120, Santiago, Chile.
Fax: +56 2 6895531

Resumen

El artículo propone abrir una discusión sobre la necesidad de definir una ontología para el nivel educacional universitario en Chile. Se discute la importancia de tener tal ontología y sus posibles aplicaciones y perspectivas para apoyar la infraestructura educacional en Chile. Se presenta un prototipo que los autores han usado en otras aplicaciones. Se aprovecha, además, de introducir el tema de ontologías, metodologías para el desarrollo éstas.

1. Introducción

¿Que es una ontología? Es una teoría filosófica sobre la naturaleza de la existencia. Los investigadores de Inteligencia Artificial (especialmente representación y adquisición del conocimiento) reencarnaron este término en su propia jerga para expresar "un entendimiento compartido y común sobre un dominio que puede ser comunicado entre personas y aplicaciones" [1]. Visto de una manera simple cualquier esquema puede ser calificado como una ontología. Sin embargo, una ontología típica tiene una taxonomía que define clases y sus relaciones, y un conjunto de reglas de inferencia.

Iniciativas a nivel mundial En el ámbito educacional existen diversos esquemas, partiendo por los modelos de bases de datos relacionales y modelos de clases UML de aplicaciones, que son manejados por las propias instituciones u organismos gubernamentales, hasta iniciativas abiertas en tecnologías tan diversas como XML, EDI [2] y LDAP [3]. En XML existen iniciativas orientadas principalmente a *learning objects* [4]. En el sector de la educación primaria está el proyecto SIF [5] orientado a soportar la interoperabilidad entre sistemas de administración educacional. El proyecto de Registro Educacional [6] donde se definen esquemas para la transcripción del historial estudiantil utilizando EDI. Para la educación superior y ligado al proyecto *internet2* existe el proyecto eduPerson [7] donde se especifican esquemas LDAP para la descripción de personas y organizaciones. Tres proyectos puntuales en el ámbito de la Web semántica que describen departamentos universitarios de ciencias de la computación, el europeo AKT [8] y los estadounidenses del proyecto Mangrove [9] y SHOE [10].

Iniciativas a nivel nacional El proyecto SIEL (Sistema integrado de información universitaria en línea) [11], dirigido por REUNA y en el cual participan numerosas universidades chilenas, tiene como proposito establecer un sistema distribuido de contenidos para los temas relevantes que producen las universidades. Este proyecto se enmarca dentro del trabajo internacional de ja-sig y su producto uPortal. Como parte de las áreas de trabajo en el proyecto SIEL destacamos la de estandarización, donde han desarrollado metadatos para diversos objetos de información como: universidad o institución, plan de estudio, programa o carrera, curriculum, asignaturas, directorio de personas y dirección. Por su parte, CONICYT con su incipiente proyecto SICTI [12], se propone tareas similares para la descripción de información relevante para la comunidad científica, con datos sobre currículo, proyectos, instituciones y resultados.

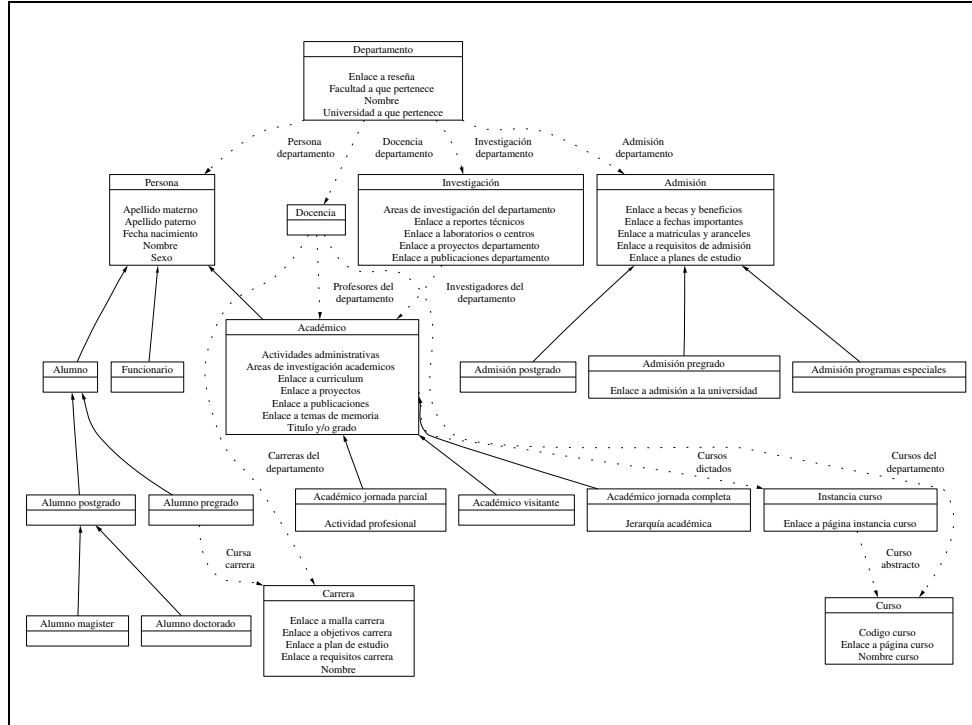


Figura 1: Sección de la ontología para el departamento.

Web semántica La Web semántica es la próxima generación de la Web, propuesta por el W3C (Consortio de la Web), donde se aumenta el nivel de granularidad al enlazar directamente datos en lugar de páginas de información. Pilares fundamentales son la sintaxis XML [13] que permite estructurar los datos, el modelo de metadatos semánticos RDF [14] que permite especificar el significado de los datos y la descripción de ontologías.

OWL El primer lenguaje de ontologías ligados con la iniciativa de la Web Semántica fue RDF Schema [15], el que ha sido denominado como un lenguaje *liviano* debido a sus limitadas funcionalidades, sin embargo RDFS es la base para las siguientes iniciativas, como son la norteamericana DAML y la europea OIL que añaden mayores características como lenguajes de ontología, luego DAML y OIL se unieron y formaron DAML+OIL [16], y enviaron la especificación al W3C para su estandarización, donde fue renombrado como OWL [17].

OWL permite al igual que RDFS: establecer una jerarquía de clases, al definir relaciones de clases y subclases; definir propiedades; establecer asociaciones de clases, mediante dominio y rango de propiedades; y crear instancias de clases. También presenta funcionalidades adicionales en comparación a RDFS, con OWL es posible: establecer sinónimos, al definir clases y propiedades equivalentes; imponer un mayor número de restricciones a los valores de propiedades; y definir axiomas mediante lógica descriptiva que permiten realizar inferencia de la información.

2. Ontología para departamentos de computación (ODC)

Importancia de una ODC Las ontologías son herramientas claves para fundar la interoperabilidad en ciertas comunidades en la Web. La importancia de tener una tal ODC en Chile es múltiple. Nombramos algunos puntos destacables:

1. Soporte a una comunidad, sea ésta temática (área de bases datos, lenguajes, ingeniería de software, etc.), de roles (comunidades de alumnos, de profesores, de regiones), y de organizaciones (todas las carreras con ciertas características, etc.).

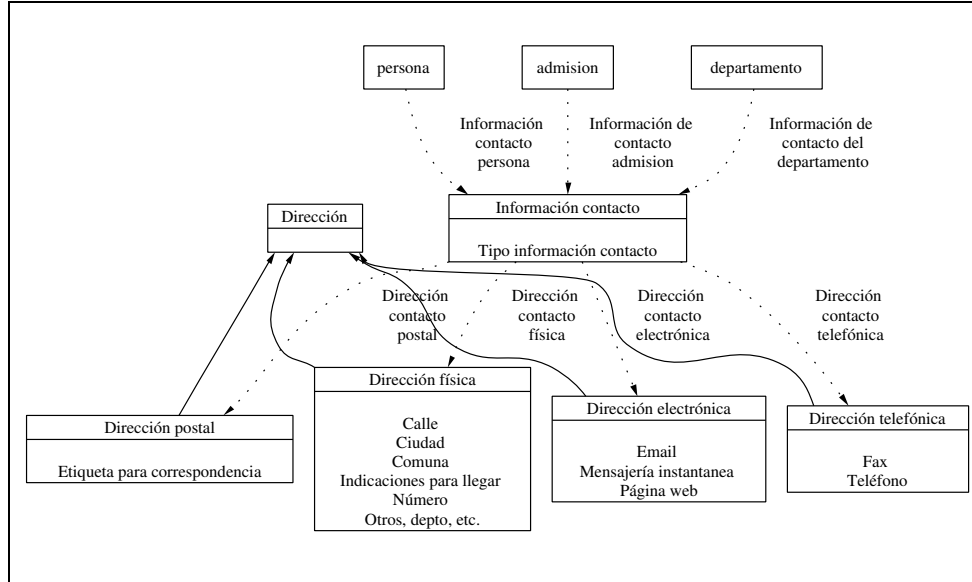


Figura 2: Sección de la ontología para información de contacto y áreas de investigación.

2. Acceso a datos para usarlos en otras aplicaciones. Uno de los grandes problemas de la Web actual es que, a pesar de haber una gran cantidad de información disponible para el mundo educacional, mucha de ésta no es “usable” (integrable con otras aplicaciones) de una forma distinta a la simple visualización en un página Web.
3. Clasificación de materiales educacionales: la ontología es un mapa conceptual que facilita y estandariza la clasificación de contenidos. Luego la búsqueda y acceso de éstos se potencia.
4. Razonar sobre información educacional: Una ontología no sólo permite recuperar información, sino razonar y hacer inferencias sobre ella.

Propuesta Una ontología es esencialmente un vocabulario estandarizado. Debe tener un nivel de aprobación social y una discusión con los actores involucrados.

Para comenzar la discusión de una tal propuesta, proponemos partir por un desarrollo inicial que servirá básicamente para fijar algunos puntos de partida básicos. Los autores han venido trabajando con un modelo desarrollado para una aplicación particular, que ha considerado algunas experiencias internacionales como el proyecto europeo AKT [8] y los estadounidenses Mangrove [9] y SHOE [10].

La formalización y modificación de la ontología, aunque aparentemente puntos secundarios, en el caso de la Web Semántica y de la interoperabilidad en la Web es clave. Como se mencionó en la introducción, el Consorcio de la Web recomienda una tal formalización, el *Lenguaje de Ontologías para la Web, OWL*. Una versión OWL de nuestra ontología está disponible en <http://www.metadatos.cl/odc/>.

Versión inicial Durante la creación de la ontología se utilizó como guía el tutorial “*Ontology Development 101: A Guide to Creating Your First Ontology*” [18], siguiendo un esquema *top-down*, se rescatan los siguientes principios: restringir el alcance de la ontología antes de comenzar su creación; y tener en cuenta que las subclases son siempre especializaciones de una superclase, se trata de una relación *is-a*. Para encontrar los términos relevantes, se buscaron los patrones comunes entre los datos que se presentan en los sitios Web de los departamentos de ciencias de la computación en Chile, además ayudó el poseer un *background* propio ya que también somos parte de un departamento. Para restringir el alcance de la ontología, sin afectar su extensibilidad, muchos de las propiedades son enlaces a páginas, las que tienen el prefijo “enlace a” (por ejemplo, ver propiedades de clase investigación). En el caso de los cursos no se creó una ontología dedicada, sino que se deja el camino abierto al mercado específico, mediante una propiedad que enlaza a la página del curso, esto tiene mucho sentido si se toma en cuenta la gran cantidad de ontologías especializadas en el

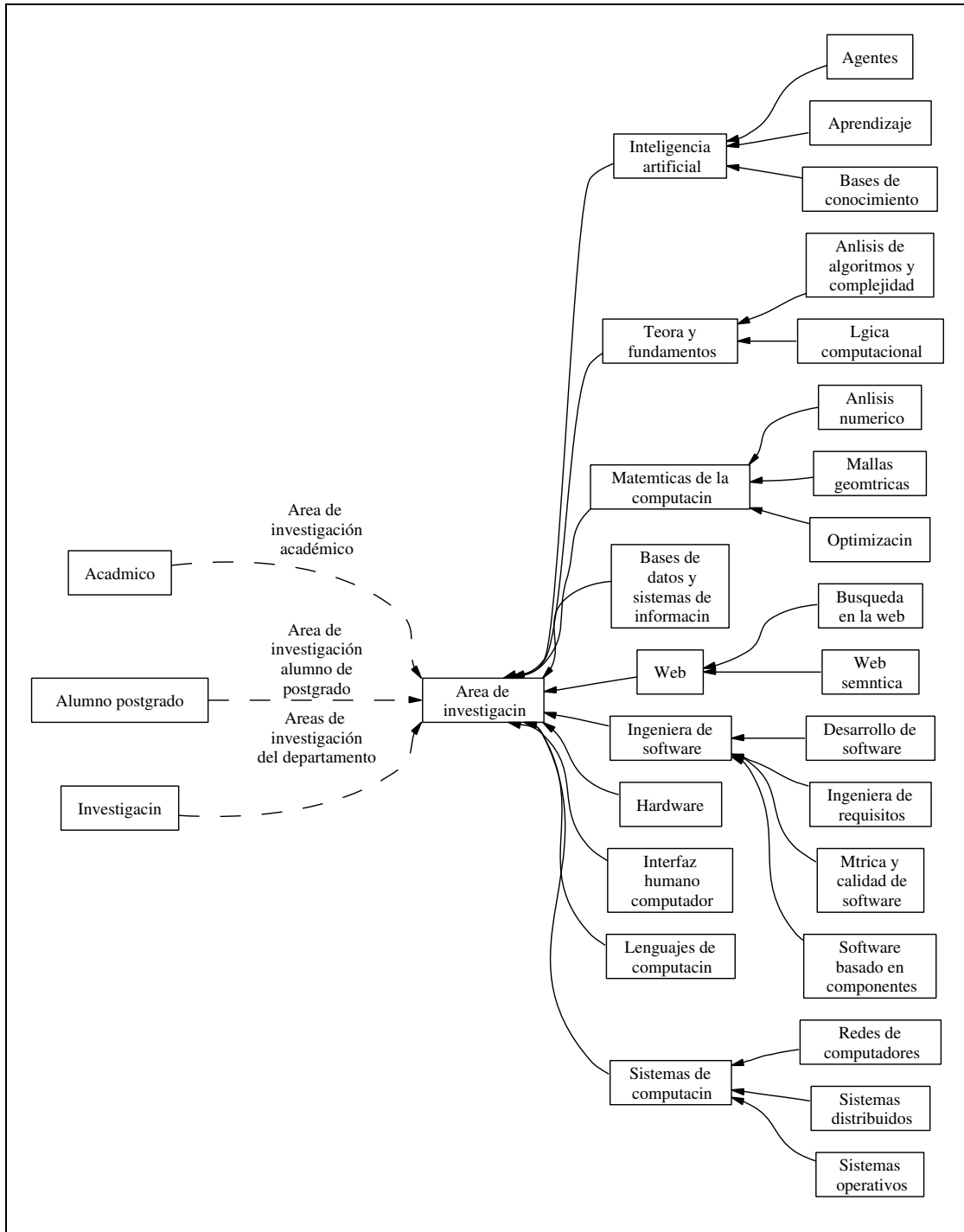


Figura 3: Sección de la ontología para información de contacto y áreas de investigación.

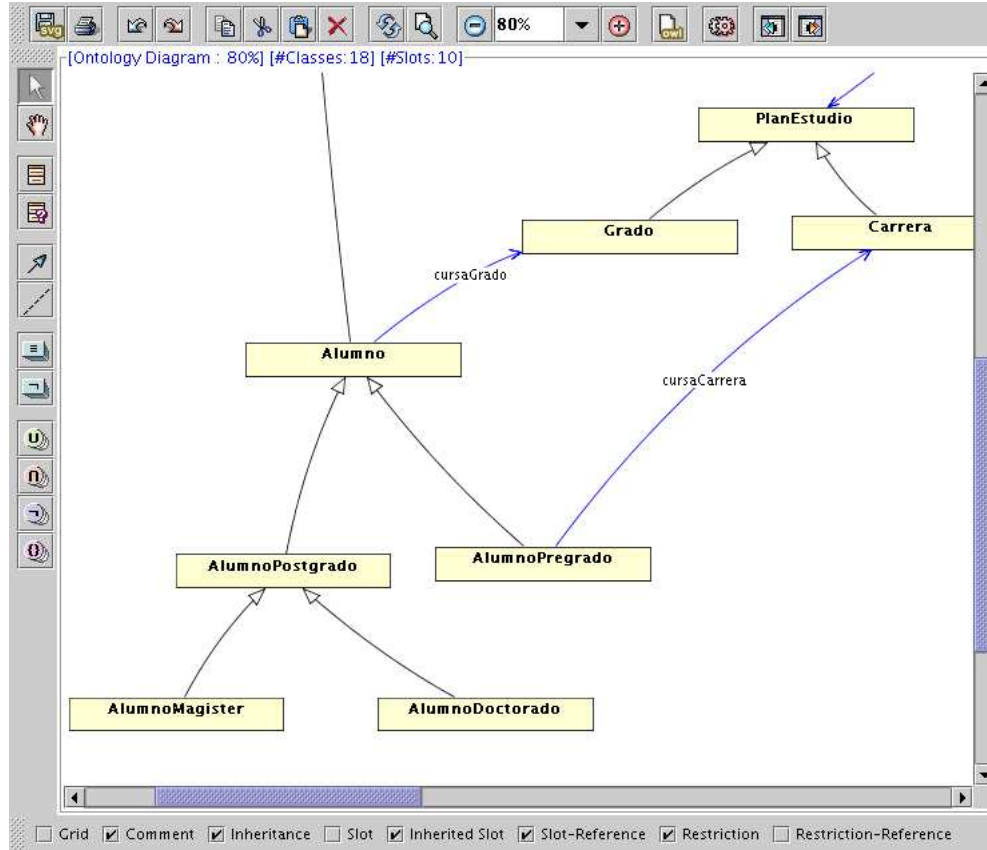


Figura 4: Cambio a la ontología en ezOWL

ámbito de *e-learning*. El resultado obtenido es una ontología que podemos dividir en tres secciones independientes: departamento, información de contacto, áreas de investigación.

Departamento es el modelo conceptual para un departamento universitario, con clases para un *departamento* que se relaciona con *investigación*, *admisión* y *docencia*. Además contiene personas, las que pueden ser diferencias de acuerdo a su rol, como son académicos, alumnos y funcionarios. Ver figura 1.

Información de contacto esta sección de la ontología es de carácter general y contiene relaciones con las siguientes subclases de la clase dirección: dirección electrónica, dirección física, dirección postal, dirección telefónica. Ver figura 3.

Áreas de investigación presenta una clasificación con clases y subclases de áreas de interés, y está basada en las áreas en que se dividen las conferencias de la Sociedad Chilena de Ciencias de la Computación. Ver figura 3.

Evolución de la ontología Para ilustrar la evolución de la ontología presentamos una posible modificación sobre la propuesta presentada en la figura 1. Pensemos en agregar una clase denominada *Grado*, vemos que tanto un grado como una carrera comparten iguales propiedades, por lo que podemos generalizar en una super-clase llamada Plan de Estudio. Con este cambio obtenemos la sección de interés presentada en la figura 4, que muestra el área de trabajo del *plug-in ezOwl* [19] para *protégé* [20].

3. Conclusiones

Hemos presentado una propuesta de ontología que creemos puede mejorar con la ayuda de los miembros de la comunidad de ciencias de la computación a los cuales extendemos la invitación a participar en la creación de la ODC.

Desarrollo de la ontología asistida por herramientas Al tener herramientas que presentan un resultado concreto a partir de una ontología, se produce una *feedback* directo y es ahí donde comienza el interés de realizar modificaciones a la ontología y añadir nuevas clases y propiedades. Para ello proponemos el uso de un anuario de departamentos universitarios y una herramienta de marcado¹ que faciliten el acercamiento a la ontología y gatillen los comentarios y sugerencias desde la comunidad.

Referencias

- [1] Thomas R. Gruber. A Translation Approach to Portable Ontology Specifications. Technical Report KSL 92-71, Knowledge Systems Laboratory Computer Science Department, Stanford University, Stanford, April 1993.
- [2] United Nations Directories for Electronic Data Interchange for Administration, Commerce and Transport. <http://www.unece.org/trade/untidd/welcome.htm>.
- [3] M. Wahl, T. Howes, and S. Kille. Lightweight Directory Access Protocol (v3). <http://www.ietf.org/rfc/rfc2251.txt>, Dec 1997.
- [4] David A. Wiley. *The Instructional Use of Learning Objects*, chapter Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy. Agency for Instructional Technology and the Association for Educational Communications and Technology, 2000.
- [5] Schools Interoperability Framework. <http://www.sifinfo.org/>.
- [6] EDI: SPEEDE/ExPRESS. <http://nces.ed.gov/edi/speedeExp.asp>.
- [7] eduPerson Object Class. <http://www.educause.edu/eduperson/>.
- [8] AKT. Hyphen.info, an information source for UK Researchers. <http://www.hyphen.info/>.
- [9] Department of Computer Science & Engineering, University of Washington. Mangrove: An Evolutionary Approach to the Semantic Web. <http://www.cs.washington.edu/research/semweb/>.
- [10] Parallel Understanding Systems Group, Department of Computer Science, University of Maryland at College Park. SHOE and DAML. <http://www.cs.umd.edu/projects/plus/DAML/>.
- [11] Sistema Integrado de Información Universitaria en Línea, SIEL. <http://www.uportal.cl/>.
- [12] Sistema de Información en Ciencia, Tecnología e Innovación, SICTI. <http://www.sicticurricula.cl/>.
- [13] Tim Bray, Jean Paoli, C. M. Sperberg-McQueen, and Eve Maler. Extensible Markup Language (XML) 1.0 (Second Edition). <http://www.w3.org/TR/2000/REC-xml-20001006>, Jan 2001.
- [14] Ora Lassila and Ralph Swick. Resource Description Framework (RDF) Model and Syntax Specification. <http://www.w3.org/TR/1999/REC-rdf-syntax-19990222/>, Feb 1999.
- [15] Dan Brickley and R.V. Guha. RDF Vocabulary Description Language 1.0: RDF Schema. <http://www.w3.org/TR/2002/WD-rdf-schema-20020430/>, Apr 2002.
- [16] Ian Horrocks, Frank van Harmelen, and Peter Patel-Schneider. DAML+OIL. <http://www.daml.org/2001/03/daml+oil-index.html>, March 2001.
- [17] Mike Dean, Dan Connolly, Frank van Harmelen, James Hendler, Ian Horrocks, Deborah L. McGuinness, Peter F. Patel-Schneider, and Lynn Andrea Stein. OWL Web Ontology Language 1.0 Reference. <http://www.w3.org/TR/owl-ref/>, July 2002.
- [18] Natalya Fridman Noy and Deborah L. McGuinness. Ontology Development 101: A Guide to Creating Your First Ontology. Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880, Stanford Knowledge Systems Laboratory, Mar 2001.
- [19] ezowl. <http://iweb.etri.re.kr/ezowl/>.
- [20] Natalya F. Noy, Michael Sintek, Stefan Decker, Monica Crubézy, Ray W. Ferguson, and Mark A. Musen. Creating Semantic Web Contents with Protégé-2000. *IEEE Intelligent Systems Journal*, 16(2):60–71, Mar 2001.

¹Disponibles en: <http://purl.org/net/depmark>