

Asignación y resolución de identificadores para un repositorio de objetos de aprendizaje basado en LOM

Roberto Javier Godoy[†], Hugo Minni[‡]

Facultad de Ingeniería y Ciencias Hídricas
Universidad Nacional del Litoral
tel: +54-342-457-5234; fax: +54-342-457-5224
[†] rjgodoy@fich.unl.edu.ar [‡] hminni@fich.unl.edu.ar

Resumen En el presente trabajo se desarrolla un análisis sobre la asignación de identificadores en un repositorio institucional de objetos de aprendizaje. Este repositorio está siendo diseñado en la Universidad Nacional del Litoral, de acuerdo al estándar IEEE 1484.12.1 Learning Object Metadata. En el contexto de aplicación, en el cual los contenidos pueden ser generados de manera descentralizada, se requiere un mecanismo que garantice la unicidad y persistencia de los identificadores. Se explica porqué los URL no resultan adecuados, y porqué el espacio formal de nombres URN `fdc` se ajusta a los requerimientos. Además, se mencionan algunos protocolos que podrían implementarse a fin de resolver tales identificadores en el sistema.

1. Introducción

Un repositorio de conocimiento es una herramienta auxiliar muy útil en la educación universitaria. Estos repositorios proveen un sistema integral para coordinar material educativo relacionado bajo la forma de objetos de aprendizaje (OA). Los OA son “recursos digitales que pueden reutilizarse para ayudar al aprendizaje” [1].

Los autores actualmente están desarrollando un repositorio de OA en la Universidad Nacional del Litoral de Santa Fe [2]. Esta Casa de Estudios cuenta con más de 80 carreras de grado y postgrado, las cuales se desarrollan en sus 9 Facultades y 4 Escuelas superiores. Un diseño robusto debería prever la posibilidad de que en un futuro estas y otras dependencias de la Universidad generen repositorios de OA propios, de manera descentralizada: cada Facultad, departamento o grupo de investigación podría mantener su propio repositorio, formando un sistema distribuido con repositorios conectados entre sí, pero capaces de operar en forma autónoma.

En un repositorio que adhiere al estándar IEEE Learning Object Metadata (LOM) [3], cada OA y cada instancia de metadatos se identifica mediante uno o más elementos `identifier`. Un elemento `Identifier` está compuesto por un elemento `Catalog` que especifica el “nombre o designación del esquema de catálogo” (e.g. ISBN, ARIADNE, URI), y un elemento `Entry` que toma un valor dentro del catálogo indicado. Si bien puede elegirse libremente qué catálogo utilizar, debería *recomendarse* un catálogo y/o un mecanismo de asignación para evitar identificadores duplicados.

Si institucionalmente se permite la generación descentralizada de contenidos sin una autoridad central de control, la asignación de identificadores no es un problema menor: un mecanismo inadecuado de asignación de identificadores podría causar que recursos distintos se asocien con el mismo identificador.

2. Identificadores

URI es uno de los catálogos que puede utilizarse para formar un identificador LOM. “Una URI (Universal Resource Identifier) puede clasificarse como un localizador, un nombre, o ambos. El término ‘Universal Resource Locator’ (URL) designa al subconjunto de URIs que, además de identificar un recurso, proveen una manera de ubicar el recurso describiendo su mecanismo primario de acceso (e.g. su dirección de red). El término ‘Uniform Resource Name’ (URN) ha sido históricamente usado para referir a URIs bajo el esquema ‘urn’ [4], que deben permanecer globalmente únicas y persistentes aún cuando el recurso deje de existir o se vuelva no disponible (...)” [5]

En algunos casos especiales el material que se almacena en el repositorio ya tiene un identificador asignado en otro catálogo (e.g. ISBN) de donde se deriva directamente un valor para `Identifier.Entry`. Otra alternativa (que se ejemplifica en el estándar LOM) consiste en relacionar cada objeto con un URL (e.g. `http:` o `ftp:`). En este último caso la resolución es muy sencilla, pero utilizar URLs como identificadores únicos no es conveniente por varias razones:

- Debe ser posible modificar la ubicación de un recurso, manteniendo sin cambios los identificadores.
- El esquema base LOM proporciona un elemento específico (`Technical.Location`) para indicar la forma de acceder al recurso.
- Un OA puede marcarse como “no disponible” (i.e. `LifeCycle.Status = unavailable`), en tal caso el objeto no tendría “ubicación” y por lo tanto no podría describirse mediante un URL.
- Los recursos del sistema pueden estar lógicamente relacionados entre sí (e.g. varios OA sobre un mismo tema) aunque su almacenamiento físico sea diferente. Un URL podría no reflejar este hecho del mismo modo que un URN.

Por su persistencia, un esquema de nombres como `urn:` resulta preferible ante los esquemas URL. Con esta estrategia todos los identificadores pertenecerían a un mismo catálogo. No obstante, puesto que el estándar LOM permite referenciar una entidad mediante más de un identificador, podría incluirse una entrada bajo el catálogo URI, y entradas bajo otros catálogos, como así también identificadores URI adicionales del mismo esquema o de esquemas distintos.

En el esquema `urn:` se definen *espacios de nombre*, con mecanismos que garantizan la unicidad de cada identificador en el sistema. Actualmente existen 29 espacios formales de nombre URN registrados [6]. Algunos de ellos, como IETF (Internet Engineering Task Force) y OASIS (Organization for the Advancement of Structured Information Standards), sirven para identificar documentos generados por las organizaciones homónimas; mientras que otros, como OID (object identifier) e IVIS (Latvian National Government Integration Project), tienen propósitos más generales.

La asignación de identificadores en un espacio de nombres URN normalmente depende de una autoridad central, que puede delegar parcialmente en otras autoridades. En estos espacios no puede asignarse un identificador válido sin consentimiento de una autoridad de nombres. Sin embargo, existen otros espacios que pueden considerarse “abiertos”: `publicid` (public identifier) [7][8], `UUID` (unique identifier) [9] y `fdc` (federated content) [10]. Estos últimos espacios no se necesitan registro o utilizan un valor que ya fue asignado con otro propósito (e.g. IEEE 802 MAC Address y dominios de Internet).

2.1. Externalización de vCard

El formato vCard 3.0 [11] es utilizado por LOM para representar información sobre los individuos que contribuyen en el ciclo de vida de un OA. En un trabajo anterior, los autores propusieron un mecanismo de “externalización de vCard” [2] utilizando la propiedad `source` para referenciar el URI de un recurso que provee información adicional.

Realizando sistemáticamente esta externalización es posible centralizar el manejo de la información personal y reflejar globalmente sus cambios. Originalmente se había sugerido usar un esquema como `ldap:` (que es un URL) pero los objetivos se cumplen mejor si los identificadores son persistentes, por lo que aquí también conviene utilizar un esquema de nombres.

Por otra parte, el valor de algunos atributos de vCard puede alternativamente ser un archivo embebido (codificado en base-64) o una referencia a un recurso externo mediante URI. Aquí también es posible utilizar URN para tales referencias, en lugar de simplemente describir la ubicación del archivo mediante URL.

3. Espacios de nombre URN

Para comparar distintos espacios de nombre se tuvieron en cuenta los siguientes criterios: unicidad y persistencia de los identificadores, facilidad para resolverlos globalmente y mecanismos de asignación.

Los identificadores públicos pueden ser formales (FPI, Formal Public Identifier) o no formales. La utilización de propietarios registrados es una práctica infrecuente e implica un determinado costo. Por otra parte, la utilización de identificadores públicos no formales o FPI con propietarios no registrados no ofrece garantías de unicidad ni persistencia. Existe una alternativa intermedia (+//IDN) que automáticamente define propietarios FPI para cada nombre de dominio de Internet, pero presenta las debilidades asociadas a la persistencia de los nombres de dominio.

Otro inconveniente de los FPI es que incluyen información ya presente en los metadatos LOM (idioma, formato, restricciones) por lo que el identificador puede quedar *inconsistente* si se modifican los metadatos.

En el espacio de nombres `urn:fdc` cada organización se distingue por la concatenación de un nombre de dominio y una fecha en la que poseía ese dominio. Así se consiguen identificadores persistentes incluso cuando un dominio deje de existir o sea cedido. No obstante, considerando que los repositorios de OA serán manejados por instituciones estables, cabe esperar que los nombres de dominio sufran pocas modificaciones.

En ambas alternativas una parte del identificador puede ser asignada libremente por la organización. Por ejemplo, para garantizar la unicidad en un repositorio institucional de LOM, podrían asignarse identificadores parciales para departamentos, cátedras, docentes, etc.

También se analizaron los UUID, que resuelven los problemas de persistencia, asignación y unicidad, pero resultan inadecuados para la interacción entre repositorios debido a la carencia de un mecanismo de resolución global.

4. Resolución de URN

En OASIS Technical Resolution 9401:1997 [12] se define un formato simple de catálogo de entidades para ser usado principalmente con aplicaciones SGML. El catálogo es una lista

ordenada de entradas que relacionan identificadores externos y nombres de entidad con identificadores de objetos del sistema (e.g., nombres de archivos). Esta técnica fue aplicada a la resolución de URN en el espacio de nombres *oasis* [13], y podría aplicarse como estrategia local en cualquier espacio de nombres.

Al implementar la resolución en un ámbito más extenso, surge la necesidad de definir *servicios de resolución*, y que permiten obtener el recurso identificado o convertir un URI en un URL o URC (Uniform Resource Characteristics) [14]. Estos servicios son I2L (URI a URL), I2C (URI a URC), I2N (URI a URN), I2R (URI a recurso), sus versiones colectivas (e.g. I2Ls, URI a URLs) y el servicio I=I (¿es URI igual a URI?) que consulta sobre equivalencia funcional.

Para que la resolución de un determinado URI sea efectiva, debe poder ubicarse el servidor que provea resolución para el mismo. A este fin puede utilizarse el Dynamic Delegation Discovery System (DDDS) [15], que define mecanismos para:

- Aplicar a una cadena reglas de transformación (expresiones regulares) obtenidas dinámicamente.
- Establecer mediante DNS (Domain Name System) una base de datos distribuida de reglas codificadas mediante Naming Authority Pointer (NAPTR) Resource Record (RR).
- Encontrar un servidor con autoridad (authoritative server).

5. Conclusión y continuación del trabajo

Se presenta una síntesis del análisis realizado para la elección de identificadores en un repositorio de OA, en un contexto de generación descentralizada de contenidos.

Como parte del diseño del repositorio de OA de la Universidad Nacional del Litoral, esto cobra gran importancia ya que no solo se aplica a los identificadores de LOM sino también a la externalización de información personal bajo este estándar.

Los URL no son adecuados para esta finalidad, pues son intrínsecamente no-persistentes. Otros catálogos, esquemas y espacios de nombres, que podrían ser aplicados en todos los casos, deben ser descartados por su difícil resolución o por no ofrecer suficientes garantías de unicidad.

Se encontró que —aunque con algunas observaciones— el espacio de nombres *urn:fdc* cumple con los requerimientos. No obstante, es necesario analizar mejor la estructura del espacio de nombres ampliando y definiendo mejor las consideraciones aquí efectuadas. También debe profundizarse el estudio de las especificaciones y protocolos, con el objetivo de encontrar mecanismos sencillos de acceso por parte de otras aplicaciones.

Para la resolución de identificadores se realizó un prototipo de servidor THHTTP integrado con la base de metadatos. Actualmente se está en proceso de incorporar DDDS (con reglas de transformación codificadas en NAPTR).

Referencias

1. WILEY D. *Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy*. 2000. <http://reusability.org/read/chapters/wiley.doc> (accedido en 2007-03).
2. R. J. GODOY, H. MINNI, G. ZARZA, Y H. LOYARTE. Design Criteria for the Development of an Institutional Learning Object Repository. En *Proceedings of XII Congreso Argentino de Ciencias de la Computación*, San Luis, 2006.
3. LEARNING TECHNOLOGY STANDARDS COMMITTEE. IEEE Standard for Learning Object Metadata. IEEE Standard 1484.12.1, Institute of Electrical and Electronics Engineers, New York, 2002.

4. R. MOATS. URN Syntax. RFC 2141, Internet Engineering Task Force, Mayo 1997.
5. T. BERNERS-LEE, R. FIELDING, Y L. MASINTER. Uniform Resource Identifier (URI): Generic Syntax. RFC 3986, Internet Engineering Task Force, Enero 2005.
6. INTERNET ASSIGNED NUMBERS AUTHORITY. Official IANA Registry of URN Namespaces, Enero 2007. <http://www.iana.org/assignments/urn-namespaces> (accedido en 2007-03).
7. N. WALSH, J. COWAN, Y P. GROSSO. A URN Namespace for Public Identifiers. RFC 3151, Internet Engineering Task Force, Agosto 2001.
8. ISO (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION). Information processing — Text and Office Systems — Standard Generalized Markup Language (SGML). ISO Standard ISO 8879:1986(E), International Organization for Standardization, Geneva, Switzerland, Octubre 1986.
9. PAUL J. LEACH, MICHAEL MEALLING, Y RICH SALZ. A Universally Unique Identifier (UUID) URN Namespace. RFC 4122, Internet Engineering Task Force, Julio 2005.
10. DAVE TESSMAN. A Uniform Resource Name (URN) Namespace for Federated Content. RFC 4198, Internet Engineering Task Force, Noviembre 2005.
11. F. DAWSON Y T. HOWES. vCard MIME Directory Profile. RFC 2426, Internet Engineering Task Force, Septiembre 1998.
12. P. GROSSO. Entity Management: OASIS Technical Resolution 9401:1997 (Amendment 2 to TR 9401), Septiembre 1997. <http://www.oasis-open.org/html/tr9401.html> (accedido en: 2007-03).
13. K. BEST Y N. WALSH. A URN Namespace for OASIS. RFC 3121, Internet Engineering Task Force, Junio 2001.
14. M. MEALLING Y R. DANIEL. URI Resolution Services Necessary for URN Resolution. RFC 2483, Internet Engineering Task Force, Enero 1999.
15. M. MEALLING. Dynamic Delegation Discovery System (DDDS). RFC 3401–3405, Internet Engineering Task Force, Octubre 2002.