

Evaluación de la calidad de metadatos en repositorios digitales de objetos de aprendizaje

Valentina Tabares Morales

Administración de Sistemas Informáticos
Universidad Nacional de Colombia-
Sede Manizales. Estudiante de Maestría
en Ingeniería de Sistemas. Universidad
Nacional de Colombia-Sede Medellín
vtabaresm@unal.edu.co
Manizales –Colombia

Néstor Darío Duque Méndez

Doctorado en Ingeniería
Universidad Nacional de Colombia-Sede
Medellín. Profesor Asociado
Facultad de Administración
Universidad Nacional de Colombia
ndduqueme@unal.edu.co
Manizales –Colombia

Julián Moreno Cadavid

Doctorado en Ingeniería
Universidad Nacional de Colombia-Sede
Medellín. Profesor Asistente
Facultad de Minas
Universidad Nacional de Colombia
jmoreno1@unal.edu.co
Medellín –Colombia

Demetrio Arturo Ovalle Carranza

Doctorat en Informatique
Université Joseph Fourier–Grenoble–
Francia. Profesor Titular
Facultad de Minas
Universidad Nacional de Colombia
dovalle@unal.edu.co
Medellín –Colombia

Rosa Maria Vicari

Doutorado em Engenharia Eletrônica e
Computadores
Universidade de Coimbra–Portugal
Professora Titular
Instituto de Informática
Universidade Federal do
Rio Grande do Sul
rosa@inf.ufrgs.br
Porto Alegre –Brasil

Resumen

Los recursos educativos almacenados en repositorios digitales requieren de metadatos que los describan adecuadamente y son fundamentales en procesos de búsqueda y recuperación. El objetivo de este artículo es realizar una aproximación al cálculo y evaluación de métricas que permitan determinar el nivel de calidad del etiquetado de recursos digitales, específicamente de Objetos de Aprendizaje (OAs). El enfoque adoptado establece cómo evaluar los recursos teniendo en cuenta tres dimensiones: completitud, consistencia y coherencia, con base en el estándar IEEE LOM. Con el fin de validar los resultados obtenidos al aplicar las métricas se realizaron pruebas experimentales con OAs sintéticos y reales, concluyendo que las mismas permiten evaluar la calidad de los metadatos involucrados. De esta manera, la propuesta presentada se convierte en un insumo relevante para procesos de evaluación automática de metadatos, así como para la administración de repositorios y procesos de recomendación de OAs.

Palabras Clave: recursos digitales, objetos de aprendizaje, repositorios, metadatos, métricas de calidad

Assessing Metadata Quality in Digital Repositories of Learning Objects

Abstract

Educational digital resources stored in repositories require metadata in order to describe them properly, and are fundamental for searching and retrieval processes. The aim of this paper is to present an approach for the calculation of metrics that determine the quality level of labeling digital resources, specifically Learning Objects (LOs). We describe how to evaluate resources considering three dimensions such as completeness, consistency and coherence, based on the standard IEEE LOM. In order to validate the results that such metrics exhibit, some experimental tests were performed with synthetic and real OAs, concluding that they allows evaluating

Cómo citar este artículo: TABARES MORALES, Valentina; DUQUE MÉNDEZ, Néstor Darío; MORENO CADAVID, Julián; OVALLE CARRANZA, Demetrio Arturo y VICARI, Rosa Maria. Evaluación de la calidad de metadatos en repositorios digitales de objetos de aprendizaje. Revista Interamericana de Bibliotecología 2013, vol. 36, n° 3, pp. 183-195.

Recibido: 2013-03-14 / **Aceptado:** 2013-10-16

the quality of the metadata involved. This way, the proposal presented turns to be a plausible and relevant input for the automatic evaluation of metadata, as well as for the LOs repository management and recommendation processes.

Keywords: digital resources, learning objects, repositories, metadata, quality metrics

1. Introducción

De acuerdo con la definición formal del International Standard Bibliographic Description for Electronic Resources (ISBD, 2004), se entiende por recurso digital todo material codificado para ser manipulado por un computador y consultado de manera directa o por acceso electrónico remoto. Los recursos digitales en línea son cada día más utilizados en las bibliotecas y en otros espacios en los que se tiene acceso a internet. Tienen como objetivo, al igual que los recursos denominados tradicionales, atender necesidades y preocupaciones de los usuarios, facilitando la obtención de información pertinente, precisa y relevante de manera ágil y oportuna (Hubp, 2004).

Un tipo específico de recurso digital son los Objetos de Aprendizaje (OAs), que son entidades generalmente entregadas a través de internet y están diseñadas con el fin de que sean utilizadas y reutilizadas en múltiples contextos educativos (Learning Technology Standards Committee, 2002; Wiley, 2001). Una de las principales características de los OAs es que contienen metadatos que los describen facilitando su búsqueda y recuperación, lo cual permite la reutilización de estos recursos en diferentes contextos educativos, lo cual es uno de sus mayores retos (Bez, et al., 2010; Morales, et al., 2005). Adicionalmente, estos deben responder a las necesidades de los usuarios para los que fueron diseñados (Duque 2009; Motelet, Baloian & Pino 2006). Más específicamente, dichos metadatos permiten identificar cada OA, conocer cómo, dónde y por quién fue desarrollado, cuál es la población objetivo a la que va dirigido, su aplicación, interactividad, características técnicas y otro tipo de información relevante, que ayuda a entender su contenido. El principal objetivo de los metadatos es facilitar la búsqueda, evaluación, recuperación y uso de los OAs.

Con el fin de facilitar la interoperabilidad entre diferentes sistemas informáticos que contengan OAs, se han

propuesto diversos estándares para estructurar sus metadatos. Algunos de los más conocidos y utilizados son los siguientes:

- **Dublin Core:** Propuesto por la DCMI (Dublin Core Metadata Initiative) para catalogar recursos digitales en general. Establece solamente 15 campos, sin embargo existen algunas iniciativas que pretenden extenderlo. Es ampliamente utilizado en el contexto de los OAs (Weibel & Koch; 2000).
- **LOM:** Como su nombre lo indica (LOM es el acrónimo inglés de *Learning Object Metadata*) este es un estándar específico para OAs y es internacionalmente reconocido. Fue propuesto por el Comité de Estándares de Tecnologías Educativas de la IEEE y diseñado como un modelo jerárquico que contiene alrededor de 50 campos de metadatos agrupados en nueve categorías (Learning Technology Standards Committee, 2002).
- **OBAA:** Es una iniciativa brasileña que propone un estándar de metadatos para OAs compatible con el estándar IEEE LOM. Su principal diferencia es que está orientado a la interoperabilidad de los OAs en plataformas como Web, TV digital y dispositivos móviles. Además permite almacenar informaciones adicionales relacionadas con accesibilidad para personas con necesidades especiales y características propias del contexto de educación brasileño (Vicari, et al., 2009).

Estos recursos educativos son alojados junto con sus metadatos en Repositorios de Objetos de Aprendizaje (ROAs), que son bibliotecas digitales especializadas, orientados a facilitar la búsqueda y recuperación de los OAs, de manera que puedan ser utilizados en diversos ambientes de e-learning. Algunos repositorios albergan los OAs de una amplia variedad de áreas temáticas, otros se centran en temas específicos o en niveles educativos. Generalmente para compartir recursos a través de un ROA se requiere de un proceso de suscripción al repositorio, y en algunos casos se exige de un proceso de revisión de los materiales. Sin embargo, la mayoría de ROAs son abiertos, en el sentido de que cualquier usuario puede acceder a los OAs que contienen.

De acuerdo a la localización de los OAs y sus metadatos, estos repositorios se pueden clasificar en los siguientes

tres tipos: 1) los que alojan los OAs, siguiendo un modelo centralizado en el que los recursos se almacenan en un único servidor (ROA local); 2) los que funcionan como portales que sólo almacenan los metadatos y proporcionan enlaces a contenidos educativos de terceros, conocidos también como Referatorios, y 3) los que combinan los dos primeros tipos (Alfano & Henderson, 2007; McGreal, 2008).

Independiente del tipo de repositorio, uno de los principales objetivos de los ROAs, y que va en consonancia con las funcionalidades de los metadatos de los OAs, mencionadas previamente, es la búsqueda y recuperación de OAs, que generalmente se realizan a través de palabras clave u otras herramientas de búsqueda avanzada. Como en cualquier proceso de búsqueda de un recurso digital, se espera que su resultado sea la obtención de manera ágil y oportuna de información pertinente, precisa y relevante (Hubp, 2004). Sin embargo, ésta no es la situación general en los ROA, donde los resultados de búsqueda a menudo son inadecuados. De hecho, según Ochoa (2008), un recurso podrá ser parte de un repositorio y nunca ser recuperado en búsquedas relevantes, debido a que la utilidad de dicho repositorio está limitada a la calidad de los metadatos que describen sus recursos. En otras palabras, por más avanzada que sea la arquitectura y procedimientos de búsqueda de un repositorio, si los metadatos de los recursos no son adecuados, las búsquedas estarán destinadas al fracaso.

Esta afirmación es compartida por varios autores que mencionan que el crecimiento en número, tamaño y diversidad de los ROAs hace que la calidad de los metadatos sea un tema cada vez más importante. Metadatos inexactos, incompletos e inconsistentes suelen ser un problema común en los repositorios, lo que implica un problema para el acceso a los recursos (Bui & Park, 2006; Nichols, Chan & Bainbridge, 2008).

Estudios como el presentado por Tabares et al. (2011) evidencian la necesidad de evaluar la calidad de los OAs desde diferentes aspectos. En este trabajo los autores evaluaron el tiempo de respuesta, cantidad de recursos, efectividad del contenido, motivación al aprendizaje y facilidad para el uso de OAs en un repositorio, desde la percepción de los usuarios. Los resultados obtenidos fueron negativos, demostrando que las búsquedas

entregaban recursos que no cubrían las necesidades definidas. La conclusión ante este panorama es que el primer paso para realizar una adecuada búsqueda y recuperación en ROAs es garantizar que los metadatos contenidos en los ROAs que se están accediendo tengan un cierto nivel de calidad. Sin embargo, esto no es una tarea fácil pues son varios los problemas detectados que afectan la calidad de los OAs y que pueden llegar a provocar desinterés en este tipo de recursos digitales (Ochoa, 2008).

Precisamente, el problema de la evaluación de la calidad de los metadatos de OAs es el foco de atención de este artículo, el cual se encuentra organizado como se especifica a continuación. En la sección 2 se discuten algunos de los trabajos relacionados con esta investigación. En la sección 3 se describe de manera detallada cada uno de los índices propuestos para evaluar dicha calidad. En la sección 4 se presentan y analizan los resultados experimentales obtenidos al aplicar las métricas tanto a una serie de OAs con metadatos sintéticos, como a un grupo de OAs reales utilizando un ROA real de uso abierto. Finalmente, en la sección 5, se exponen las conclusiones y trabajo futuro.

2. Trabajos relacionados

Se han propuesto varios enfoques para evaluación general de los OAs desde diferentes perspectivas. Algunos trabajos proponen valoraciones manuales realizadas por expertos (Morales et al., 2007; Morales, et al., 2005; Kay & Knaack, 2008), otros plantean la utilización de métricas propias de la Ingeniería de Software aplicadas al contexto de los OAs (Cuadrado-Gallego, 2005; Noor et al., 2009), otros usan estrategias de clasificación de acuerdo al uso de los recursos y la valoración de los usuarios (Duval, 2006; López et al., 2012; Yen, et al., 2010). Finalmente, otros investigadores se centran en determinar la calidad de los metadatos que describen los OAs, como es el caso de la investigación presentada en este artículo.

Para determinar el nivel de calidad de los metadatos, se deben contemplar diferentes aspectos que fueron identificados en varios trabajos revisados. Uno de dichos aspectos es la *completitud*, que indica si los metadatos describen los objetos tanto como sea posible. Esta

característica es evaluada para comprobar si cada elemento de los metadatos es efectivamente una instancia del estándar correspondiente (Motelet et al., 2006) y si cada instancia contiene datos, midiendo qué tanta información está disponible para el recurso. La calidad de los metadatos disminuye con la ausencia de elementos básicos que se utilizan para la búsqueda y recuperación (Bruce & Hillmann, 2004), por lo tanto el criterio de completitud se centra en determinar cuantitativamente la calidad de los metadatos (Menéndez, et al., 2012).

Un segundo aspecto de relevancia en la evaluación de calidad de los metadatos es la *consistencia*, la cual permite estimar el nivel de conformidad con el estándar de metadatos o las reglas establecidas por el repositorio que almacena los OAs (Bruce & Hillmann, 2004). En el trabajo de Menéndez, et al. (2012) se propone un modelo para evaluar la calidad de los OAs partiendo de las recomendaciones definidas en el estándar IEEE LOM, aplicando reglas de consistencia con las que deben cumplir los metadatos y estableciendo dos niveles de conformidad: La conformidad base que se da cuando los metadatos pueden incluir elementos extendidos o adicionales a los definidos por el estándar, y la conformidad estricta que se presenta cuando el OA contiene únicamente metadatos definidos por el estándar. Por su parte, Park & Tosaka (2010) evalúan las prácticas en control de calidad de los metadatos de repositorios digitales a través de una encuesta que permite identificar la consistencia como criterio fundamental en la evaluación de los metadatos.

Nichols et al. (2008) y Hughes (2005) presentan herramientas para la evaluación de los metadatos, específicamente para repositorios que trabajan con el protocolo OAI-PMH, por la forma como son obtenidos los metadatos. En los dos casos se evalúan diferentes aspectos asociados a la calidad de los metadatos, resaltando la evaluación de completitud y consistencia, que en el caso de Nichols et al. (2008) se basa únicamente en el estándar Dublin Core.

Un tercer aspecto relevante que debe ser considerado es la *coherencia*, la cual permite determinar el grado con el que los metadatos describen el mismo recurso. En el trabajo de Bruce & Hillmann (2004) se analiza la correlación existente entre algunos campos de un grupo de metadatos. En el trabajo presentado por Sanz et al.

(2009) se desarrolla un método de evaluación utilizando algunos elementos del estándar IEEE LOM y verificando, entre varias características consideradas, la cohesión entre los metadatos para determinar el nivel de calidad y reusabilidad de los recursos. Margaritopoulos et al. (2008) también evalúan los recursos digitales definiendo reglas que imponen o prohíben valores en ciertos campos, teniendo en cuenta las dependencias que existen entre algunos campos.

Uno de los trabajos que incluye más aspectos para determinar la calidad de los metadatos es el presentado por Ochoa & Duval (2006), el cual propone métricas basadas en los mismos parámetros de calidad utilizados para la revisión manual de metadatos: la integridad, precisión, procedencia, el cumplimiento de las expectativas, la consistencia lógica y coherencia, oportunidad y accesibilidad, entre otros. Los resultados preliminares sugieren que las métricas son sensibles a las características de calidad en los metadatos y recomiendan un trabajo más amplio para validar y calibrar los indicadores propuestos.

Como conclusión de esta sección, es posible afirmar que la *completitud*, *consistencia* y *coherencia* son los aspectos que más se tienen en cuenta y que tienen un mayor impacto en la calidad de los metadatos. Sin embargo, existen investigaciones que incluyen aspectos adicionales como son precisión, accesibilidad y redundancia. En la Tabla 1 se muestra una clasificación de los trabajos revisados de acuerdo con tales aspectos, presentados en orden cronológico descendente.

Tabla 1. Aspectos analizados para la evaluación de la calidad de metadatos

Trabajo	Completitud	Consistencia	Coherencia	Otros Criterios
(Menéndez, et al. 2012)		X		
(Park & Tosaka, 2010)		X		X
(Sanz, Sánchez & Dodero, 2009)			X	X
(Nichols, Chan & Bainbridge, 2008)	X	X		
(Margaritopoulos, et al., 2008)			X	X
(Ochoa & Duval, 2006)	X	X	X	X
(Hughes, 2005)	X	X		X
(Stvilia, et al., 2004)		X	X	
(Bruce & Hillmann, 2004)	X		X	X

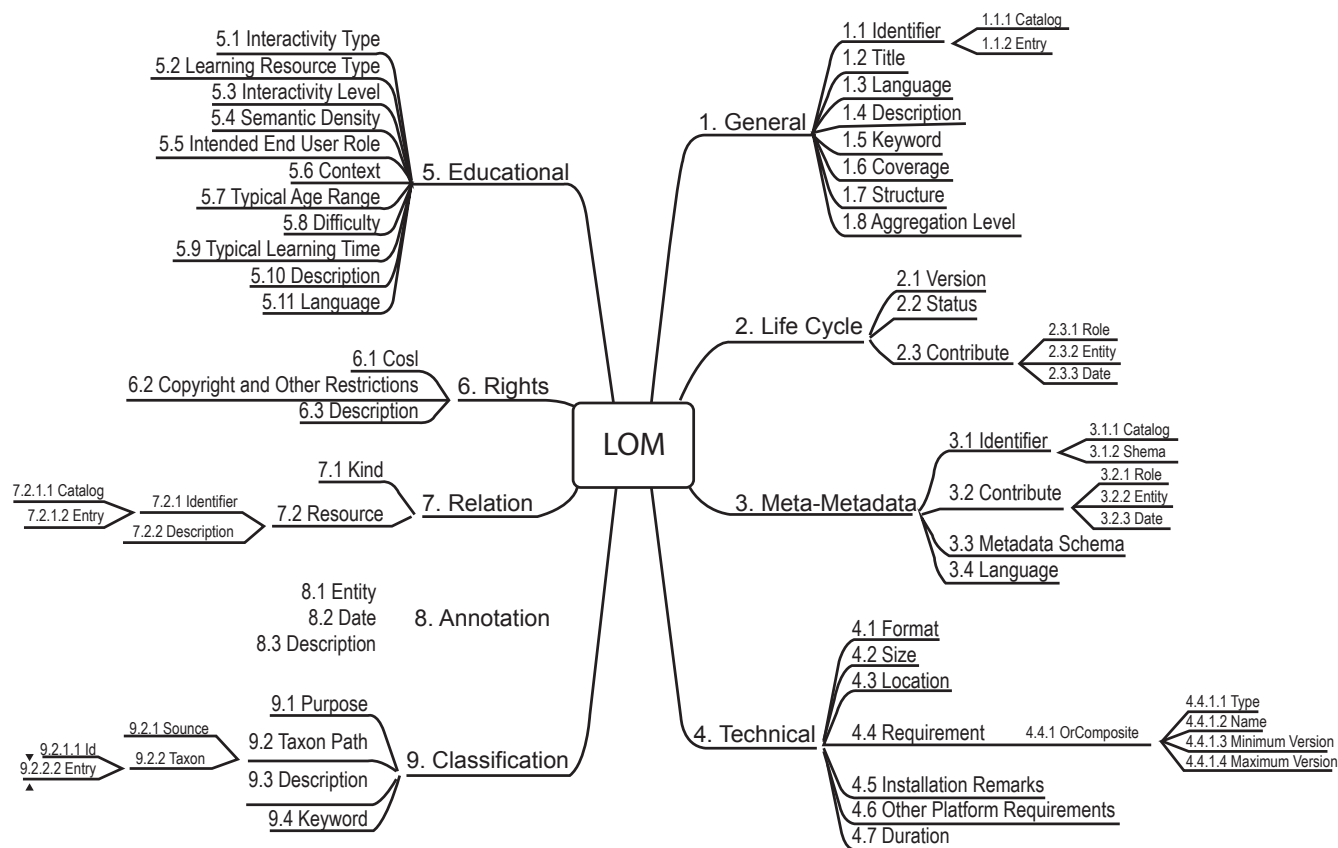
Como se puede apreciar en la revisión de los trabajos relacionados, son pocos los que cubren simultáneamente los aspectos de Completitud, Consistencia y Coherencia. Además no se presenta un modelo genérico que permita evaluar todos los metadatos incluidos en el estándar IEEE LOM, especialmente en el caso de la Completitud, ya que se menciona cómo podría ser calculada, pero no se especifica cómo determinar los pesos para cada metadato analizado. De igual forma, se encuentra que en estos trabajos no hay una profunda validación de las métricas propuestas.

3. Propuesta de evaluación basada en Métricas de Calidad

Como se describió en la sección anterior, existen diversos aspectos o criterios a considerar a la hora de deter-

minar la calidad de los metadatos de OAs (Morales et al., 2007). En este trabajo se propone el uso de tres de ellos y se especifica una métrica particular para cada uno a partir del estándar de metadatos IEEE LOM. La elección de tal estándar se debió a que es específico de OAs, es internacionalmente reconocido, y es uno de los que ofrece mayores posibilidades para describir y especificar tales recursos. En la Figura 1 se muestra su estructura jerárquica dividida en categorías. Sin embargo, es importante mencionar que dicha elección no es necesariamente una limitación pues si se desea aplicar las métricas propuestas a OAs que están almacenados en ROAs pero que están descritos en otro estándar, es posible definir una equivalencia con el estándar LOM, como es el caso de una de las validaciones presentadas en la siguiente sección.

Figura 1. Estructura jerárquica del estándar IEEE LOM (Barker, 2005)



A continuación se describe en detalle cómo realizar el cálculo de las métricas propuestas.

3.1. Métrica de completitud

En términos generales, la métrica de completitud consiste en revisar un subconjunto de los campos que componen los metadatos para determinar si contienen algún valor o en el caso de los campos con múltiples valores si contienen por lo menos una instancia.

Diversos trabajos determinan dicho subconjunto según la importancia relativa que consideren tienen ciertos campos. Bui & Park (2006), Hughes (2005) y Ochoa (2008), evalúan sólo algunos campos, los cuales son considerados “los más frecuentemente utilizados”. Por su parte en Rodríguez et al. (2011) se concluye que los metadatos usados principalmente con fines de búsqueda, por tanto de interés para la completitud, son aquellos asociados con la información general, el ciclo de vida y las características educacionales.

Tabla 2. ROAs revisados para el cálculo de pesos de metadatos

#	Nombre	URL
1	MERLOT	http://www.merlot.org/
2	GATEWAY	http://www.thegateway.org/
3	MARICOPA	http://www.mcli.dist.maricopa.edu/mlx/
4	GLOBE	http://www.globe-info.org/
5	ARIADNE	http://ariadne.cs.kuleuven.be/AriadneFinder/
6	FEB	http://feb.ufrgs.br
7	SMETE	http://www.smete.org/smete/
8	AGREGA	http://www.proyectoagrega.es
9	FUNES	http://funes.uniandes.edu.co/cgi/search/simple
10	LA FLOR	http://laflor.laclo.org/
11	Colombia Aprende	http://64.76.190.172/drupalM/
12	ROAp	http://roap.medellin.unal.edu.co/

Diferente a tales trabajos, la propuesta presentada en este artículo considera la totalidad de los campos considerados en el estándar IEEE LOM. Esto no significa que a todos los campos se les dé la misma importancia. Por el contrario, se propone el uso de un sistema de pesos para considerar la importancia relativa de los mismos. Para determinar dichos pesos se revisaron 12 ROAs diferentes, presentados en la Tabla 2. En cada repositorio se evaluó cuáles metadatos eran usados en las búsquedas avanzadas y cuáles eran exhibidos a los usuarios en los resultados de las búsquedas. De esta for-

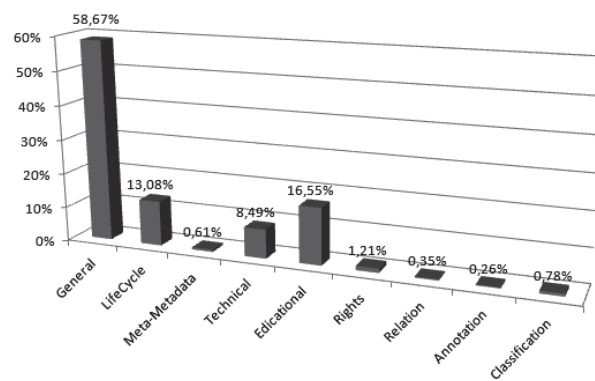
ma se obtuvo la Tabla A1 del Anexo A con valores binarios: 1 si el metadato i se considera en el ROA j , y 0 en caso contrario. Así, el peso del metadato i , corresponde a la suma en j de ambas revisiones, llevando luego este valor a porcentaje respecto a la totalidad de los metadatos, elevando previamente al cuadrado para enfatizar la importancia relativa.

Como resumen de este procedimiento, la Tabla 3 recoge los cinco metadatos con mayor peso según su utilización en los repositorios analizados, mientras que la Figura 2 muestra la concentración de pesos según las categorías de LOM.

Tabla 3. Top 5 de metadatos más usados para los 12 ROAs seleccionados

Metadatos LOM	Pesos
General.Title	22,18%
General.Keyword	19,50%
General.Description	12,48%
LifeCycle.Contribute.Entity	12,48%
Educational.LearningResourceType	8,67%

Figura 2. Distribución de los pesos de acuerdo a las categorías IEEE LOM



Una vez calculados los pesos para todos los metadatos, el cálculo de la métrica completitud para un OA específico consiste en revisar cada campo de los metadatos dando un valor de 1 si tiene algún valor o instancia ó 0 en caso contrario. Estos valores se multiplican luego por el peso correspondiente y el resultado final de la métrica corresponde a la sumatoria de estos resultados. De este procedimiento puede intuirse que el rango para esta métrica es $[0,1]$, donde un valor de 0 significa que el OA no tienen ningún metadato diligenciado, mientras

que un valor de 1 que todos los metadatos están presentes. La fórmula correspondiente es:

$$\text{Compleitud: } \sum_1^n (K_i * M_i)$$

Donde $K_i \in [0,1]$ es el peso del metadato i , y M_i es el valor binario que indica si dicho metadato en el OA tiene algún valor o instancia.

3.2. Métrica de consistencia

Si la métrica de completitud determina para cada metadato si existen o no valores, la métrica de consistencia determina si dichos valores son o no “consistentes”. En este caso la consistencia se evalúa a partir de la especificación del estándar de metadatos, el cual establece si un determinado metadato puede tomar valores libres, o si por el contrario existe una lista de valores posibles. En esta investigación se revisó detenidamente el estándar IEEE LOM y se sacaron los valores posibles de los metadatos que así lo establecen, tal como se muestra en la Tabla 4. De esta manera, para calcular la métrica de consistencia se revisa para un OA si existe algún va-

lor para cada uno de dichos metadatos y, en caso afirmativo, se asigna un valor de 1 si cumple con el valor permitido o de 0 en caso contrario. Posteriormente, la sumatoria de estos valores es dividida por la cantidad de comparaciones posibles, obteniendo así el valor final de la métrica.

Al igual que en el caso de la métrica de completitud, el rango de la métrica de consistencia es [0,1], donde un valor de 0 significa que el OA tiene todos los metadatos presentados en la Tabla 4 y que los mismos son inconsistentes. Entre tanto un valor de 1 significa que el OA tiene todos los metadatos presentados en la Tabla 4 y que los valores de éstos son consistentes. La fórmula correspondiente es:

$$\text{Consistencia: } \sum_1^n \frac{S_i}{P}$$

Donde S_i es igual a 0 si el metadato i es inconsistente o 1 si es consistente, y P es la cantidad de comparaciones evaluadas. En caso de que no se pueda evaluar ninguna comparación ($P=0$) la consistencia toma un valor de “No Aplica”.

Tabla 4. Comparaciones para evaluar la métrica de consistencia

Metadatos LOM	Valores Posibles
General.Agregation Level	1, 2, 3, 4
General.Structure	atomic, collection, networked, hierarchical, linear
LifeCycle.Contribute.Role	author, publisher, unknown, initiator, terminator, validator, editor, graphical designer, technical implementer, content provider, technical validator, educational validator, script writer, instructional designer, subject matter expert
LifeCycle.Status	draft, final, revised, unavailable
Meta-Metadata.Contribute.Role	creator, validator
Educational.Interactivity Type	active, expositive, mixed
Educational.Learning Resource Type	exercise, simulation, questionnaire, diagram, figure, graph, index, slide, table, narrative text, exam, experiment, problem, statement, self assessment, lecture
Educational.Interactivity Level	very low, low, medium, high, very high
Educational.Semantic Density	very low, low, medium, high, very high
Educational.Intended End User Role	teacher, author, learner, manager
Educational.Context	school, higher education, training, other
Educational.Difficulty	very easy, easy, medium, difficult, very difficult
Rights.Copyright and Other Restrictions	yes, no
Rights.Cost	yes, no
Classification.Purpose	discipline, idea, prerequisite, educational objective, accessibility, restrictions, educational level, skill level, security level, competency

3.3. Métrica de coherencia

Diferente a la métrica de consistencia, cuyo objetivo es revisar un metadato versus sus valores posibles, la métrica de coherencia compara y revisa el valor de un metadato en relación con el valor de otro. Tanto desde el punto de vista conceptual como de la especificación del estándar, existen metadatos que tienen una alta correlación y este hecho determina su coherencia. En este caso se revisó de nuevo detenidamente el estándar IEEE LOM y se sacaron las comparaciones posibles que se presentan en la Tabla 5.

Una diferencia fundamental entre esta investigación y los trabajos relacionados, descritos en la sección anterior, es que en vez de dar un valor binario a cada comparación, se establece un valor escalonado que da cuenta de la pertinencia de cada combinación de valores. Así por ejemplo (ver comparación 1 de la Tabla 5), aunque lo esperado es que un OA cuyo metadato “General/ Estructura” tenga un valor de “Atómico” y a su vez un valor de 1 en el metadato “General/ Nivel de agregación”, en la práctica se puede presentar que tenga un valor de 2 o 3, e incluso en situaciones atípicas un valor de 4.

Tabla 5. Comparaciones para evaluar la coherencia

#	Metadatos y Valores		Resultado
1	General.Structure	General.AgregationLevel	
	atomic	1	1
	atomic	2	0,5
	atomic	3	0,25
	atomic	4	0,125
	collection, networked, hierarchical, linear	1	0,5
	collection, networked, hierarchical, linear	2, 3, 4	1
2	Educational.InteractivityType	Educational.InteractivityLevel	
	active, mixed	very high, high, medium, low, very low	1
	expositive	very high, high	0
	expositive	medium	0,5
	expositive	low, very low	1
3	Educational.InteractivityType	Educational.Learning Resource Type	
	active	exercise, simulation, questionnaire, exam, experiment, problem statement, self assessment	1
	active	diagram, figure, graph, index, slide, table, narrative text, lecture	0
	expositive	exercise, simulation, questionnaire, exam, experiment, problem statement, self assessment	0
	expositive	diagram, figure, graph, index, slide, table, narrative text, lecture	1
	mixed	exercise, simulation, questionnaire, exam, experiment, problem statement, self assessment, diagram, figure, graph, index, slide, table, narrative text, lecture	1

De manera similar al caso de la métrica de consistencia, para calcular la métrica de coherencia se revisa para un OA si existe alguna de las combinaciones válidas presentadas en la Tabla 5 y, en caso afirmativo, se asigna el valor de la combinación correspondiente, o 1 en caso contrario. Posteriormente, la sumatoria de estos valores es dividida por la cantidad de comparaciones posibles, es decir para este caso de 3, obteniendo así el valor final

de la métrica. Nuevamente, el rango de esta métrica es [0,1] y la fórmula correspondiente es:

$$Coherencia: \sum_{1}^n \frac{R_i}{Q}$$

Donde R_i se obtiene a partir de la Tabla 5 en caso que exista la combinación y Q representa la cantidad de

comparaciones evaluadas. Si no es posible evaluar ninguna comparación, la coherencia toma un valor de “No Aplica”.

4. Validación y análisis de resultados experimentales

Con el fin de validar experimentalmente las métricas propuestas en este artículo, éstas fueron implementadas en dos situaciones diferentes, una controlada y otra con datos reales obtenidos a partir de una consulta en

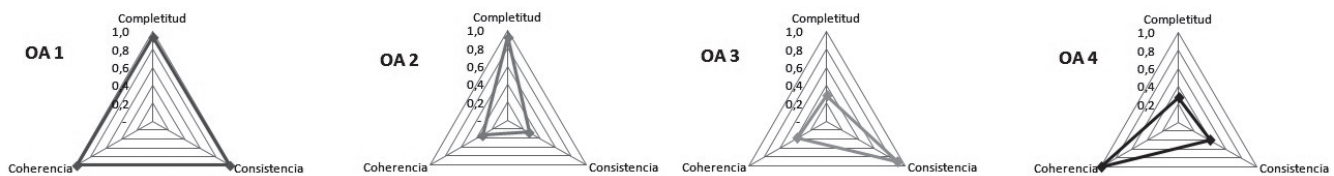
línea. Para la primera situación, se crearon cuatro OAs ficticios con metadatos sintéticos, los cuales se resumen en la Tabla 6 y los resultados obtenidos se presentan en la Figura 3.

En este caso puede apreciarse que las métricas cambian sustancialmente de un OA a otro. El primer objeto es muy completo, consistente y coherente. El segundo objeto es muy completo, pero no muy consistente ni coherente. El tercer objeto es muy consistente, pero no muy completo ni coherente. Finalmente el cuarto objeto es muy coherente, pero no muy completo ni consistente.

Tabla 6. Metadatos de cuatro OAs sintéticos

Metadato	OA 1	OA 2	OA 3	OA 4
General.Title	OA Sintético 1	OA Sintético 2		
General.Language	Español	Español		
General.Description	OA prueba	OA prueba		
General.Keyword	Validación	Validación		
General.Structure	atomic	#\$%&	atomic	atomic
General.Agregation Level	1	1	4	1
LifeCycle.Status	draft	#\$%&	draft	#\$%&
LifeCycle.Contribute.Role	author	#\$%&	author	#\$%&
LifeCycle.Contribute.Entity	Escuela ABC	Escuela ABC	Escuela ABC	Escuela ABC
Technical.Format	pdf	pdf		
Technical.Location	URL			
Educational.Interactivity Type	active	active	expositive	expositive
Educational.Learning Resource Type	exercise	#\$%&	narrative text	diagram
Educational.Interactivity Level	very low	very low	high	medium
Educational.Semantic Density	very low	#\$%&	active	#\$%&
Educational.Intended End User Role	teacher	#\$%&	teacher	#\$%&
Educational.Context	school	#\$%&	school	#\$%&
Educational.Difficulty	very easy	#\$%&	very easy	#\$%&

Figura 3. Resultados de la validación de las métricas para cuatro OAs sintéticos

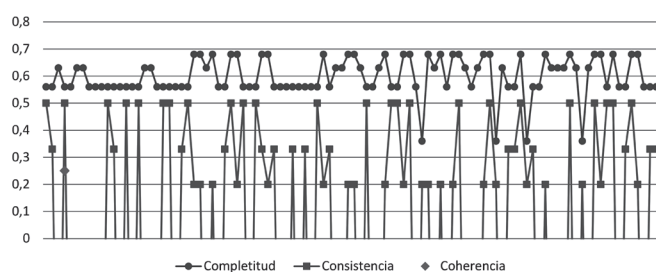


Esto lleva a demostrar por una parte que la propuesta de evaluación basada en métricas de calidad permite representar claramente las características reales de los metadatos que describen los OAs, y por otra que las métricas determinan un factor diferenciador de la calidad del etiquetado de tales metadatos.

Para la segunda situación de experimentación se recuperaron 100 OAs de la Federación Educa Brasil – FEB (<http://feb.ufrgs.br>), correspondientes a los primeros 100 resultados de la búsqueda con la palabra clave “cáncer”. En la Figura 4 se muestran las métricas obtenidas para estos OAs.

En este caso se puede observar que los valores de la completitud para la mayoría de los OAs oscilan entre 0.5 y 0.7, lo cual indica que los metadatos de dichos OAs incluyen algunos de los presentados en la Tabla 3, considerados como “relevantes”, pero ninguno alcanza siquiera a tener esos 5 al tiempo. La consistencia por su parte es muy fluctuante, oscilando entre 0 y 0.5, lo cual da indicio de la existencia de muchos metadatos que no cumplen con la especificación de LOM en cuanto a valores posibles. Finalmente, para el caso de la coherencia solo fue posible evaluar uno de los OAs, encerrado en un círculo, el cual obtuvo un valor muy bajo (0.25).

Figura 4. Valores de las métricas para 100 OAs recuperados del repositorio FEB de Brasil



5. Conclusiones y trabajo futuro

Tomando en consideración que el éxito en la búsqueda y recuperación de OAs dentro de repositorios nacionales e internacionales, para apoyar procesos de enseñanza-aprendizaje, depende en gran medida de la calidad de los metadatos que los describen, la propuesta de evaluación de estos recursos digitales, presentada en este artículo, se focaliza en la definición de tres métricas diferentes para determinar su nivel de calidad. De hecho, puede considerarse que este trabajo de investigación lo que propone es un índice, solo que no agregado (un único valor) sino compuesto (la unión de las tres métricas). En este sentido, esta propuesta se diferencia de otros trabajos relacionados pues se examinan tres “dimensiones” diferentes de los metadatos, a partir de métricas en una escala común pero no correlacionadas.

Es importante señalar que otros factores diferenciadores subyacen en las métricas en sí. En el caso de la métrica de completitud no sólo se asignan valores binarios a la presencia o ausencia de cada metadato, sino que los mismos son ponderados según un sistema de pesos

determinados – que establecen su nivel de pertinencia – a partir de un análisis con 12 ROAs representativos tanto a nivel nacional como internacional. En el caso de la métrica de consistencia se realiza una revisión minuciosa de las especificaciones del estándar de metadatos, en este caso IEEE LOM, para determinar si hay valores por fuera de dicha especificación. Finalmente, en el caso de la métrica de coherencia se establecen, igualmente a partir de las especificaciones del estándar, las combinaciones válidas o erróneas de valores entre metadatos, pero en vez de usar valores binarios para su calificación se propone el uso de valores discretos escalonados.

De las pruebas experimentales realizadas tanto con OAs sintéticos como reales de un ROA conocido, se pudo determinar que la propuesta permite en efecto evaluar la calidad de los metadatos involucrados, pero adicional a esto, permite caracterizar estos metadatos. Esto es de vital importancia pues termina convirtiéndose en un insumo relevante e indispensable tanto para procesos de evaluación automática de los metadatos, como también para procesos de recomendación de los OAs en sistemas de e-learning.

Un hallazgo adicional durante esta investigación, específicamente en la situación de una búsqueda real, fue que el ROA empleado (la Federación Educa Brasil – FEB), si bien congrega una enorme cantidad de OAs – más de 40.000 – y tiene cientos de accesos diarios, presenta algunos problemas de calidad de metadatos de los OAs, principalmente respecto a la completitud y coherencia de los mismos.

Cabe mencionar por último, que parte del trabajo futuro de esta investigación es la incorporación de nuevas métricas de calidad en ROAs tanto para los metadatos (e.g. congruencia entre metadatos con valores libres), como para los OAs en sí (e.g. disponibilidad o velocidad de acceso y recuperación).

6. Agradecimientos

El trabajo de investigación presentado en este artículo fue financiado parcialmente por el proyecto de la Vicerrectoría de Investigación de la Universidad Nacional de Colombia titulado: “Mejoramiento de las capacidades académicas de visibilidad y cooperación entre gru-

pos iberoamericanos de investigación en informática educativa - Convocatoria Nacional - Fortalecimiento de Alianzas Estratégicas 2012” con código Quipú 202010011122, y por el proyecto de Colciencias titulado: “ROAC Creación de un modelo para la Federación de OA en Colombia que permita su integración a confederaciones internacionales” de la Universidad Nacional de Colombia – Sedes Manizales y Medellín, con código 1119-521-29361.

7. Referencias

1. ALFANO, C. S., & HENDERSON, S. L. (2007). Repositories. In *Learning Objects for Instruction: Design and Evaluation*, 16 – 28.
2. BARKER, P. (2005). What is IEEE Learning Object Metadata/IMS Learning Resource Metadata? *CETIS Standards Briefing Series, JISC (Joint Information Systems Committee of the Universities' Funding Councils)*.
3. Bruce, T. R., & Hillmann, D. I. (2004). The Continuum of Metadata Quality: Defining, Expressing, Exploiting. In *Metadata in Practice*.
4. BUI, Y., & PARK, J. (2006). An Assessment of Metadata Quality: A Case Study of the National Science Digital Library Metadata Repository.
5. CUADRADO, J. (2005). Adaptación de las Métricas de Reusabilidad de la Ingeniería del Software a los Learning Objects. *RED. Revista de Educación a Distancia*.
6. DUQUE, N. D. (2009). *Modelo Adaptativo Multi-Agente para la Planificación y Ejecución de Cursos Virtuales Personalizados - Tesis Doctoral*. Universidad Nacional de Colombia - Sede Medellín.
7. DUVAL, E. (2006). LearnRank: Towards a real quality measure for learning. *Handbook on quality and standardisation in E-learning*, 457–463.
8. HUGHES, B. (2004). Metadata Quality Evaluation: Experience from the Open Language Archives Community. *Digital Libraries: International Collaboration and Cross-Fertilization*, 320–329.
9. ISBD. (2004). International Standard Bibliographic Description for Electronic Resources. Draft for World Wide Revision November 26, 2012, from http://www.ifla.org/VII/s13/guide/isbder_ww2-1-04.pdf
10. KAY, R. H., KNAACK, L., & VAIDEHI, V. (2008). A Multi-component Model for Assessing Learning Objects: The Learning Object Evaluation Metric (LOEM). *Australasian Journal of Educational Technology*, 24(5), 574–591.
11. Learning Technology Standards Committee. (2002). IEEE Standard for Learning Object Metadata. *Institute of Electrical and Electronics Engineers, New York*.
12. LÓPEZ, V. F., DE LA PRIETA, F., OGIHARA, M., & WONG, D. D. (2012). A model for multi-label classification and ranking of learning objects. *Expert Systems with Applications*, 39(10), 8878–8884. doi:10.1016/j.eswa.2012.02.021
13. LUGO HUBP, M. (2004). El impacto de los recursos digitales en la bibliotecas. In UNAM (Ed.), *Administración de servicios de información* (p. 173).
14. MARGARITPOULOS, T., MAVRIDIS, I., MARGARITPOULOS, M., & MANITSARIS, A. (2008). A Conceptual Framework for Metadata Quality Assessment. *Proc. Int'l Conf. on Dublin Core and Metadata Applications*, 104–113.
15. MCGREAL, R. (2008). A Typology of Learning Object Repositories. In *Handbook on Information Technologies for Education and Training*, 5 – 28. doi:10.1007/978-3-540-74155-8_1
16. MENÉNDEZ, V. H., CASTELLANOS, M. E., VIDAL, C., & SEGURA, A. (2012). Un Modelo de Calidad de Objetos de Aprendizaje basado en la Semántica de sus Metadatos. *LACLO 2012 - Séptima Conferencia Latinoamericana de Objetos y Tecnologías de Aprendizaje*, 3(1).
17. MORALES, E., GARCÍA, F. J., BARRÓN, Á., BERLANGA, A. J., & LÓPEZ, C. (2005). Propuesta de Evaluación de Objetos de Aprendizaje. In *II Simposio Pluridisciplinar sobre Diseño, Evaluación y Descripción de Contenidos Educativos, SPEDECE*. 2005.
18. MORALES, E., GIL, A., & GARCÍA, F. (2007). Arquitectura para la Recuperación de Objetos de Aprendizaje de calidad en Repositorios Distribuidos. *SCHA: Sistemas Hipermedia Colaborativos y Adaptativos. II Congreso Español de Informática CEDI 2007*, 1, 31–38.
19. MOTELET, O., BALOIAN, N., & PINO, J. A. (2007). Learning Object Metadata and Automatic Processes: Issues and Perspectives. In *Learning Objects: Standards, Metadata, Repositories & LCMS*, 185–218.
20. NICHOLS, D., CHAN, C., & BAINBRIDGE, D. (2008). A Tool for Metadata Analysis. Hamilton, New Zealand.
21. NOOR, S. F. M., YUSOF, N., & HASHIM, S. Z. M. (2009). A Metrics Suite for Measuring Reusability of Learning Objects. *Ninth International Conference on Intelligent Systems Design and Applications*, 961–963. doi:10.1109/ISDA.2009.114

22. OCHOA, X. (2008). *Learnometrics: Metrics for Learning Objects*. Katholieke Universiteit Leuven.
23. OCHOA, X., & DUVAL, E. (2006). Quality Metrics for Learning Object Metadata. *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications*, 1004–1011.
24. PARK, J.-R., & TOSAKA, Y. (2010). Metadata Quality Control in Digital Repositories and Collections: Criteria, Semantics, and Mechanisms. *Cataloging & Classification Quarterly*, 48(8), 696–715. doi:10.1080/01639374.2010.508711
25. RODRÍGUEZ, P. A., ISAZA, G., & DUQUE, N. D. (2011). Búsqueda personalizada en Repositorios de Objetos de Aprendizaje a partir del perfil del estudiante. *CAVA 2011*.
26. SANZ, J., SÁNCHEZ, S., & DODERO, J. M. (2009). Reusability Evaluation of Learning Objects Stored in Open Repositories Based on Their Metadata. *Communications in Computer and Information Science*, 46(2), 193–202. doi:10.1007/978-3-642-04590-5_18
27. STVILIA, B., GASSER, L., TWIDALE, M. B., SHREEVES, S. L., & COLE, T. W. (2004). Metadata Quality for Federated Collections. *Proceedings of the Ninth International Conference on Information Quality (ICIQ-04)*, 111–125.
28. TABARES, V., DUQUE, N., & MORENO, J. (2011). Análisis experimental de la utilidad en la recuperación de objetos de aprendizaje desde repositorios remotos. In *Congreso Internacional en Ambientes Virtuales de Aprendizajes Accesibles y Adaptativos - CAVA 2011*.
29. VICARI, R., BEZ, M., MARQUES, J., RIBEIRO, A., GLUZ, J. C., PASSERINO, L., ... ROESLER, V. (2010). Proposta Brasileira de Metadados para Objetos de Aprendizagem Baseados em Agentes (OBAA). *Novas Tecnologias na Educação*.
30. VICARI, R., GLUZ, J. C., SANTOS, E. R., THOMPSEN PRIMO, T., LONGHI, L. H., BORDIGNON, A., BEHAR, P., & PASSERINO, Liliana Maria Machado, Raymundo Carlos Roesler, V. (2009). Proposta de Padrão para Metadados de Objetos de Aprendizagem Multiplataforma. *Projeto OBAA*.
31. WEIBEL, S. L., & KOCH, T. (2000). The Dublin Core Metadata Initiative: Mission, Current Activities, and Future Directions. *D-Lib Magazine*, 6(12).
32. WILEY, D. A. (2001). Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy, 2830(435), 1–35.
33. YEN, N. Y., SHIH, T. K., MEMBER, S., CHAO, L. R., & JIN, Q. (2010). Ranking Metrics and Search Guidance for Learning Object Repository. *IEEE Transactions on Learning Technologies*, 3(3), 250–264.

Anexo A

Tabla A1. Resultados de la revisión de ROAs para cálculo de peso de metadatos

ROA Metadatos LOM	1	2	3	4	5	6	7	8	9	10	11	12	SUB	1	2	3	4	5	6	7	8	9	10	11	12	SUB	TOTAL	TOTAL	PESOS
	EN BÚSQUEDAS AVANZADAS												TOTAL	EN RESULTADOS												TOTAL	TOTAL	%	
General.Identifier.Catalog												1	1													0	1	1	0,09%
General.Identifier.Entry							1					1	2												0	2	4	0,35%	
General.Title	1	1				1		1				1	5	1	1	1	1	1	1	1	1	1	1	1	11	16	256	22,18%	
Genera.Language	1		1	1				1		1			6												0	6	36	3,12%	
General.Description	1	1						1				1	4	1	1	1	1	1	1				1	1	8	12	144	12,48%	
General.Keyword	1	1	1	1	1	1	1	1	1	1	1	1	12							1				1	1	3	15	225	19,50%
General.Coverage												1	1												0	1	1	0,09%	
General.Structure												1	1												0	1	1	0,09%	
General.Agregation Level								1				1	2								1				1	3	9	0,78%	
LifeCycle.Version												1	1												0	1	1	0,09%	
LifeCycle.Status									1			1	2												0	2	4	0,35%	
LifeCycle.Contribute.Role												1	1												0	1	1	0,09%	
LifeCycle.Contribute.Entity	1			1	1	1	1	1	1	1		1	9	1						1		1			3	12	144	12,48%	
LifeCycle.Contribute.Date												1	1												0	1	1	0,09%	
Meta-Metadatas.Identifier.Catalog												1	1												0	1	1	0,09%	
Meta-Metadatas.Identifier.Entry												1	1												0	1	1	0,09%	
Meta-Metadatas.Contribute.Role												1	1												0	1	1	0,09%	
Meta-Metadatas.Contribute.Entity												1	1												0	1	1	0,09%	
Meta-Metadatas.Contribute.Date												1	1												0	1	1	0,09%	

ROA Metadatos LOM	1	2	3	4	5	6	7	8	9	10	11	12	SUB TOTAL	1	2	3	4	5	6	7	8	9	10	11	12	SUB TOTAL	TOTAL	TOTAL ^2	PESOS
Meta-Metadatos.Metadata Schema												1	1													0	1	1	0,09%
Meta-Metadatos.Language												1	1													0	1	1	0,09%
Technical.Format	1	1		1	1			1			1		7								1				1	2	9	81	7,02%
Technical.Size												1	1													0	1	1	0,09%
Technical.Location	1											1	2						1							1	3	9	0,78%
Technical.RequirementCompositeType												1	1													0	1	1	0,09%
Technical.RequirementCompositeNam												1	1													0	1	1	0,09%
Technical.Requirement.OrComposite. MinimumVersion												1	1													0	1	1	0,09%
Technical.Requirement.OrComposite. MaximumVersion												1	1													0	1	1	0,09%
Technical.Installation Remarks												1	1													0	1	1	0,09%
Technical.OtherPlatformRequirements												1	1													0	1	1	0,09%
Technical.Duration												1	1													0	1	1	0,09%
Educational.Interactivity Type												1	1		1	1										3	4	16	1,39%
Educational.Learning Resource Type	1	1		1	1		1	1	1	1	1	1	10													0	10	100	8,67%
Educational.Interactivity Level												1	1													0	1	1	0,09%
Educational.Semantic Density												1	1													0	1	1	0,09%
Educational.Intended End User Role												1	1													0	1	1	0,09%
Educational.Context	1	1		1	1		1			1		1	7													0	7	49	4,25%
Educational.Typical Age Range		1						1				1	3		1											1	4	16	1,39%
Educational.Difficulty												1	1													0	1	1	0,09%
Educational.Typical Learning Time												1	1													0	1	1	0,09%
Educational.Description		1										1	2													0	2	4	0,35%
Educational.Language												1	1													0	1	1	0,09%
Rights.Cost	1	1										1	3													0	3	9	0,78%
Rights.CopyrihtandOtherRestrictions	1											1	2													0	2	4	0,35%
Rights.Description												1	1													0	1	1	0,09%
Relation.Kind												1	1													0	1	1	0,09%
Relation.Resource.Identifier.Catalog												1	1													0	1	1	0,09%
Relation.Resource.Identifier.Entry												1	1													0	1	1	0,09%
Relation.Resource.Description												1	1													0	1	1	0,09%
Annotation.Entity												1	1													0	1	1	0,09%
Annotation.Date												1	1													0	1	1	0,09%
Annotation.Description												1	1													0	1	1	0,09%
Classification.Purpose												1	1													0	1	1	0,09%
Classification.Taxon Path.Source								1				1	2													0	2	4	0,35%
Classification.Taxon Path.Taxon.Id												1	1													0	1	1	0,09%
Classification.Taxon Path.Taxon.Entry												1	1													0	1	1	0,09%
Classification.Description												1	1													0	1	1	0,09%
Classification.Keyword												1	1													0	1	1	0,09%

[Evaluación de la calidad de metadatos en repositorios digitales de objetos de aprendizaje]