

A Recommender System for Learning Objects on a Learning Management System: Implementation Experience

Jacqueline Solís-Céspedes
Tec Digital
Instituto Tecnológico de Costa Rica
Email: jacsolis@itcr.ac.cr

Mario Chacón-Rivas
Tec Digital
Instituto Tecnológico de Costa Rica
Email: machacon@itcr.ac.cr

Resumen—This paper aims to present the experience of implementing a recommender system for selecting learning objects (LO) for lessons through the Learning Management System (LMS) of Costa Rica Institute of Technology. Our implementation approach is made up of three components: A user model based on the descriptions of the lessons of a course given in an Instructional Design (ID), a serie of items which constitute the metadata bank of resources described using IEEE Standard for Learning Object Metadata model (IEEE-LOM) and harvested by Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) and finally, an utility function derived from the weighting between the collaborative assessment of LO and similarity measures to compare LO's metadata with descriptions of activities given for course. As validation mechanism, ID's tests were conducted in order to determine whether or not recover resources for lessons and precision tests with a pilot plan remained pending.

Keywords—Learning Object Metadata, Recommender Systems, Educational Metadata, LOM-IEEE, SQL, OAI-PMH

1. Introducción

La cantidad de recursos disponibles en los Repositorios de Objetos de Aprendizaje (ROA) puede resultar abrumadora para los usuarios que buscan reutilizar y aprovechar los materiales generados por otros profesionales. Del mismo modo, establecer criterios para filtrar de forma manual los materiales que podrían ser útiles para impartir una lección, es una tarea que requiere de mucho tiempo y experiencia por parte de quién los clasifica.

Las sistemas recomendadores permiten modelar estos procesos de forma automática o semiautomática, agilizando las tareas de filtrado y selección de recursos y además, permitiendo una trazabilidad de las cambios en cuanto a tendencias de búsqueda y contextos en los que se encuentra un mismo usuario.

El presente artículo pretende exponer la experiencia de implementación del Recomendador Híbrido de Objetos de Aprendizaje sobre el Sistema de Gestión del Aprendizaje (LMS, por sus siglas en inglés) del Instituto Tecnológico de Costa Rica (TEC); así como las diversas propuestas

de solución por las que ha pasado desde su planteamiento inicial hasta su puesta en producción.

Nuestro enfoque de implementación se centró en modelar un sistema recomendador para las actividades descritas en un diseño instruccional, donde cada actividad contiene una serie de descriptores que marcan el eje temático hacia el cuál se ubica una sesión en particular. De modo, que utilizando este insumo, se buscan recursos dentro de un banco de datos local que podrían resultar de interés para utilizar en la lección.

Para exponer la experiencia, se distribuyó el documento en secciones de la siguiente manera: La subsección 1.1 hace referencia a los trabajos relacionados que se han realizado tanto en TEC como en otras instancias internacionales mediante el uso de recomendadores híbridos para objetos de aprendizaje.

Seguidamente, la Sección 2 contiene la propuesta de solución que fue planteada con las funcionalidades tanto a nivel de usuario como de administrador. La Sección 3 corresponde a la metodología aplicada, la cuál se centra en aplicar múltiples iteraciones sobre los procesos de creación de modelos para mejorar la propuesta de implementación. Aquí se mencionan los enfoques bajo los que se intentó llevar a término la propuesta y las causas por las que fue necesario realizar las modificaciones.

La Sección 4 corresponde a un análisis cualitativo de los metadatos cosechados como items a recomendar, y finalmente la Sección 5 presenta las conclusiones derivadas del proceso de implementación de la solución.

1.1. Trabajo Relacionado

Según la IEEE, un Objeto de Aprendizaje (OA) se define como “cualquier entidad, digital o no, que puede ser usada para el aprendizaje, educación o capacitación” [1]. Esto quiere decir que cualquier elemento didáctico (digital o físico) puede ser considerado como un OA y esto incluye los tres niveles de jerarquización propuestos por Redeker: cursos completos, cursos parciales y unidades de aprendizaje [2].

Consecuentemente nace el concepto de Repositorio de Objetos de Aprendizaje (ROA) ante la necesidad de almacenar y compartir OAs digitales con el fin de que puedan ser

reutilizados [3] y aparecen instituciones como el IMS Global Consortium (compuesto por una serie de organizaciones y universidades distribuidas a nivel mundial) con especificaciones para empaquetar y transferir contenido usando XML con el fin organizarlos [4].

Actualmente, la disponibilidad de recursos, más la cantidad de repositorios abiertos que ponen a disposición diversos materiales educativos, hacen que el proceso de filtrado sea largo y difícil para el usuario. De este modo, resultan de gran valor el uso de herramientas tecnológicas que puedan mejorar su experiencia, realizar sugerencias de apoyo en la toma de decisiones y brindándole la libertad de utilizarlas o no. Estos elementos se traducen en el incremento de la fidelidad de los usuarios, incrementos en la satisfacción del usuario, incrementos en ventas, entre otros [5].

Se han desarrollado muchos aportes para la implementación de interfaces de consumo de OAs con el fin de facilitar su portabilidad entre las plataformas educativas y los LCMS¹, por ejemplo, en el caso del Instituto Tecnológico de Costa Rica, se están realizando múltiples esfuerzos por desarrollar, publicar y consumir OAs. Por un lado, el área de comunicación visual del TecDigital, ha tomado bajo su responsabilidad el proceso de virtualización de cursos, el cual, se ha concentrado en la producción de materiales educativos para cursos completos en lugar de desarrollar unidades mínimas exportables a distintos contextos.

Por otro lado, está el Repositorio de Objetos de Aprendizaje Institucional, el cuál “reúne, conserva y difunde a través del acceso abierto a los documentos en formato digital resultantes de la actividad académica y científica del TEC” [6] los cuáles pueden ser: OAs, informes (tanto de graduación como de investigación), materiales didácticos, publicaciones, entre otros. Además, este repositorio se encuentra cosechado desde el Repositorio Nacional de Costa Rica, conocido como “kimuk”.

Los primeros intentos de incorporar las funcionalidades de los sistemas recomendadores en procesos de aprendizaje se remontan aproximadamente al año 2000 y se han desarrollado en múltiples contextos, a pesar de que difieren significativamente de sus contrapartes comerciales.

En los contextos centrados en el profesor, las recomendaciones deben estar relacionadas con metas específicas y pueden incluir una variedad de tareas como la búsqueda de nuevos materiales, repaso de conocimientos, ilustraciones, entre otras. De este modo, su efectividad se puede evaluar de acuerdo al cumplimiento del conjunto de metas definidas inicialmente [7].

En la actualidad, existen múltiples sistemas recomendadores educativos con enfoques híbridos en los que se utilizan modelos como las cadenas de markov, enfoques colaborativos, técnicas de clústering, redes de aprendizaje, entre otras [7], [8]. Donde se ha visto que la efectividad de un algoritmo en un contexto no garantiza su rendimiento en otro, de modo que el análisis profundo de las características de los usuarios y de su contexto de acción son necesarias para una adecuada selección de algoritmos [7].

1. LCMS: Learning Content Management Systems.

Algunos ejemplos de enfoques híbridos aplicados a procesos de enseñanza son RACOFI², el cuál utiliza filtros basados en reglas y colaborativos [9]. ISIS-Tutor³, el cuál es uno de los primeros sistemas hipermedia con soporte de navegación adaptativo y recomienda rutas de navegación utilizando información social y metadatos de los estudiantes y actividades de aprendizaje [10]. LORM⁴ que utiliza una agregación de las calificaciones de los estudiantes y un modelo de correlación de experiencias registradas por otros estudiantes al interactuar con los OAs [11] y Delphos, que aplica la metodología desarrollada para recomendar recursos reutilizables de acuerdo a las necesidades de un profesor a partir de un banco de recursos disponibles en el repositorio AGORA⁵ [12], [13].

2. Solución Propuesta

Como se menciona en la introducción del artículo, nuestro objetivo se centró en modelar un sistema recomendador para las actividades descritas en un diseño instruccional (DI), donde cada actividad contiene una serie de descriptores que marcan la temática hacia la cuál se ubica una sesión en particular. De modo, que utilizando este insumo, se buscan OAs dentro de un banco de recursos que podrían resultar de interés para utilizar en la lección.

Sin embargo, para poner en funcionamiento el recomendador fue necesario generar una serie de funcionalidades adicionales que sólo son de interés para los administradores del servicio. Por este motivo se segmentó el acceso a la aplicación mediante dos perfiles: uno para los usuarios estándar y otro para los administradores de la plataforma.

2.1. Funcionalidades para un usuario

Un usuario convencional es aquel que sólo va a realizar operaciones de búsqueda, descarga y evaluación de las recomendaciones que le retorna la aplicación. Este perfil está típicamente ocupado por profesores o usuarios de la plataforma que deseen realizar búsquedas dentro de las colecciones.

Existen dos tareas típicas para un usuario: las búsquedas realizadas a partir de los DI y las búsquedas sencillas sobre la totalidad de los OAs disponibles.

Como se puede apreciar en la figura 1, que corresponde a la pantalla de inicio de la aplicación, tenemos una interfaz gráfica de usuario (GUI por sus siglas en inglés) dividida en dos secciones: la primera fila o sección de opciones está asociada a la selección de recursos para una búsqueda utilizando como referencia los datos almacenados en el DI, mientras que la segunda fila corresponde a las búsquedas que pueden realizar los usuarios por términos específicos.

2. RACOFI: Rule-Appling Collaborative Filtering.

3. ISIS-Tutor “es un entorno de aprendizaje inteligente para apoyar el aprendizaje de un lenguaje de formato de impresión de un sistema de recuperación de información CDS/ISIS/ (ISIS para abreviar)” [10].

4. LORM: Learning Object Recommendation Model.

5. AGORA “Ayuda para la Gestión de Objetos Reutilizables de Aprendizaje”.

Learning Object Recommender System

Welcome Jacqueline Solis, this tool guides you through the process of searching for Learning Objects.

Instructional Design

Introducción a bases de datos

Activity

Sesion 1



or

Search or topic

Figura 1. Pantalla de bienvenida para un usuario de la aplicación.

El modelo de recomendación para los profesores con DIs se basa en tres componentes:

- Modelo del curso: Construido a partir de un DI, donde un DI puede estar asociado a varios cursos y un mismo curso puede ser impartido por uno o varios profesores en forma conjunta.
- Los items: Representados por metadatos de OAs que vienen dados en múltiples idiomas y presentan descripciones dadas en lenguaje natural.
- Monitoreo de las transacciones: Donde se implementa una trazabilidad sobre las selecciones y la navegación entre los recursos sugeridos para obtener nuevos descriptores para el modelo o perfil de curso.

Así que para aprovechar estos tres componentes se pasó por varios modelos de implementación, donde se utilizaron desde reglas de asociación hasta funciones de utilidad.

Por otro lado, en las búsquedas sencillas el usuario puede introducir un texto y buscar en los repositorios o colecciones para encontrar aquellos metadatos que tengan alguna similitud con el texto introducido. De modo que en ambos casos se toma como entrada: 1) un texto o conjunto de textos (las descripciones textuales de las actividades) como elemento a buscar dentro de las colecciones, y 2) un repositorio en donde realizar la búsqueda.

La Figura 2 muestra la interfaz de búsqueda para un usuario en una búsqueda sencilla dentro del repositorio de OAI cosechado desde una URL. Esta GUI está dividida en secciones:

- Sección de criterios de búsquedas: donde aparecen las palabras de búsqueda (aquí se despliega la lista de actividades cuando se trata de una búsqueda por DI) y la lista de repositorios o colecciones disponibles.
- Sección de botones: donde aparece la opción de “Búsqueda Local”. Inicialmente existió una opción para “Repetir búsquedas” que funcionaba principalmente para el caso de los repositorios donde se realizan las búsquedas mediante SQL. En este caso, lanza una búsqueda sobre el ROA a partir del último metadato cosechado y los agrega a la lista de los elementos disponibles de forma local. La búsqueda local se realiza sobre los elementos que ya han sido cosechados. En el caso de los metadatos de colecciones cosechadas mediante OAI-PMH todos se encuentran de forma local, ya que la cosecha se realiza sobre la colección completa.

Learning Object Recommender System

Word or text search

WBS

Repository

http://riadme.gnet.gr/riadme-repository/services/oai (

Back to top

Search Results

Text - WBS

Learning Object	Release Date	Collection	URL
<input type="checkbox"/> Structure, description and coding system	1500-01-01	oai	http://winds-app.ft.fraunhofer.de/cgi-bin/WebObjects/windscoai.waa/showContent?id=3678.io
<input type="checkbox"/> exercise 1 roll-up of a Work Breakdown Structure	1500-01-01	oai	http://winds-app.ft.fraunhofer.de/cgi-bin/WebObjects/windscoai.waa/showContent?id=3684.io
<input type="checkbox"/> The Work Breakdown Structure applied to the project study	1500-01-01	oai	http://winds-app.ft.fraunhofer.de/cgi-bin/WebObjects/windscoai.waa/showContent?id=3609.io
<input type="checkbox"/> Example of WBS application	1500-01-01	oai	http://winds-app.ft.fraunhofer.de/cgi-bin/WebObjects/windscoai.waa/showContent?id=3678.content
<input type="checkbox"/> none	1500-01-01	oai	http://winds-app.ft.fraunhofer.de/cgi-bin/WebObjects/windscoai.waa/showContent?id=3682.content

There are 5 records (Showing from 1 to 5)

First page Previous 1 Next Last page

Figura 2. Vista de usuario: Búsqueda sencilla de palabras en un repositorio.

- Sección de resultados: Se despliega un datatable con los resultados de la búsqueda.

Los resultados que se desplegarán al usuario final dependen de los metadatos de los elementos cosechados y del DI. De modo, que están condicionadas por las descripciones textuales disponibles, así como por las descripciones brindadas por el profesor en las actividades del DI.

2.2. Funcionalidades para un administrador

Las opciones para el administrador de momento cuentan con dos secciones de trabajo. Una consiste en la funcionalidad para importar los nombres de los catálogos y los datos para cosecharlos, mientras que la segunda constituye un bloque de funciones de limpieza del esquema de datos. Las opciones de cosecha, como se puede ver en la Figura 3, requieren de un URL que apunta al repositorio a consultar y a partir de ahí, se pueden obtener los nombres de las colecciones que contiene.

Una vez recuperados los nombres de las colecciones, se le muestra al administrador una serie de datos relacionados con los catálogos. Estos incluyen el estatus, el URL al que pertenece el catálogo, una breve descripción del catálogo, la especificación (congruente con el setspec del OAI-PMH) y una sección de notas para almacenar anomalías detectadas durante la cosecha.

El estatus de un catálogo puede ser uno de tres valores:

- Registrado: Indica que el nombre del catálogo ha sido registrado pero no se ha realizado ninguna acción con él todavía.
- En Cosecha: El administrador dió clic sobre el elemento en la tabla, de modo que se inició la cosecha de los metadatos contenidos en esa colección de forma asincrónica.

- **Importado:** Cuando termina de importar los elementos de una colección, la llamada se encarga de cambiar el estado de la colección a “importado”, para indicar que todos los elementos disponibles al momento de la cosecha, fueron recuperados.

Harvesting Options OAI-PMH
Enter the path of the repository to harvest its collection names:

Enter the URL

Catalogs available
The following list shows the catalogs available in the repositories that share resources with OAI registered to data.
To import metadata from a collection, select the desired row with status “Registered” and the system will begin to harvest.
As you finish harvesting the collection will change the status of the catalog to “Importado”.
A collection registered does not mean that its metadata have been harvested. In addition, a collection will appear as available to users until it has been harvested.

Showing 10 records Filter by:

Status	LOR URL	Catalog Name	Specification	Notes
Importado	http://ariadne.gnet.gr/ariadne-repository/services/oai	Metadata originating from ARIADNE	ARIADNE	
Importado	http://ariadne.gnet.gr/ariadne-repository/services/oai	Metadata originating from aua.gr	aua.gr	
Importado	http://ariadne.gnet.gr/ariadne-repository/services/oai	Metadata originating from bmukk.gv.at	bmukk.gv.at	
Importado	http://ariadne.gnet.gr/ariadne-repository/services/oai	Metadata originating from cnx.org	cnx.org	
Importado	http://ariadne.gnet.gr/ariadne-repository/services/oai	Metadata originating from jem.thematic.net	jem.thematic.net	
Importado	http://ariadne.gnet.gr/ariadne-repository/services/oai	Metadata originating from laprof.eu	laprof.eu	
Importado	http://ariadne.gnet.gr/ariadne-repository/services/oai	Metadata originating from nsdl.org	nsdl.org	
Importado	http://ariadne.gnet.gr/ariadne-repository/services/oai	Metadata originating from oai.mace-project.eu	oai.mace-project.eu	
Importado	http://ariadne.gnet.gr/ariadne-repository/services/oai	Metadata originating from oai.members.ariadne.org	oai.members.ariadne.org	
Importado	http://ariadne.gnet.gr/ariadne-repository/services/oai	Metadata originating from open.jorum.ac.uk	open.jorum.ac.uk	

There are 18 records (Showing from 1 to 10) Previous Next

Cleaning Options
The following options allow clear all harvested records.

- Cleaning SQR's elements includes profiles of the activities and all elements preprocessed.
- Cleaning OAI-PMH's elements attempts to delete only metadata elements harvested via OAI-PMH.
- Stored collections identifiers can overlap.
- SQR repositories's validation would be enabled or disabled according to the valid keys generator service and harvesting service responses.

Figura 3. Vista del administrador: Registro de ROAs

Los datos asociados a la descripción del catálogo, el URL y la especificación se recuperan mediante métodos definidos en el OAI-PMH. Mientras que el espacio para notas

fue agregado para capturar detalles como las cosechas sin resultados obtenidos o los errores durante la recuperación de metadatos.

Las opciones de limpieza disponibles para el administrador fueron creadas para facilitar los cambios entre los mecanismos de recuperación de OAs (por lo que no son tareas de uso frecuente ni siquiera para los administradores) aunque resultan necesarias al realizar las pruebas asociadas a control de calidad vigentes en el TecDigital. La descripción de las funcionalidades implementadas es la siguiente:

- “Borrar registros SQR”: Toma todos los elementos cosechados mediante SQR y los elimina del esquema. Además, reinicia las secuencias que generan las llaves.
- “Borrar registros OAI-PMH”: Toma los IDs de las tablas de cosechas (elementos cosechados mediante OAI-PMH) y los elimina de la lista de metadatos preprocesados. Dado que pueden existir intersecciones entre los elementos cosechados mediante SQR y OAI-PMH, no inicializa las secuencias generadoras de los identificadores.
- “Borrar registros (Todos)”: Actualmente está deshabilitada. Sería una operación en conjunto entre el Borrado de SQR y de OAI-PMH más la reinicialización de las secuencias generadas para las tablas de OAI-PMH
- “Validar disponibilidad de interfaces SQR”: Toma los servicios web que generan las llaves para conexión anónima y los de consulta de registros para probar que retornen una respuesta. De no ser así, marca el repositorio como inactivo en la base de datos para que no le aparezca más al usuario al realizar consultas.
- “Detener cosechas”: Detiene todos los procesos que se encuentren activos “En Cosecha”. Para esto, marca los repositorios como “Importados” para evitar que ingresen a la siguiente iteración de recopilación de recursos.

3. Metodología de Investigación

Como metodología de investigación se trabajó por fases. Primero se realizó un (1) análisis de los datos disponibles para trabajar, seguidamente se (2) modeló un posible esquema de mapeo de textos y se (3) construyó una implementación para generar la recomendación.

Los pasos (2) y (3) se repitieron en varias ocasiones para corregir omisiones en el modelo y fue necesario modificar en múltiples ocasiones los mecanismos de recuperación de ítems. Para agregar estas mejoras en las cosechas y algoritmos como se describe a continuación.

3.1. Análisis del contexto y los datos

El TEC, al ser una institución de educación superior, requiere de un alto volumen materiales educativos para satisfacer la demanda de conocimiento de los estudiantes.

De modo que aprovechar los recursos generados por otros profesores y profesionales representa una necesidad.

En este sentido, los repositorios de objetos de aprendizaje constituyen una oportunidad ante la creciente demanda de educación en todos los países, la capacidad limitada de la enseñanza tradicional, el esfuerzo y costo que implica la construcción de materiales de aprendizaje multimedia [14]. En la actualidad existen múltiples iniciativas para estandarizar los OA y convertirlos en elementos portables entre las plataformas y muchas de estas iniciativas se basan en el uso de metadatos para describir el contenido de cada ítem. Dos de las iniciativas más importantes en cuanto a estándares para intercambiar y recuperar información digital son LOM-IEEE y Dublin Core (DC), aunque existen otros formatos con amplia difusión como los son MODS, MARC_XML y ORE con especial difusión en el ambiente bibliotecario.

En nuestro caso particular, nos interesan las características pedagógicas de los OAs para determinar la población hacia la que están enfocados, así como los tipos de datos involucrados, los niveles de dificultad entre otras para intentar incorporarlas al proceso de recomendación.

El estándar de LOM-IEEE busca "facilitar la búsqueda, evaluación, adquisición, y uso de los objetos de aprendizaje por parte de los estudiantes o instructores o procesos automáticos de software" [1], mediante una serie de descriptores que abarcan desde aspectos técnicos hasta características pedagógicas de los OAs. Por otro lado, DC es un estándar formal de intercambio de recursos web y se concentra en quince aspectos técnicos y de contexto del recurso a compartir [15].

Por la necesidad de los descriptores pedagógicos, la posibilidad de mapear elementos especificados en LOM-IEEE sin perder información hacia DC (pero no a la inversa), y por la existencia de interfaces en SQI y OAI-PMH en Ariadne (que forma parte del Global Consortium) disponibles para la experimentación, seleccionamos LOM-IEEE como el esquema conceptual inicial a utilizar para recuperar las estructuras de los metadatos que describen a los OAs.

3.2. Recopilación de Datos

Durante el desarrollo de este proyecto se han utilizado tres mecanismos asociados a la recopilación de los datos de los items, los cuales se mencionan a continuación.

La primer versión de la aplicación se denominó "Prototipo de asistente inteligente para sugerir actividades en el diseño instruccional", y utilizó conceptos básicos de minería de datos para asignar los pesos a las posibles reglas de asociación de los metadatos de las unidades de aprendizaje disponibles en el LMS.

Para esta implementación se utilizó como base el esquema de extracción de conocimientos propuesto en la patente 7113917 para "realizar recomendaciones personalizadas de items representadas dentro de una base de datos" [16], y tras la experimentación con la asignación de pesos y la selección de las palabras claves en este primer modelo, se identificaron elementos que debían ser modelados nuevamente para ajustarse a las necesidades del LMS y a la evolución de los

proyectos que proveen los insumos de datos.

En una segunda versión del núcleo, se utilizó una serie de interfaces de Simple Query Interface (SQI) para realizar las consultas sobre ROAs. Esta implementación funcionó de forma efímera al generar las recomendaciones y se basó en el principio de realizar cosechas solo de los metadatos de los objetos que respondían a los textos extraídos de las actividades, para aprovechar las características del PLPQLv1 implementada en algunos repositorios.

Como se puede apreciar en la figura 4, el SQI es un API

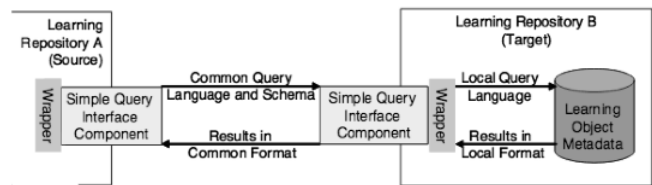


Figura 4. Intercambio de información entre dos repositorios usando SQI. Tomado de [17].

de comunicaciones donde se utilizan una serie de métodos tanto sincrónicos como asincrónicos; donde "los métodos propuestos pueden ser utilizados por una fuente para la configuración y presentación de consultas a un sistema de destino y recuperar resultados de ella" [17].

Para utilizar este API, se necesita tener un identificador de sesión, que puede ser registrada o anónima. En nuestro caso, la implementación se llevó a cabo utilizando sesiones anónimas y consultas sincrónicas para recuperar los metadatos de los OAs dados en LOM.

Este mecanismo para recuperar items presentó dos líneas principales de problemáticas, la primera estaba asociada al tiempo de respuesta, ya que dependía directamente de la respuesta de los servicios web del cliente para SQI y del manejo de los modelos vectoriales para representar a los documentos. El segundo se dió a partir de la ausencia de interfaces SQI disponibles para realizar las consultas, ya que a partir de octubre del 2014 ninguno de los servicios web conocidos brindaba el servicio de cosechas.

Finalmente, la tercera versión de recuperación se basa en la utilización de OAI-PMH para cosechar colecciones completas desde repositorios en la nube y posteriormente se preprocesan en el servidor de aplicación para analizar la similitud con las actividades.

La iniciativa de OpenArchives busca proveer un framework de interoperabilidad basado en cosecha de metadatos y que se denomina Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH); donde hay dos tipos principales de participantes del proceso: los proveedores de datos que son los administradores del sistema que ofrecen soporte para el OAI-PMH y administran los repositorios y los proveedores de servicios, quienes usan los metadatos recolectados vía OAI-PMH como una base para crear servicios con valor agregado [18].

Esta última implementación es la que se está utilizando actualmente en la aplicación y presenta problemas con el

rendimiento del servidor principal durante la cosecha de metadatos, debido al volumen de ítems de las colecciones.

3.3. Preprocesamiento de información

Para el preprocesamiento de los metadatos de los ítems se utilizaron dos fases: la primera utilizando funciones del xml-bus y XPath⁶ para segmentar los nodos de cada OA con el fin de extraer los ítems que fueron considerados como significativos para describir un OA; y una segunda fase, en la que se utilizan funciones de PostgreSQL para extraer los textos (en múltiples idiomas) que van a ser comparados con las descripciones de las actividades.

El esquema de LOM está compuesto por una serie de meta-elementos que describen características del OA y se agrupan por categorías: “general”, “life-cycle”, “meta-metadata”, “educational”, “technical”, “educational”, “rights”, “relation” y “annotation” [1]. Sin embargo estas estructuras no son estrictas en su mayoría y esto provoca que una buena parte de las descripciones estén ausentes o colapsadas sobre el nodo padre; lo que dificulta su extracción y análisis.

3.4. Análisis de similitudes e hibridación

De un modo similar al proceso de recuperación de datos, el modelo de análisis de similitudes y de hibridación para generar la recomendación ha pasado por una serie de iteraciones en busca de mejoras en la calidad de las recomendaciones.

En primer lugar es necesario distinguir entre dos orígenes de información:

- Diseño instruccional: Es la representación en lenguaje natural de la organización de las lecciones.
- Metadatos de los OAs: Constituyen los ítems a utilizar en el sistema y se extraen a partir de representaciones en LOM-IEEE. Su contenido es facilitado por quienes suben los OAs a los repositorios y están dados en lenguaje natural en múltiples idiomas.

Los datos aportados por el DI aportan los descriptores textuales de la lección y a partir de ellos se automatiza la construcción de los perfiles de cursos. Por otro lado, metadatos de cada OA aportan las descripciones de los ítems, los cuales son evaluados mediante modelos vectoriales para establecer la similitud que comparten con las descripciones de cada actividad.

El primer modelo que se utilizó para intentar realizar una correspondencia entre los DIs y los ítems, se basó en reglas de asociación extraídas a Priori, e incorporó reducciones a partir de las medidas de soporte y confiabilidad obtenidas. En este primer enfoque, se trabajó sobre unidades de aprendizaje almacenadas directamente sobre el LMS y dichas unidades fueron generadas a partir de exeLearning con metadatos introducidos manualmente en español; sin embargo no era un esquema de trabajo sostenible y los tiempos de respuesta no eran deseables en una aplicación

web con múltiples usuarios

El segundo modelo utilizado incorporó la base de las estructuras para la medición de similitudes. Se estableció la aritmética vectorial como mecanismo para comparar los descriptores de las actividades y de los ítems.

La ventaja fundamental de este modelo consistió en la incorporación del SQI como mecanismo para obtener los ítems desde ROAs en la nube, aunque su análisis se realizaba de forma efímera. En cuanto a la construcción de los perfiles de los cursos, se incorporó el uso del TF-IDF y de la similitud de cosenos durante la extracción de los descriptores y seguidamente, se integraron en el proceso de extracción de la utilidad de cada ítem. Además, se establecieron las primeras reglas de reducción de textos personalizadas para nuestro contexto y se intentó incorporar el uso de las medidas de similarity y Levenshtein disponibles en el motor de base de datos para aplicarlos en la corrección ortográfica y reducciones de cardinalidad del conjunto de descriptores disponibles.

Entre las desventajas que se encontraron en esta iteración del modelo podemos citar el tiempo de respuesta. Las búsquedas y el preprocesamiento se realizaba al momento de disparar una consulta y al mapear los datos a un modelo vectorial se encontraron similitudes de coseno iguales a cero. Posteriormente detectamos que los ROAs aplican un proceso de radicación con los textos que le son facilitados en las búsquedas mediante SQI, de modo que el hecho de que un OA fuera retornado bajo un texto de búsqueda no implica que el término contenido dentro de sus metadatos sea el mismo solicitado.

Un tercer modelo para generar la recomendación incorporó el algoritmo RAKE [19] para reducir la cardinalidad de los textos a analizar en los metadatos de los OAs cosechados, sin embargo quedó pendiente un proceso de detección de idioma para mejorar la extracción de las palabras claves y frases.

4. Análisis de Resultados

Una de las características por las que fue elegido LOM-IEEE como estándar está asociado a la disponibilidad de metadatos con información de tipo educativo, que resultan sumamente valiosos al categorizar los contextos y perfiles de quienes los van a utilizar.

Al realizar un análisis de los metadatos cosechados, se encontró que la gran mayoría de ellos se encuentran incompletos, de modo que no es posible aprovecharlos en la medida que se planeó al realizar la selección de LOM.

En esta sección se exponen las características de una muestra de metadatos de 109865 OAs; de los cuáles 86031 presentan algún tipo de metadato registrado bajo la etiqueta educacional (un 78.3 % de la muestra).

Además, se expondrán las características de los metadatos de los catálogos de una muestra de 2993 URLs que apuntan a diversas interfaces en OAI-PMH. Sobre esta muestra se mapearon los metadataPrefix con el fin de determinar el porcentaje de ellos que es aprovechable para el modelo de funcionamiento propuesto en nuestra solución.

6. XPath: <https://www.postgresql.org/docs/8.4/static/functions-xml.html>

4.1. Metadatos de colecciones cosechadas

Los metadatos utilizados para esta evaluación fueron cosechados de uno de los servicios de OAI-PMH de Ariadne⁷ y vienen dados en LOM-IEEE. Además, su preprocesamiento fue realizado mediante una mezcla de funcionalidades en tcl mediante el paquete tDom⁸, XPath 1.0 en PostgreSQL 8.4 y tareas específicas de extracción construidas en R⁹.

La categoría de metadatos educativos describe un OA en término de sus características pedagógicas [1], sin embargo, la disponibilidad de los mismos varía de un ROA a otro. En nuestro caso particular, durante el intento de la aplicación de los metadatos educativos al proceso de recomendación encontramos que la existencia o no de las etiquetas no implica que sean significativos o útiles para el proceso de selección.

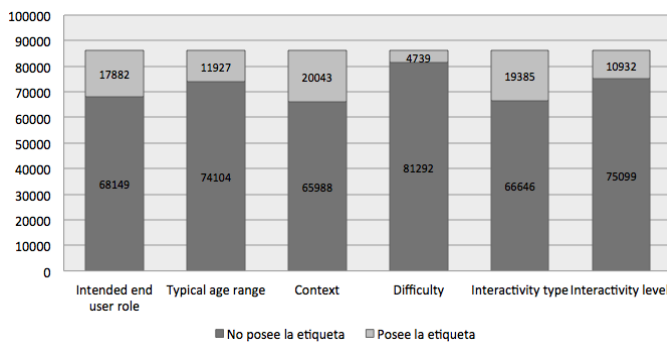


Figura 5. Frecuencia de Aparición de etiquetas de la categoría Educativa.

La Figura 5 muestra la cantidad de apariciones de las distintas etiquetas que componen el grupo de metadatos educativos en LOM. Estos elementos fueron obtenidos de 86031 OAs en los que se identificó al menos una de estas etiquetas, los cuales representan un 78.3% de la muestra total sobre la que se ha probado la propuesta de implementación.

A pesar de que se puede apreciar una amplia difusión en estos metadatos, en una revisión en la que se incluyó la frecuencia de aparición de cada uno de los valores específicos de las etiquetas, se encontró que a pesar de estar referenciadas, en la gran mayoría de los casos existe una etiqueta sin contenidos descriptivos reales que se puedan aprovechar. Otra situación detectada, es que a pesar de que en los valores de la etiqueta se está aportando un dato correcto y especificado dentro del LOM-IEEE, este queda sesgado al criterio de quien cargó el OA en el ROA; de modo que la realidad bajo la que fue etiquetado un OA (por ejemplo en complejidad) no va a ser la misma que para un sistema educativo en otra región.

7. OAI-PMH de Ariadne utilizado: <http://ariadne.grnet.gr/ariadne-repository/services/oai>

8. tDom: <http://wiki.tcl.tk/1948>

9. R: <https://www.r-project.org/>

Tabla 1: Diccionario de datos para Tabla resumen de Metadatos Educativos

ID Source	Significado
1	lomv1.0
2	lrev3.0
Otros	Agrupar las opciones minoritarias.
-	No disponible / colapsado sobre el nodo padre.

La Tabla 1 presenta el diccionario de datos utilizado para la columna "Source" de la tabla 2: *Resumen de los metadatos educativos encontrados*. En la tabla 1 se muestran los identificadores seleccionados para los valores de mayor frecuencia de aparición dentro de la etiqueta "Source" y se ofrece una breve descripción del contenido de los mismos. Los elementos identificados bajo el ID de "Otros" corresponden a una agregación de los elementos restantes que fueron encontrados en dicha columna.

Otro elemento rescatable es el ID "-", bajo el cuál se agruparon los elementos que no contenían ningún valor dentro de las etiquetas del grupo, o bien, poseen el contenido colapsado dentro de su nodo padre.

Tabla 2: Resumen de los metadatos educativos

Learning Source Type			Interactivity Type		
Source	Value	Freq.	Source	Value	Freq.
-	-	49933	-	-	69681
1	other	20821	1	active	546
1	narrative	8788	1	expositive	34887
1	figure	5325	1	mixed	608
1	none	4096	1	none	4073
1	open activity	4083	1	undefined	28
Otros		16819	-	expositive	42

Intended End User Role			Typical Age Range		
Source	Value	Freq.	Language	Value	Freq.
-	-	71176	-	-	77089
1	learner	37316	x-none	14-18	20806
1	teacher	1132	x-t-lre	22-u	4149
1	manager	93	en	18-u	3214
2	other	68	-	18-u	1511
-	learner	42	-	20-25	560
Otros		38	Otros		2536

Context			Difficulty		
Source	Value	Freq.	Source	Value	Freq.
-	-	69079	-	-	105132
1	bve	20806	1	difficult	478
1	Higher education	11645	1	easy	1051
1	continuing education	4149	1	medium	3067
2	vocational education	1063	1	difficult	10
2	library	696	1	very easy	127
Otros		2427			

En la Tabla 2 se muestra un resumen del contenido real almacenado dentro de cada una de las etiquetas de metadatos asociadas al grupo "Educational". Dicha tabla

está compuesta por seis agrupaciones correspondientes a los metadatos de “Learning Resource Type”, “Interactivity Type”, “Intended End User Role”, “Typical Age Range”, “Context” y “Difficulty”. Para cada una de estas subetiquetas se seleccionaron los siete pares de valores con la mayor cantidad de apariciones a lo largo del subconjunto de datos seleccionado.

Es importante rescatar que para todas las categorías recopiladas en esta tabla el valor predominante es la ausencia de datos, de modo que su existencia no aporta información al proceso de recomendación.

Otro hallazgo interesante está relacionado con la contribución en la elaboración de los OAs. En este aspecto, encontramos que sólo el 34.7% de la muestra contiene alguna etiqueta asociada a este rubro, para un total de 38116 OAs.

En este subconjunto de datos, se seleccionaron todos los elementos que aparecen en bajo la etiqueta “Contribute” de los metadatos de los OAs, de modo que pueden contabilizarse más items debido a que la especificación de LOM-IEEE permite la definición de uno o más elementos con esta etiqueta en distintas secciones de los metadatos.

Tabla 3: Roles registrados bajo la etiqueta “Contribute”

Id	Rol	Frecuencia
1	Autor	34973
2	Publisher	1415
3	Creator	1107
4	Desconocido	353
5	Content provider Validator	154
6	Editor	91
7	Technical Implementer	6
8	Subject matter expert	3
9	Technical Validator	2
10	Initiator	2
11	educational validator	10

La Tabla 3 contiene la frecuencia de aparición de los roles registrados en aquellos metadatos que poseen referencias bajo la etiqueta “Contribute”. El rol etiquetado como “Desconocido” centraliza los registros en los que aparecen nombres o datos dentro del campo, pero sin ninguna referencia a un rol en particular, o bien, están etiquetados como “unknown”.

La presencia o ausencia de un autor detrás de un OA debería ser un indicador de la calidad del mismo, ya que nos brinda una referencia del responsable de la creación y validación de su contenido y sus descripciones.

Bajo esta premisa, se decidió realizar un análisis de componentes principales (PCA, por sus siglas en inglés) para visualizar si existe alguna correlación entre las descripciones de los metadatos, el título y la existencia de algún contribuyente asociado al OA.

Para la validación de esta hipótesis se definieron las siguientes variables:

- autor: Identificador de los roles de contribución asociados.
- urls_tech: Cantidad de URLs para el OA registradas

bajo la etiqueta Technical.

- urls_descrip: Cantidad de URLs para el OA registradas bajo alguna descripción del OA.
- anno: Año para el que existe algún registro de colaboración para el OA.
- cl_descrip_general: Cantidad de letras existentes en la descripción contenida en el metadato General.
- cl_descrip_clasificacion: Cantidad de letras existentes en la descripción contenida en el metadato Classification.
- cl_descrip_educational: Cantidad de letras existentes en la descripción contenida en el metadato Educational.
- pd_general: Cantidad de palabras existentes en la descripción contenida en el metadato General.
- pd_clasificacion: Cantidad de palabras existentes en la descripción contenida en el metadato Classification.
- pd_educational: Cantidad de palabras existentes en la descripción contenida en el metadato Educational.
- cl_titulo: Cantidad de letras existentes en el título del OA
- palabras_titulo: Cantidad de palabras existentes en el título del OA.

Como se puede observar en la Figura 6, al realizar un PCA de estas variables se obtuvo un 57.66% de inercia en los datos. Además, el círculo de correlaciones nos indica que no existe ningún tipo de correlación entre las descripciones brindadas en el grupo de Classification, General y Educational; sin embargo, sí existe una correlación fuerte y positiva entre la cantidad de palabras y letras del título del OA y la descripción brindada en el grupo de información General.

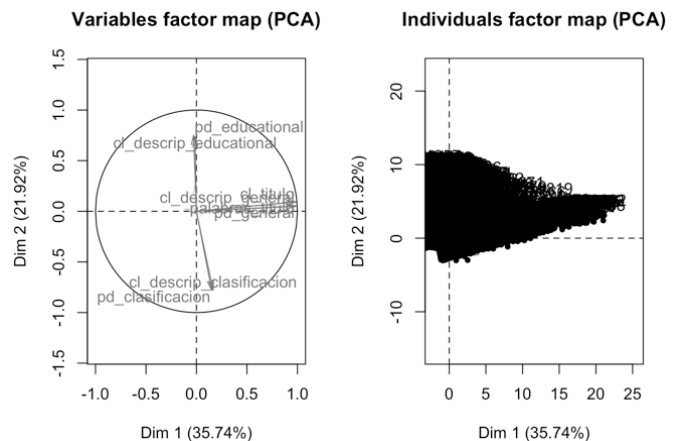


Figura 6. Análisis de componentes principales para las descripciones contenidas en los metadatos.

En el mismo círculo se excluyeron cuatro variables al resultar poco representativas (menos de 10% de representación) las cuales fueron: la cantidad de URLs ubicadas tanto en la sección Técnica como en el contenido de los metadatos, el año y los roles de contribución. Por esta razón, tomar en

cuenta la existencia de contribuyentes no es un parámetro con representación significativa para esta muestra de datos. La exclusión de las URLs nos plantea una nueva interrogante, debido a que nuestro proceso de cosecha se implementó sólo sobre los metadatos y si no se dispone de un URL asociado no es posible recuperar el recurso descrito. Finalmente, al graficar los individuos se encontró que la mayoría de los metadatos de los OAs analizados poseen valores en los rubros asociados al título y la descripción general; mientras que existe una pequeña porción de ellos que contiene descripciones en el grupo de clasificación y también existe un grupo importante que posee tanto descripciones generales y educativas, mas no de tipo clasificatorio. En resumen, la existencia de las características pedagógicas dentro de las descripciones en LOM podría llegar a constituir una herramienta valiosa en la medida que los autores y creadores de OAs se preocupen por realizar un etiquetado adecuado de los mismos. Sin embargo, en la actualidad aún no es del todo aprovechable en los sistemas de información.

4.2. Metadatos de colecciones no cosechadas

Una de las problemáticas detectadas durante las fases de control de calidad al trabajar con OAI-PMH es que muchas de las colecciones recuperadas al registrar un repositorio nuevo no agregaba elementos al banco de recursos en LOM-IEEE. Por este motivo se inició una búsqueda de posibles causas que inciden en este comportamiento y se procedió a realizar una revisión minuciosa de los URLs con los que contamos para la experimentación.

Tras detectar que la causa estaba asociada al metadataPrefix de los metadatos que intentamos cosechar, nos dimos a la tarea de explorar una muestra de 2993 URLs que apuntan a diversas interfaces en OAI-PMH (no todos activos o con las interfaces para compartir recursos activas) disponibles en el sitio oficial de OpenArchives¹⁰.

De dicha muestra, 1233 URLs no retornaron ningún metadataPrefix disponible al realizar la consulta mediante el verbo "ListMetadataFormats" de la especificación del OAI-PMH, o bien no fue posible acceder a la URLs. Esto representa aproximadamente un 41.20% de repositorios inactivos o sin respuesta.

Tras cuantificar y normalizar los metadatos de las colecciones, se identificaron un total de 159 tipos distintos de metadataPrefix en los repositorios consultados y se encontró que la mayoría de las URLs con respuesta posee entre 3 y 6 metadataPrefix distintos en sus colecciones, alcanzando un pico de 868 URLs con 5 metadataPrefix distintos en sus colecciones.

La Figura 7 muestra que los 20 tipos más comunes de metadataPrefix están asociados a dublin core (oai_dc, qualified-dublin-core, simple-dublin-core, qdc, etc), marc (oai_marc, marcxml, marc), rfc (rfc1807), nlm, mets, rdf, didl, uketd_dc, ore, mods, etdms (oai_etdms, etdms), dim, xoai y juni2. Donde, dublin core y marc aparecen contabi-

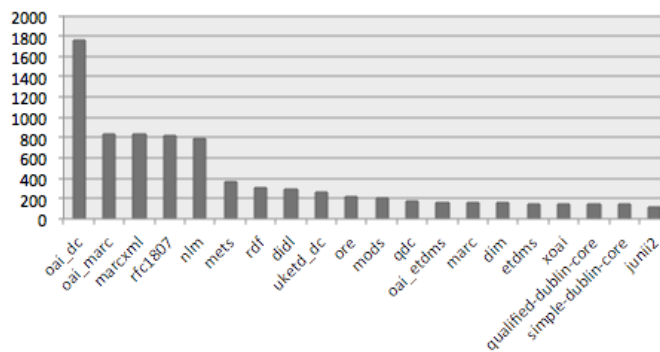


Figura 7. Top 20 metadataPrefix en la muestra

lizados más de 2040 y 1836 veces, respectivamente. Además, al realizar una clasificación manual para determinar el número exacto de apariciones de cada tipo de metadataPrefix (Ver Figura 8), encontramos que el 39% de las colecciones registradas están asociadas o poseen esquemas de metadatos basados en dublin core, un 21% están asociadas a marc y solamente un 0.33% de la muestra poseen colecciones con LOM-IEEE (29 apariciones de 8727 registros de catálogos).

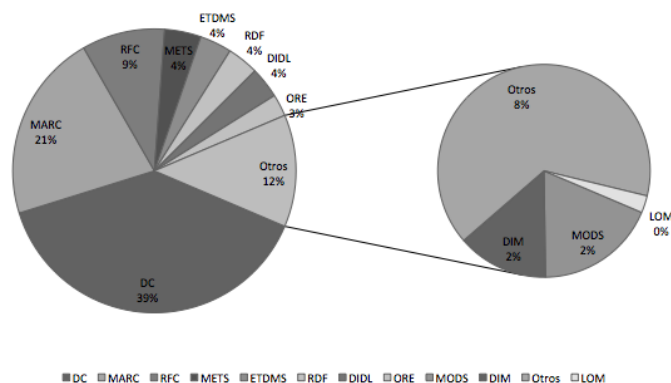


Figura 8. Porcentaje de aparición de los elementos de cada familia de metadataPrefix.

Esta diferencia en la tendencia de los metadatos, al menos en la muestra seleccionada, resulta crítica para nuestra propuesta; ya que, al vincularnos solamente con LOM-IEEE estamos perdiendo la oportunidad de aprovechar millones de recursos que se encuentran descritos en otros formatos y podrían ser útiles para las lecciones.

Esta situación se replica en los metadatos del Repositorio Nacional de Costa Rica (kimuk), el cuál ofrece soporte a seis formatos de metadatos¹¹: uketd_dc, dim, oai_dc, marc, etdms, xoai, qdc, rdf, ore, mods, mets y didl y actualmente no podemos aprovechar en el recomendador de objetos de aprendizaje por el formato de los metadatos.

10. Lista de repositorios disponible en <https://www.openarchives.org/Register/BrowseSites>

11. Obtenido en <http://kimuk.conare.ac.cr:8090/xoai-2.0/request?verb=ListMetadataFormats>

4.3. Consideraciones para la precisión de las recomendaciones

La propuesta inicial de implementación de un esquema de recomendaciones híbrido para los metadatos de OA utilizando componentes basados en contenidos y colaborativos resulta funcional, sin embargo, los componentes de la función de utilidad requieren reevaluaciones importantes.

Como se mencionó en la sección 4.1, la existencia de etiquetas con las características pedagógicas de los OA no implica que sean utilizables y en la sección 4.2 nos quedó claro que estas características tampoco son la tendencia en cuanto a la exposición de los recursos de aprendizaje (al menos para la muestra a la que tuvimos acceso).

Las precisiones alcanzadas en los catálogos donde los metadatos vienen dados en múltiples idiomas (y el español no es uno de los prioritarios) para un DI dado en español tienden al 5%. De hecho, se puede considerar que las asociaciones de los contenidos en este tipo de colecciones resultan de la casualidad ante la existencia de los términos con escritura similar buscados en otro idioma, o bien, que están contenidos dentro de una palabra. Por ejemplo el término “ER”, que puede estar haciendo referencia al modelo “Entidad Relación” en el contexto de Tecnologías de Información en español, al “Emergency Room” en el contexto médico en inglés o al pronombre “Er” en alemán.

Estas medidas de precisión varían entre los catálogos disponibles y mejoran al realizar las búsquedas sobre metadatos en español y en el contexto apropiado. Además, a través de la navegación y evaluación de los recursos sobre la herramienta, se ajusta progresivamente la función de utilidad que va refinando la recomendación.

Otro factor que demostró tener gran peso sobre la calidad de las recomendaciones es la precisión con la que un profesor puede describir los contenidos que espera cubrir durante su lección. Este componente es el que marca la calidad del perfil que se va a construir y a partir de él se deriva todo el proceso de recomendación.

La validación de la precisión de las recomendaciones sobre un entorno de producción es una de las tareas que están pendientes a corto plazo, sin embargo ya se ha visto que las descripciones concretas que incluyen conceptos de contextos específicos retornan mejores resultados que aquellas dadas en un lenguaje coloquial.

5. Conclusiones

La experiencia de implementación del Recomendador Híbrido de Objetos de Aprendizaje generó conocimiento sobre la utilidad teórica y práctica de los metadatos, los procesos de recomendación en entornos de educativos y la integración de múltiples herramientas para generar soluciones integrales que ha sido compilada en las siguientes conclusiones.

En cuanto a los mecanismos de cosecha, podemos decir que a pesar de los beneficios que aportan las consultas sobre los ROAs mediante SQI, la inestabilidad de las interfaces las descarta como mecanismo eficiente para ser tomado en

cuenta en entornos de producción o de alta concurrencia. Por otro lado, el uso de OAI-PMH, a pesar de tener desventajas como el desconocimiento de la cardinalidad de cada catálogo al iniciar una cosecha, ofrece un mecanismo estable, constante y estándar para compartir metadatos.

Con respecto al análisis realizado sobre los metadatos de los recursos cosechados, tenemos que rescatar que a pesar del esfuerzo realizado por incorporar las etiquetas de los descriptores aún no es posible utilizarlos de forma estándar y su contenido no siempre es aprovechable por los sistemas informáticos. Además, existen sesgos asociados a la subjetividad de quién carga el OA en los repositorios e introduce su descripción que pueden afectar los procesos de recomendación.

La presencia o ausencia de uno o más autores en los metadatos de un OA no es un indicador de la calidad de las demás descripciones del recurso. Esta etiqueta obtuvo menos de un 10% de representación al ser evaluada dentro de la muestra de datos y la mayoría de los autores sólo aportan las descripciones bajo la etiqueta General de la especificación de LOM-IEEE.

Para la muestra de datos analizada se encontró una correlación fuerte y positiva entre la cantidad de palabras y letras existentes en el título y la descripción del OA bajo la etiqueta General; esto se puede interpretar como que en una buena parte de los OAs tanto la descripción como su título tiene prácticamente la misma longitud (y posiblemente el mismo contenido).

También bajo esta misma muestra, menos de un 10% de los metadatos incluían URLs para acceder a los OAs dentro de la información técnica o en las descripciones de recurso. Esto implica que al realizar cosechas sólo de los metadatos estamos perdiendo el acceso al objeto en sí y esto genera ruido en la base de datos. Además, plantea la necesidad de implementar otra interface que permita traer el OA junto con su descripción para evitar las inconsistencias.

Al analizar la distribución de distintos esquemas de metadatos en múltiples ROAs se hace más que evidente que es necesario eliminar las dependencias de nuestro recomendador con LOM-IEEE y que es necesario enfocarlo en un esquema que sea escalable a otros estándares como Dublin Core y MARC. Como parte de estos resultados se encontró que alrededor de 0.33% de los catálogos que fueron consultados poseían LOM-IEEE, contra un 39% para DC, 21% para Marc y 9% para el RFC. Además, esto nos permitiría aprovechar los recursos almacenados en el Repositorio Nacional de Costa Rica.

Además, a pesar de que se han obtenido pobres precisiones al realizar búsquedas sobre ROAs con múltiples idiomas, es un hecho conocido que estas precisiones mejoran significativamente al realizar búsquedas en español, con una buena descripción dada en el DI y el entrenamiento del sistema. De modo que está pendiente la validación del mismo en un entorno apropiado de búsquedas como lo sería el ROA La Flor, el cuál contiene metadatos en español que eliminarían el sesgo idiomático en una nueva secuencia de pruebas.

Como parte de las propuestas para intentar mejorar los procesos de recomendación y el etiquetado específico de los

nuevos OAs generados en el TEC; se está trabajando para implementar una tropicalización de LOM-IEEE con los elementos mínimos necesarios que puedan aportar información útil y concisa tanto para el ROA Institucional como para el recomendador; del mismo modo, se realizó un censo para establecer las etiquetas aceptables y sus significados dentro de un marco común de trabajo. De este modo, dicha adaptación permitirá refinar las búsquedas al menos en los recursos generados por la misma institución.

5.1. Trabajo futuro

En cuanto a trabajos futuros existen múltiples áreas de acción y aplicación de los diversos componentes asociados al recomendador.

Es necesario cambiar el proceso de cosecha de metadatos para incorporar un ente intermediario que se encargue de la recuperación constante de recursos y sea externo al servidor de aplicación del LMS. Las pruebas de rendimiento que se han realizado a la fecha, reflejan problemas con el manejo de estos dos elementos en el mismo servidor.

El establecimiento dinámico y automático de contextos e idiomas para los OAs es otro elemento en el que se debe trabajar. Este componente va más allá de las descripciones que se puedan aportar mediante LOM-IEEE y las soluciones propuestas deben ser portables a los diferentes esquemas de metadatos existentes con el fin de garantizar la escalabilidad de las nuevas soluciones.

También se espera construir adaptaciones para generar recomendaciones a los modelos de usuario de los estudiantes, los cuáles representan sus conocimientos y necesidades, y posteriormente incorporar rutas adaptativas durante la sugerencia de materiales a los estudiantes de un curso.

Otra utilidad que se desea incorporar es un mecanismo de evaluación automática de la calidad en los metadatos, pero para eso primero se debe partir de los estudios y trabajos que han sido realizados por otros investigadores mediante diversas rúbricas para establecer las características de completitud y contenidos que aplican para nuestras necesidades específicas.

En cuanto a las validaciones pendientes, se esperan realizar a corto plazo las pruebas formales de rendimiento en un entorno de producción con más usuarios concurrentes del sistema y las de precisión con un plan piloto entre profesores de la institución.

En cuanto a las mejoras en la funcionalidad a nivel de usuario, se desea agregar funcionalidades que le permitan al usuario editar los textos de los perfiles generados para las actividades para agregar o quitar etiquetas. Del mismo modo, si se les permite editar los metadatos de los OAs facilitaría la tarea de selección de recursos y mejoraría la precisión de las recomendaciones. Otra mejora es la posibilidad de integrar los URLs de los OAs directamente a los cursos de la plataforma.

Referencias

- [1] LTSC.IEEE, "Draft standard for learning object metadata," Jul. 2002. [Online]. Available: http://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf
- [2] G. Redeker, "An educational taxonomy for learning objects," in *The 3rd IEEE International Conference on Advanced Learning Technologies, 2003. Proceedings, 2003*, pp. 250–251.
- [3] B. Bannan-Ritland, N. Dabbagh, and K. Murphy, "Learning object systems as constructivist learning environments: Related assumptions, theories, and applications," *The Instructional Use of Learning Objects: Online Version. Consultado el*, vol. 24, 2000.
- [4] "Ims global learning consortium," <http://www.imsglobal.org/>. [Online]. Available: <http://www.imsglobal.org/>
- [5] F. Ricci, L. Rokach, and B. Shapira, *Introduction to recommender systems handbook*. Springer, 2011.
- [6] CONARE, "Kimuk repositorio nacional de costa rica," <http://kimuk.conare.ac.cr/Contents/Home?section=networks>, 2016, accessed: 2016-04-26.
- [7] N. Manouselis, H. Drachsler, R. Vuorikari, H. Hummel, and R. Koper, "Recommender systems in technology enhanced learning," in *Recommender systems handbook*. Springer, 2011, pp. 387–415.
- [8] H. Drachsler, K. Verbert, O. C. Santos, and N. Manouselis, "Panorama of recommender systems to support learning," in *Recommender systems handbook*. Springer, 2015, pp. 421–451.
- [9] M. Anderson, M. Ball, H. Boley, S. Greene, N. Howse, S. McGrath, and D. Lemire, "Racofi: A rule-applying collaborative filtering system," 2003.
- [10] P. Brusilovsky and L. Pesin, "Adaptive navigation support in educational hypermedia: An evaluation of the isis-tutor," *Journal of computing and Information Technology*, vol. 6, no. 1, pp. 27–38, 1998.
- [11] K. H. Tsai, T. K. Chiu, M. C. Lee, and T. I. Wang, "A learning objects recommendation model based on the preference and ontological approaches," in *Advanced Learning Technologies, 2006. Sixth International Conference on*. IEEE, 2006, pp. 36–40.
- [12] A. Zapata, V. H. Menendez, M. E. Prieto, and C. Romero, "A hybrid recommender method for learning objects," *IJCA proceedings on design and evaluation of digital content for education (DEDCE)*, vol. 1, pp. 1–7, 2011.
- [13] A. Z. González, D. M. E. P. Méndez, and D. C. R. Morales, "Modelo híbrido de recomendación de objetos de aprendizaje."
- [14] "Globe alliance," 2012. [Online]. Available: <http://www.globe-info.org/>
- [15] D. C. M. Initiative *et al.*, "Dublin core metadata element set, version 1.1," 2012.
- [16] J. A. Jacobi, E. A. Benson, and G. D. Linden, "Personalized recommendations of items represented within a database," Sep. 26 2006, uS Patent 7,113,917.
- [17] B. Simon, D. Massart, F. Van Assche, S. Ternier, E. Duval, S. Brantner, D. Olmedilla, and Z. Miklós, "A simple query interface for interoperable learning repositories," in *Proceedings of the 1st Workshop on Interoperability of Web-based Educational Systems*. CEUR, 2005, pp. 11–18.
- [18] C. Lagoze, H. Van de Sompel, M. Nelson, and S. Warner, "Open archives initiative - protocol for metadata harvesting - v.2.0," <http://www.openarchives.org/OAI/openarchivesprotocol.html>, Jun. 2002. [Online]. Available: <http://www.openarchives.org/OAI/openarchivesprotocol.html>
- [19] M. W. Berry and J. Kogan, "Text mining," *Applications and Theory. West Sussex, PO19 8SQ, UK: John Wiley & Sons*, 2010.