

Generación Dinámica de Redes de Colaboración en México y América Latina (GReCo)

**Proyecto apoyado por la Corporación Universitaria para el Desarrollo
de Internet (CUDI)**

Reporte Semestral

**Dr. Alfredo Sánchez
Universidad de las Américas Puebla**

**Dr. Juan Carlos Lavariega
ITESM Monterrey**

**Ing. Arianna Becerril
Universidad Autónoma del Estado de México**

Junio 2009

Generación Dinámica de Redes de Colaboración en México y América Latina (GReCo)

Reporte Semestral, enero - junio 2009

Proyecto financiado por la Corporación Universitaria para el Desarrollo de Internet CUDI

Investigadores responsables:

Dr. Alfredo Sánchez,
Universidad de las Américas- Puebla, UDLAP
COORDINADOR

Dr. Juan Carlos Lavariega,
Instituto Tecnológico y de Estudios Superiores de Monterrey

Ing. Arianna Becerril,
Universidad Autónoma del Estado de México

Resumen

En este documento se reportan las actividades realizadas en el primer semestre del proyecto “Generación Dinámica de Redes de Colaboración en México y América Latina” (GReCo). Los avances más sobresalientes incluyen una revisión exhaustiva de los proyectos relacionados, lo cual se reporta en documentos que servirán de referencia, así como el diseño, construcción y evaluación preliminar de prototipos para los componentes de software que se producirán como parte del proyecto. El reporte está organizado en los siguientes apartados: presentación del proyecto GReCo, reporte de avances y actividades, documentación generada, referencias y directorio de participantes.

Presentación

El presente proyecto se enfoca a generar redes que reflejan la afinidad temática y autoral entre documentos y permitir su visualización y búsqueda para detectar dinámicamente y oportunidades de colaboración entre autores e instituciones.

Tomando como base las colecciones disponibles en la Red Abierta de Bibliotecas Digitales (RABiD), entre las que se cuentan revistas especializadas y de divulgación, documentos digitalizados, así como tesis de licenciatura y posgrado, se aprovecharán mecanismos existentes basados en OAI-PMH y se desarrollarán otros según los protocolos de las colecciones participantes para recolección de sus metadatos.

Como parte del proyecto, se construirán índices y se aplicarán técnicas de inferencia para medir la proximidad semántica entre documentos y generar una representación interna de las correspondientes redes entre autores y temas. Utilizando esta representación, se ofrecerán varios servicios de acceso para explorar las redes, incluyendo búsquedas semánticas, interfaces de visualización de redes y grupos resultantes, así como respuestas a preguntas específicas acerca de oportunidades potenciales de colaboración.

El proyecto inicia con las colecciones que se ofrecen a través de RABiD, pero se planea aplicar a todas las instituciones integrantes de CUDI, así como a repositorios a nivel latinoamericano que permitan detectar textual y visualmente oportunidades de colaboración y apoyar la generación de propuestas y proyectos conjuntos.

Avances del proyecto

El proyecto tiene avances de acuerdo a lo programado en las tres instituciones participantes. En este resumen se mencionan los logros más importantes, dejando los detalles para los reportes técnicos de cada institución.

En la UDLAP se produjeron tres propuestas de interfaces para redes de colaboración, para cada una de las cuales se construyeron y se evaluaron prototipos de baja fidelidad. Dos de las interfaces se refieren a la visualización de redes de colaboración, para las que se han propuesto mecanismos innovadores específicos para el tipo y volúmenes de datos que se manejan en GReCo, mientras que la tercera se enfoca a ofrecer una alternativa de respuestas a preguntas específicas sobre redes de colaboración.

En primer lugar se propuso una interfaz basada en grafos pesados y filtros que facilita la exploración y visualización de redes de colaboración a partir de una institución o un autor. El uso de colores y nodos de diferentes tamaños resaltan las características de las redes y facilitan su interpretación por el usuario.

La segunda interfaz introduce el concepto de *onto-starfish*, el cual enriquece mediante el uso de ontologías la visualización de *Star-fish* propuesta en proyectos anteriores. *Star-fish* promueve el uso de campos de estrellas (*starfields*) a la vez

que permite explorar zonas específicas de un espacio de datos manteniendo el contexto mediante el uso de la técnica de “ojo de pescado”. Onto-starfish facilita la ubicación de entidades de interés en un campo de estrellas agregando una semántica familiar para el usuario al utilizar ontologías en cada eje de la visualización. Asimismo, onto-starfish permite el uso de múltiples ojos de pescado sobre un solo campo de estrellas.

La interfaz para el módulo de respuestas a preguntas se planteó de manera que aprovecha los hábitos desarrollados por los usuarios para utilizar motores de búsqueda mediante palabras clave. De esta manera, el módulo será un híbrido de respuestas a preguntas y recuperación de información.

Para alimentar las interfaces descritas, un proceso crucial es la inferencia de relaciones de colaboración a partir de las colecciones de documentos disponibles. En este sentido ITESM ha estado explorando y probando algoritmos de agrupamiento como como K-means, FIHCy K-NN, así como mecanismos para construcción de ontologías. En este sentido se está considerando el uso de OntOAIr, desarrollado en la UDLAP y que continúa siendo objeto de investigación en la Universidad Politécnica de Puebla. Un avance importante ha sido la definición de las estructuras de datos necesarias para representar las redes de colaboración resultantes del proceso de inferencia.

Finalmente, para hacer posible el proceso de inferencia es necesario garantizar el acceso uniforme a las colecciones digitales participantes. En UAEM, gran parte del esfuerzo se ha enfocado a afinar los procesos de cosecha de datos utilizando los servidores de metadatos basados en OAI-PMH. Se han analizado las coincidencias y discrepancias en los metadatos que reporta cada colección y se han generado esquemas para facilitar la unificación de atributos. Para asegurar que cada componente de la arquitectura de GReCo recibe los atributos que requiere para funcionar, se han hecho ajustes en el contenido de cada atributo que genera en particular Redalyc, la colección a cargo de UAEM. Otro avance de utilidad es el análisis de alternativas de sistemas para manejo de bases de datos nativas XML, el cual hace posible la selección del software más adecuado en este nivel de manejo de datos de la arquitectura.

Difusión

Los integrantes de GReCo participaron en la Reunión de Primavera de CUDI, celebrada en Cholula, Puebla, en la que se presentó el proyecto y sus avances al momento. Destaca en este rubro el curso de “Herramientas para gestión de colecciones de publicaciones electrónicas de acceso abierto”, impartido por UAEM también durante la reunión en Cholula. Asimismo, se han realizado presentaciones al interior de las instituciones.

Documentos elaborados

Reportes técnicos por institución

UDLAP, 2009. Reporte Técnico Semestral GReCo. Junio.

ITESM, 2009. Reporte Técnico Semestral GReCo. Junio.

UAEM, 2009. Reporte Técnico Semestral GReCo. Junio.

Otros documentos

UDLAP, 2009. Evaluación de modelos existentes de visualización. Mayo.

UDLAP, 2009. Evaluación de herramientas para implementación de interfaces de visualización. Junio.

Referencias

Corporación Universitaria para el Desarrollo de Internet A.C. (CUDI)

<http://www.cudi.edu.mx>

GReCo

Portal del proyecto: <http://ict.udlap.mx/greco>

Wiki del proyecto GReCo: <http://ict.udlap.mx/wikis/greco>

Colección de revistas en UAEM (Redalyc)

<http://redalyc.uaemex.mx/>

Colección de tesis en ITESM (Phronesis)

http://copernico.mty.itesm.mx/phronesis/mty/Default_en.html

Colección de tesis en UDLA (Tales)

<http://ciria.udlap.mx/tesis>

Laboratorio de Tecnologías Interactivas y Cooperativas (ICT)

<http://ict.udlap.mx>

Directorio de Participantes***UDLAP***

Dr. J. Alfredo Sánchez, investigador principal

Federico Hernández Bolaños, coordinador técnico del proyecto

Alfredo Ramos Lezama, becario

Efren Hernández, becario

Victor Madrigal, becario

Rosa Paredes, colaboradora

ITESM

Dr. Juan Carlos Lavariega, investigador principal

Eric Santiago Balam Sabido, becario

Martha Sordia Salinas, colaboradora

Ramón López Osuna, colaborador

UAEM

Ing. Arianna Becerril, investigador principal

Josué Isaac Legorreta Pérez, becario