

Developing an automatic crawling system for populating a digital repository of professional development resources: A pilot study

Jung-ran Park, y otros

Journal of electronic resources librarianship, ISSN 1941-1278, Vol. 28, n. 2, 2016

Se realiza un estudio para el desarrollo de un repositorio digital sostenible de recursos para profesionales, en el área de los estándares y tecnologías emergentes para la organización de datos y gestión en bibliotecas. El proyecto se engloba dentro del denominado "*Building a workforce of Next Generation Librarians for 21st Century Global Information Access*", llevado a cabo entre los años 2014-2017 por el Institute of Museum and Library Services (IMLS). Las bibliotecas, archivos y museos tienen un papel fundamental como portal para la diversidad cultural y de recursos, de ahí la importancia de realizar un estudio acerca de los recursos digitales. En el análisis previo de las publicaciones existentes sobre este tema, destaca la inexistencia de contenidos versados en la catalogación y la educación sobre metadatos, para que los estudiantes de biblioteconomía lo utilicen en su formación. Así como otros buscadores, como Google, se basan para sus búsquedas en grandes y costosas infraestructuras, este proyecto pretende construir un repositorio digital con un motor de búsqueda, capaz de mejorar la precisión y rapidez en la recuperación de la información a un coste razonable. Toda la infraestructura lógica se basa en una combinación de contenido similar y estructuras enlazadas. Se seleccionaban páginas web con contenido relevante en recursos de educación que eran enviados al motor de búsqueda, utilizando palabras clave. El objetivo era diseñar una dinámica de trabajo para rastrear, monitorizar y clasificar objetos web en categorías de desarrollo profesional. Con estas premisas, se realizó un prototipo para el mejorar el sistema de rastreo automático y el repositorio resultante mantendría por sí mismo los recursos en educación para la comunidad de catalogación y metadatos. El siguiente paso incluye el desarrollo de mecanismos para el tratamiento de recursos identificados en el repositorio, la generación de metadatos para cada recurso y el desarrollo de una interfaz de usuario que facilitará el empleo del repositorio, de manera que se produzca un mantenimiento propio en los recursos de educación para la comunidad de metadatos.

Resumen elaborado por la Sección de Documentación Bibliotecaria