

Représenter en RDF, interconnecter et visualiser en graphe des jeux de métadonnées archivistiques de provenances multiples: un projet de prototype

Anila Angjeli, Florence Clavaud y Stéphanie Roussel

La Gazette des archives, ISSN 0016-5522, n. 245, 2017, p. 157-172

Los Archivos Nacionales, la Biblioteca Nacional (BnF), el Servicio Interministerial de Archivos de Francia (SIAF) y el CRIHN de la Universidad de Montreal se han asociado para elaborar un prototipo que permite representar en RDF, interconectar y visualizar en gráficos conjuntos de metadatos archivísticos. El proyecto quiere demostrar que es posible: representar de manera rigurosa los datos archivísticos producidos por muchas instituciones; interconectar y enriquecer lo conjuntos de datos; y visualizar estos conjuntos de forma gráfica. Se trata de establecer a prueba de fallos la utilidad de esta hipótesis. El proyecto se inició en 20015 liderado por los Archivos Nacionales y se prevé que esté listo en 2017. Se creará una aplicación web accesible a todos, que integrará archivos RDF resultantes de la conversión de conjuntos de metadatos. Se prevé que el software tenga licencia abierta. Cada una de las instituciones promotoras tiene objetivos diferentes. El SIAF no produce metadatos, pero difunde normas y buenas prácticas, por lo que el prototipo será útil a su vocación pedagógica. Los Archivos Nacionales han realizado numerosos proyectos de búsqueda de información, pero sus metadatos no están "semantizados", lo que limita su uso. Este proyecto les permitirá avanzar en la aplicación de la tecnología para la representación en forma RDF y la interconexión y visualización de los metadatos generados y difundidos por el sistema de información archivística. La BnF conserva colecciones de archivos, pero no tiene una herramienta de descripción archivística detallada. Este proyecto le permitirá visualizar su trabajo, valorando la riqueza de la información que saca a la luz. También servirá para interconectar sus datos con los de los Archivos Nacionales. El proyecto también será útil para los historiadores, facilitando la investigación. El prototipo ambiciona superar dos desafíos tecnológicos: una representación RDF de calidad de metadatos archivísticos, y una visualización de los datos. Hasta el momento no había una ontología general de referencia para los archivos. Este vacío se verá colmado con la aparición en 2017 de la primera versión de la ontología RDF/OWL Records In Contexts-Ontology (RIC-0), que se ha utilizado como ontología de referencia para este proyecto. Respecto a la visualización de los datos, la combinación de gráficos y de una representación de la dimensión temporal es infrecuente, y ninguna aplicación es utilizada para el proyecto, por lo que habrá que desarrollar librerías ya

existentes. Los conjuntos de metadatos elegidos debían ser reales, y se seleccionó los proporcionados por la gestión de monumentos históricos y los de bibliotecas públicas. Un primera análisis muestra que os conjuntos de metadatos están en gran medida solapados y que son complementarios. En este momento del proceso, antes de la conversión a RDF, ya se han identificado varios problemas, como la divergencia de puntos de vista en la representación del tiempo o la importancia de conservar las huellas de procedencia de la información.

Resumen elaborado por Antonio Rodríguez Vela