

La publicación en Linked Data de registros bibliográficos: modelo e implementación

Jose A. Senso, Wenceslao Arroyo Machado

Revista española de documentación científica, ISSN-e 1988-4621, Vol. 41, n. 4, 2018, p. 1-20

Desde que la Biblioteca del Congreso de los Estados Unidos anunciase su intención de profundizar en entornos abiertos y compartidos en el trabajo bibliotecario hasta la actualidad, las bibliotecas están viendo cómo se están transformando la mayoría de herramientas, lenguajes y sistemas que emplean para realizar su trabajo. Esta adaptación lleva un tiempo, pero parece claro que los conceptos de datos abiertos, Linked Data y RDF son la base de los modelos, lenguajes y mecanismos para describir y catalogar en el futuro. Catalogar es un proceso complejo y arduo que, si se hace de manera concienzuda, requiere de mucho esfuerzo y profesionalidad por parte de las personas encargadas de llevarlo a cabo. El objetivo de lograr la interoperabilidad que permitir compartir datos, recursos y esfuerzos, pasa por dos principios: no abandonar la web y aprovecharse de las bases tecnológicas y filosóficas que de ella emanan. Los principios sobre los que se sustenta el sistema de publicación Linked Data se establecieron en 2010 sobre un mecanismo que aporta hasta un máximo de 5 estrellas en función a cómo se compartan los datos. Cada vez son más los proyectos realizados en este entorno de las bibliotecas con este método como principal protagonista. Son muchos los factores que han determinado el empleo de diferentes metodologías para lograr la conversión de los datos. A partir del estudio de todos los proyectos, se propone una metodología

para lograr ese objetivo basada en seis etapas que permite la implementación de esos datos alcanzando las cinco estrellas a las que debería tender cualquier proyecto. También se llega a la conclusión de que existe la posibilidad de automatizar los procesos de extracción de datos, la limpieza y la generación de grafos RDF. Esta metodología ha permitido poner en evidencia la necesidad de que tanto los programas de automatización de bibliotecas como la política bibliotecaria permitan incorporar más métodos para la extracción de datos bibliográficos. En lo que respecta a la modelación de los datos, es muy posible que muchos de los problemas encontrados tengan solución elaborando un vocabulario u ontología propia. Para finalizar, es importante destacar la cantidad de herramientas disponibles que permiten acometer prácticamente todas las etapas de esta metodología. Algunas, como la expuesta en este trabajo, no tendrían sentido si las bibliotecas contaran, a través de su propio programa de automatización, de mecanismos para publicar en Linked Data sus registros bibliográficos. La mayoría de bibliotecas están abocadas a retrasar su salto hacia este nuevo método. Por ese motivo esta metodología tiene razón de ser.

Resumen realizado por José María Amate Sánchez