

## Linked data para la gestión del conocimiento compartido en bibliotecas, archivos y museos (LAM)

### Síntesis operativa

Las bibliotecas, archivos y museos (LAM) poseen una gran cantidad de datos y recursos que a menudo han permanecido ocultos en los catálogos y archivos. Aprovechar la posibilidad de mejorar y difundir esta información a un público más amplio, es una oportunidad para enriquecer aún más la World Wide Web, fomentar una cultura de apertura al conocimiento y crear una serie de ventajas para todos los elementos de la cadena de la información. De esta manera se puede preservar el patrimonio cultural mundial en beneficio de las generaciones futuras y se puede mantener la investigación en sus idiomas originales, preservando la variedad y la vitalidad cultural.

Los linked data son las herramientas que van a permitir que todo esto sea posible.

El término linked data se refiere a un grupo de reglas para la publicación y conexión de los datos estructurados en el web y se basa en el concepto de relación cualificada. El World Wide web se funda en la creación de conexiones hipertextuales entre los documentos del web, que constituyen su base para la navegación.

La idea central de los linked data consiste en utilizar conexiones hipertextuales no sólo para identificar los documentos web, sino también entidades arbitrarias del mundo real con el fin de crear una verdadera red de datos unívocos y certificados. Se supera así el problema de la integración de catálogos de objetos diferentes, gracias a un sistema que los interconecta independientemente de su tipología (libros, obras, etc) y de su origen (bibliotecas, museos, archivos, etc), creando un tejido de informaciones extremadamente más funcional y completo para quien realiza búsquedas y al mismo tiempo convirtiendo el conocimiento compartido en un instrumento colaborativo para las instituciones culturales.

Los linked data ponen las bases para ulteriores mejoras en las dinámicas de colaboración entre instituciones y para la reutilización de los datos en contextos diferentes, consintiendo la identificación y posibilidad de discovery más eficientes basados en el web.

Este nuevo enfoque es para las bibliotecas, archivos y museos una buena ocasión para crear lazos de unión entre catálogos y ofrecer a los usuarios instrumentos de búsqueda más sofisticados y siempre en línea con la evolución de las nuevas tecnologías web mientras que para los productores de datos se configura como un paso decisivo para la creación y la promoción de nuevas formas de cooperación.

Los potenciales beneficios para todo el sector de las humanidades, que hoy en día y más que nunca está a riesgo de marginalización, son significativos.

Flexible y sostenible, este enfoque cooperativo permite la creación de un sistema bien estructurado para la organización de datos, teniendo en cuenta las diversas tradiciones culturales y explotando las oportunidades que ofrecen las nuevas tecnologías. Las figuras clave del sector podrán contribuir a lado de organizaciones más pequeñas, listas para compartir su experiencia y la singularidad de sus recursos.

Respetando la nueva filosofía abierta de compartir y reutilizar datos, el concepto de registros de autoridad tradicionales también debe evolucionar, y desplazar el núcleo de la discusión de la determinación de si el control debe tener lugar a nivel central o local, hacia una noción más flexible centrada en la identificación de la entidad y las relaciones entre las entidades. Bibliotecas, archivos, museos, editoriales y proveedores se involucrarán en la generación de nuevos datos y en el descubrimiento de nuevos recursos, superando los límites de sus dominios específicos para crear oportunidades de enriquecimiento de datos que anteriormente hubieran sido impensables.

En un contexto de renovación de los lenguajes y estándares de la catalogación, que ha visto en los últimos años la aparición de los lineamientos RDA (Recursos, Descripción y Acceso) y la publicación del Modelo de Referencia de Bibliotecas (LRM) de IFLA, nace el modelo de datos BIBFRAME, que actualmente es tema de discusión y desarrollo dentro de la comunidad bibliotecaria y otras. La intención a la base del framework es abrir las posibilidades que ofrecen los linked data a bibliotecas, archivos y museos, proporcionando una mayor interoperabilidad, visibilidad y disponibilidad para todo tipo de recursos.

El objetivo de los componentes y herramientas desarrollados dentro de la familia de iniciativas Share, basadas en linked data, que incluye Share-VDE (Virtual Discovery Environment), Share-ART (con las bibliotecas de historia del arte del Instituto Max Planck), y Share-MUSIC (una experimentación en el sector musical) es el de crear un entorno útil para la gestión del conocimiento. Las interfaces de búsqueda avanzada mejorarán la experiencia del usuario y proporcionarán resultados más amplios a bibliotecas, archivos y a sus usuarios. Los bibliotecarios, archiveros y conservadores de museos se beneficiarán de avances tales como el desarrollo continuo de las funciones de catalogación de acuerdo con los estándares de la web semántica, la integración de los procesos de un entorno colaborativo con los sistemas y herramientas locales. El objetivo es ayudar a revelar la riqueza de datos en las colecciones existentes, que a menudo permanecen ocultos o no expresados en un catálogo tradicional.

En el desarrollo de la iniciativa, los registros de bibliotecas, archivos y museos se transforman en linked data, a través de la identificación de la entidad, la reconciliación y el enriquecimiento de los datos. Los atributos se utilizan para identificar de forma unívoca a una persona, una obra u otras entidades, con formas variantes conciliadas para formar un clúster de datos relacionados con la misma entidad. Los datos se concilian posteriormente con los registros de autoridad tradicionales y otras fuentes externas para enriquecerlos y crear una red de información y recursos. El resultado es una base de datos de relaciones, abierta a toda la comunidad, y una base de conocimiento de clústeres en el RDF común. La

base de datos utiliza los paradigmas de la web semántica, pero también permite a las bibliotecas participantes continuar gestionando sus datos lo más independientemente posible.

Entre las principales áreas de actividad:

- enriquecimiento de registros MARC con URI;
- conversión de datos a MARC en RDF utilizando el vocabulario BIBFRAME y otras ontologías añadidas, si es necesario;
- creación de una plataforma de discovery basada en una adaptación del modelo BIBFRAME, desarrollada para ofrecer un servicio de discovery en linked data;
- creación de una base de datos de relaciones y cluster de recursos (Sapientia Cluster Knowledge Base) accesible en RDF;
- implementación de componentes para la interacción directa con los datos, que consienten validación, actualización, control a largo plazo y mantenimiento de los cluster y de los URI que identifican la entidad;
- procedimientos de actualización automática o por lotes;
- distribución de lotes / automatizada de datos a bibliotecas;
- implementación progresiva de ulteriores casos de uso según el orden establecido por la comunidad. Entre estos: catalogación por copiado, catalogación original, API (Application Programming Interface) para ILS, retro conversión para sistemas de adquisición y administración locales basados en formatos MARC, informes.

Se fomenta la colaboración de las instituciones participantes para que las decisiones sobre los próximos pasos en la implementación a gran escala en la comunidad de bibliotecas, archivos y museos se basen en fundamentos sólidos derivadas de los conjuntos de datos utilizados. Los datos utilizados no se convierten simplemente en RDF, sino que también se enriquecen con identificadores y se interconectan, para permitir su uso en el entorno de linked data.

Estos procesos pueden contribuir a la construcción de un modelo realista de la actividad del usuario, abordando potenciales problemas y ventajas de pasar al entorno de linked data. Share-VDE se desarrolla según un principio de funcionalidad, proporcionando diversos entornos e interfaces de usuario para la creación y enriquecimiento de datos, ofreciendo flujos de trabajo que respondan a las diferentes necesidades de bibliotecarios, profesionales, académicos, investigadores y estudiantes participantes.

---

Casalini Libri ([www.casalini.it](http://www.casalini.it)) es una agencia bibliográfica que produce registros bibliográficos y de autoridades; proporciona libros y revistas a las bibliotecas; ofrece una variedad de servicios de desarrollo de colecciones; trabaja en colaboración con editoriales y bibliotecas para proporcionar contenido electrónico.

@Cult ([www.atcult.it](http://www.atcult.it)) es una compañía de desarrollo de software especializada en la conversión de datos para LD; proporciona sistemas de biblioteca integrados (ILS) y herramientas de discovery que brindan soluciones tecnológicas eficaces e innovadoras para mejorar la gestión de la información y el intercambio de conocimientos.