

Colecciones digitales

Digital Collections

Resumen

En este trabajo se analiza el concepto de colecciones digitales (CD) desde las aproximaciones de bibliotecas digitales, bibliotecas virtuales y archivos digitales, hasta las agrupaciones de objetos digitales más diversas. Se toman en cuenta los enfoques de humanidades digitales así como la concepción de las colecciones digitales como datos y las colecciones de datos.

Del análisis surge que bajo el término hay enfoques muy diversos. Las aproximaciones provienen de las disciplinas de origen, el contexto en que se formulan o las instituciones que las sustentan entre otros muchos elementos que las atraviesan.

Las CD se consideran entonces enmarcadas en el concepto de objeto fronterizo, una entidad flexible que se adapta a las necesidades y limitaciones de los diversos grupos que lo emplean aunque lo suficientemente robusta para mantener una identidad común en diferentes ámbitos. Las CD se usan en el ámbito de bibliotecas, archivos y museos y se asocian a servicios de valor agregado muy heterogéneos.

Se analizan las fases del ciclo de vida de las CD entre los distintos enfoques concluyendo que existen muchas correlaciones entre las etapas y los elementos que se plantean entre distintas ópticas.

Por último se desarrollan las buenas prácticas y los aspectos que se consideran más significativos en el diseño de las CD : estándares, interoperabilidad, licenciamiento, accesibilidad, datos semánticos, gestión, preservación, servicios, interacción y aplicaciones de software libre.

Palabras clave: colecciones digitales, objetos fronterizos, ciclo de vida, aplicaciones software libre

Abstract

This paper analyzes the concept of digital collections (DC) from the point of view of digital libraries, virtual libraries and digital archives, including the most diverse groupings of digital objects. Also are considered approaches to digital humanities and the conception of digital collections as data and Data collections.

From the analysis is derived that under the term there are very different approaches. These ones come from the original disciplines, the context in which they are formulated or the institutions that support them, among many other elements that participate transversely.

The DC are then considered subsumed in the concept of a border object, that is, a flexible entity that adapts to the needs and limitations of the various groups that use it, although it is robust enough to maintain a common identity in different areas. DC are used in libraries, archives and museums and are associated with very heterogeneous value-added services. The phases of the life cycle of the DC between the different approaches are analyzed, concluding that there are many correlations between the stages and the elements present using different optics.

Finally, good practices that include aspects that are considered most significant in the design of the DC are developed: standards, interoperability, licensing, accessibility, semantic data, management, preservation, services, interaction and free software applications.

Keywords: digital collections, border objects, life cycle, free software applications

1. Introducción

Vivimos en un mundo digital por lo que existe una gran proliferación de documentos digitales de todo tipo. Estos documentos pueden ser objetos que son concebidos digitales, representaciones digitales de objetos físicos o remezclas de otros objetos digitales. En las instituciones culturales estos objetos se agrupan recibiendo tratamientos técnicos muy diversos.

Estos tratamientos pueden estar influenciados por la visión profesional o institucional y no por un marco conceptual o práctico pertinente que pueda establecer una perspectiva más amplia o vinculada.

Se busca establecer que una institución cultural puede adoptar más de un enfoque y que éstos pueden vincularse y abrir un panorama más abarcativo y rico. Un archivo de una colección de un poeta, tratado en forma archivística, puede incluir un vínculo a su biblioteca tratado en forma bibliotecológica, y a una colección digital de objetos de su pertenencia tratados en forma museística. Los museos pueden realizar exposiciones virtuales de las representaciones de sus objetos físicos e incluir otras informaciones contextualizadoras del hecho cultural que brinden una perspectiva más profunda utilizando estándares técnicos de disciplinas afines. En suma, adoptar una mirada facetada sobre las colecciones digitales abre una perspectiva interrelacionada de mundos distintos que se vinculan.

El propósito de este trabajo es problematizar el concepto de colecciones digitales y aportar una posible concepción teórica y lineamientos prácticos para su tratamiento técnico.

La metodología parte de una revisión bibliográfica que busca establecer un marco conceptual. Este marco conceptual se desarrolla en la sección 2, recorriendo conceptos cercanos y esquemas de tratamiento y se concluye que resultan isomorfos.

En la sección 3 se realiza una recopilación de elementos que se pueden considerar buenas prácticas en el tratamiento técnico, que agrupan aspectos como la adhesión a estándares de metadatos, licenciamiento, accesibilidad, semántica, preservación digital y servicios de valor agregado, entre otros.

Finalmente en la sección 4 se realiza un relevamiento sintético de aplicaciones de software libre que corporizen el marco conceptual y las buenas prácticas pretendiendo aportar así un panorama completo desde la planificación hasta la implantación de las colecciones digitales.

2. Marco Conceptual

Existe cierta dispersión nominativa en la definición de lo que en este ámbito denominaremos colecciones digitales (CD) pero que puede designarse como bibliotecas digitales, bibliotecas virtuales, bibliotecas electrónicas, archivos digitales o inclusive repositorios.

Según Ferrer et. al. (2002) las bibliotecas digitales son una manifestación de la adaptación de los servicios tradicionales frente a la aparición de la tecnología.

Witten , Bainbridge y Nichols (2003) plantean la idea de una biblioteca digital como una colección de objetos digitales diversos (textos, imágenes, videos y audio) pero con una organización que abarca la selección, creación, mantenimiento y acceso.

Sorli y Merlo (2000) definen la biblioteca digital como:

“Una colección de documentos electrónicos convenientemente organizada y dispuesta para su uso. Colección, organización y difusión son los tres aspectos básicos que distinguen a las bibliotecas. En el caso de las bibliotecas digitales el marco es el mismo, ya que se trata de formar colecciones de documentos electrónicos, organizarlas con criterios biblioteconómicos y ponerlas a disposición de los usuarios a quienes pueda interesar”.

Según Méndez (2001) una biblioteca digital debe cumplir los elementos que designa:

biblioteca: debe implicar, selección, organización y colección.

digital: debe implicar que los ítems de la colección estén en formato de codificación numérica, y que la organización y acceso a todos los servicios sea también digital, a través de la Web.

Si una biblioteca es un conjunto organizado de documentos, una biblioteca digital debe ser un conjunto organizado de documentos electrónicos.

Tramullas (2002) sostiene que el concepto de biblioteca digital, es algo que está en continuo cambio y que va más allá de una colección digitalizada de una unidad de información a disposición de una comunidad de usuarios, y que es un proceso que debe tener en cuenta todo el ciclo de vida de los documentos de la colección, su procesamiento técnico y

los servicios que proporciona, haciendo énfasis de la interactividad en los mismos y la posibilidad de ofrecer valor añadido para los usuarios.

Por otra parte las definiciones de biblioteca digital (Calhoun, 2014), resultan a veces similares a los del término repositorio. En este trabajo no consideraremos el término repositorio, que aunque cercano, se correspondería desde el punto de vista de la interactividad y la arquitectura de la información a una situación diferente. En un repositorio, una institución permite la ingesta de recursos digitales en esa plataforma con una línea de montaje de aporte de metadatos y asignación temática que culmina con el ingreso del recurso digital al repositorio, pasando luego por un proceso de revisión que se culmina con la disposición del recurso en colecciones para el usuario final.

Parecería entonces que el concepto de biblioteca digital se desarrolla en la medida que las bibliotecas han incorporado a sus colecciones recursos digitales o digitalizados que se agrupan o bien en forma de colecciones separadas o junto al resto de los recursos no digitales quedando disponibles en forma virtual con una interfaz de usuario (OPAC). Paulatinamente los recursos digitales empiezan a predominar entonces se posibilita además del registro bibliográfico la visualización del recurso digital o su descarga. Estos recursos digitales pueden estar en muy diversos formatos: imagen, video, texto, música, entre otros, y ser de tipologías documentales también muy diversas: informes, artículos, películas, etc.

La idea de colección como un agrupamiento con especialización o profundidad en cuanto al contenido referente a una temática, empieza también a deslindarse de la biblioteca y lo que podría verse en algunos sistemas integrados como estanterías o exposiciones aparece entonces como un producto separado que puede estar vinculado al sistema integrado de gestión bibliotecaria o no.

Los archivos, museos, galerías y todo tipo de instituciones culturales empiezan a utilizar estas colecciones digitales como exposiciones virtuales. Un archivo literario realiza una exposición de los poemas digitalizados, un museo propone una recorrida por las representaciones de las obras y las biografías de los pintores a través de colecciones digitales en la web. Incluso los proyectos que amplían las concepciones más tradicionales las utilizan. Una colección digital en un proyecto de rescate de la memoria territorial de una zona puede exponer videos de las personas relevantes del barrio contando la historia de los sitios emblemáticos, incluyendo fotografías de esos sitios, documentos de creación o comentarios de la prensa alusivos a los lugares, entre otros objetos digitales que perfilan el propósito del proyecto. En suma las colecciones digitales se aplican en todo tipo de emprendimientos que persigan un objetivo y rescaten documentos con ese fin específico.

Un concepto cercano a las CD son los sistemas de gestión de colecciones – Collections Management Systems – que organizan, controlan y gestionan los objetos que forman parte de una colección sea esta museística o archivística. En estos casos los sistemas se interesan en registrar información de gestión como la localización del objeto museístico, su procedencia, información de curaduría, información de preservación o conservación, historia de exposiciones en las que ha participado el objeto, entre otros datos.

Los archivos y museos tienen sistemas de gestión y acceso diferentes a los usados en las bibliotecas que corresponden a otras necesidades y situaciones pero confluyen en la creación de colecciones digitales con fines como la difusión o el apoyo a actividades de investigación.

En este sentido las colecciones digitales constituyen un objeto limítrofe, un objeto que está en la frontera entre distintas concepciones institucionales, profesionales y técnicas.

Según Star y Griesemer (1989) un objeto limítrofe es una entidad flexible que se adapta a las necesidades y limitaciones de los diversos grupos que lo emplean aunque es lo suficientemente robusta para mantener una identidad común en diferentes ámbitos. Estos objetos pueden ser abstractos o concretos. La gestión de objetos limítrofes es clave para el desarrollo y mantenimiento a través de mundos que se intersectan y administrar las tensiones entre visiones heterogéneas y divergentes, así como también la cooperación en la comprensión, lo cual conduce a hallazgos generales que permiten establecer reinterpretaciones entre esas visiones. Los autores analizan los objetos limítrofes en el caso de un museo de historia natural y la elaboración de métodos y protocolos para su tratamiento.

Bowker y Star (1999) plantean como las clasificaciones, las categorizaciones y en general esa segmentación de la representación del mundo y los estándares utilizados para esos mundos, valorizan algún punto de vista y silencian otros. Estas categorizaciones por tanto tienen implicancias políticas y sociales, por lo cual tener claro estos aspectos en la práctica plantea una perspectiva más abarcadora. En particular la segmentación en los estándares en las colecciones

digitales entendidos éstos como “un conjunto de reglas acordadas para la producción de objetos” se extiende a más de una comunidad de práctica que aplican métricas heterogéneas.

Anderson (2007) opina que los objetos fronterizos son un concepto particularmente útil para reconocer artefactos informativos (libros, documentos, registros, citas u otras representaciones), porque son formas materiales y sociales. Las nociones de adhesión y alineamiento posicionan estos aparatos representacionales como articuladores de prácticas y tecnologías que son asimiladas a prácticas materiales y discursivas. Los contextos interdisciplinarios favorecen la interpretación de los conceptos en ciencia de la información.

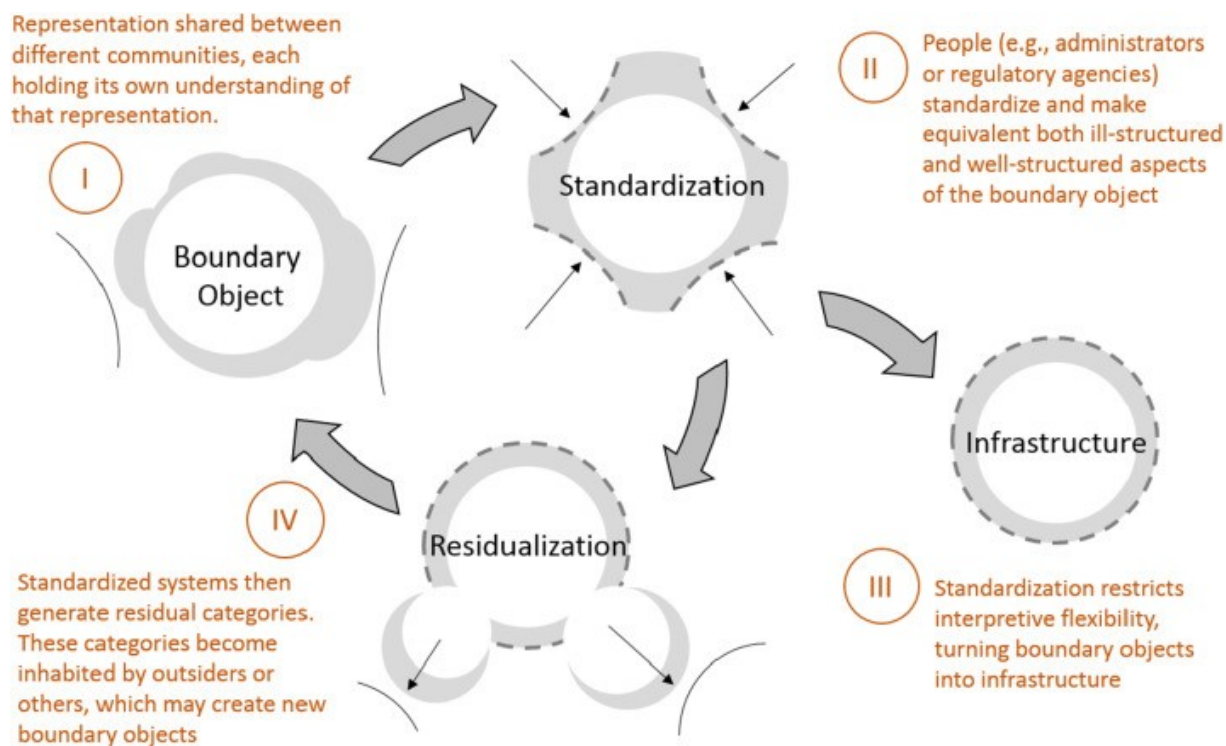


Fig. 1. Ciclo de vida del objeto fronterizo. Extraído de (Steger, 2018)

Según Steger et. al. (2018) los objetos fronterizos son entidades cambiantes que siguen un proceso dinámico o ciclo de vida a lo largo del tiempo como se muestra en la figura 1. Inicialmente lo difuso del objeto puede complicar su operacionalización. Esto encamina a una *estandarización*, o sea, un conjunto de reglas y definiciones para que la información pueda ser transferida a otros mundos y facilitar la colaboración.

El objeto fronterizo por su caracterización, pasa de tener un significado general a uno específico, en un proceso iterativo que la *estandarización* restringe y convierte en una *infraestructura*. En un sistema de información en esta fase se encapsulan los estándares y clasificaciones, se evalúan alternativas y se define el producto final.

En la etapa de *infraestructura* se naturalizan los conflictos porque se regula y formaliza la comunicación, se ignora la realidad diversa fijando una propuesta. En esta fase también se construye un espacio de colaboración entre distintos mundos. De ella pueden emerger opciones no consideradas, que se denominan objetos residuales, que pueden tornarse a su vez en objetos fronterizos que avancen hacia un nuevo ciclo.

Dentro de lo que se denominan objetos residuales se detallan los servicios culturales (Chan et al., 2012). Los servicios culturales son aquellos en los que la gente obtiene beneficios que no son materiales, que tienen que ver con la inspiración, la identidad cultural, la recreación, y agrego, la protesta, o la denuncia. Estos servicios culturales están íntimamente ligados a sus entornos.

Un caso de servicio cultural que constituye una colección digital es el del proyecto “Voces en el muro” de la Universidad Tecnológica Metropolitana que implementó un archivo visual con el registro de fotografías digitales de lo que se escribió en los muros de la ciudad de Santiago, en Chile, dando testimonio de la protesta del pueblo chileno en medio de una crisis social y política, en 2019.

Esta colección digital se alimentó con las fotografías que las personas tomaban de las consignas que se escribían en los muros y puede verse en:

<https://www.voceselmuro.cl/>

Podemos asimilar el ciclo de vida del objeto fronterizo con lo que plantea Tramullas (2007) al analizar que una colección debe enfocarse en el usuario y sus necesidades de información.

Las fases que plantea este autor para el ciclo de vida de la colección son:

1. Estudio de las necesidades de información
2. Definición de objetivos de la colección y la comunidad potencial de usuarios
3. Análisis de costos y beneficios
4. Selección de componentes de la colección
 1. selección de contenidos (documentos, metadatos, esquemas de clasificación y taxonomías)
 2. definición de flujos de trabajo
 3. criterios de calidad
5. Implementación de la colección
 1. Soluciones técnicas
6. Publicación, mantenimiento, control, monitorización de uso
7. Evaluación y mejora
8. Preservación digital

Haciendo un parangón con lo planteado para los objetos fronterizos, la fase de estandarización comprendería el análisis de los componentes de la colección digital, los metadatos, las taxonomías, los derechos de los objetos, la forma de ingesta, los aspectos de interoperabilidad, entre otros. La fase de infraestructura comprendería el análisis y desarrollo de las soluciones tecnológicas y los sistemas de información a implementar. La fase de objetos residuales serían todas las actividades a posteriori de la publicación de la colección, todos los aspectos de monitoreo, evaluación, mejora y actividades de preservación digital que pueden dar origen a replanteos de la colección en un ciclo iterativo o, a nuevos emprendimientos informacionales derivados.

Las colecciones digitales pueden ser un apoyo a los servicios tradicionales de bibliotecas, archivos y museos que realizan exhibiciones virtuales seleccionando representaciones digitales de sus colecciones especiales: libros raros, cartas históricas, manuscritos, esculturas, etc.

Estas exhibiciones son la puerta de acceso a los recursos que ofrece la Institución, recursos que a veces no son fácilmente accesibles físicamente, por una cuestión de deterioro o por otras razones. Las exhibiciones de colecciones digitales permiten crear narrativas para acceder de una forma diferente a los objetos de una colección física relacionada o pueden constituir colecciones digitales que han nacido digitales y pueden gestionarse o mostrarse de esta forma como una estantería virtual.

Las humanidades digitales (HD) constituyen un espacio que comparte la tecnología y la humanística y es un área muy vasta que abarca entre muchos asuntos el diseño y mantenimiento de colecciones digitales.

En una investigación sobre colecciones digitales en Aracne, una Red de humanidades digitales y letras hispánicas, (Alvite y Pena, 2019) se plantea una metodología con los siguientes bloques de análisis de colecciones digitales:

1. Colección: descripción del proyecto, contenidos y objetivos.
2. Arquitectura de la información: accesibilidad, interoperabilidad, tecnologías, estándares, modelos de datos y herramientas de organización del conocimiento.
3. Interfaz de usuario: prestaciones de búsqueda y navegación, presentación y visualización de resultados, objetos digitales y sistemas de ayuda

4. Valores añadidos: interrelaciones, contextualización, enriquecimiento semántico, integración de herramientas tecnológicas derivadas, sinergias con colecciones digitales patrimoniales.

Puede verse aquí también cierta analogía con las otras propuestas de fases en colecciones digitales y también caracterizaciones a tener en cuenta.

Un aspecto a destacar en esta investigación es que aparece la consideración de abordar la implementación de tecnologías semánticas (publicación en RDF y datos abiertos enlazados), la incorporación de herramientas de enriquecimiento semántico o la integración de aplicaciones para la explotación de datos (geolocalizadores, analizadores sintácticos, semánticos, etc.).

Los datos semánticos permiten interconectar información de distintas fuentes, construyendo una red de conocimiento a partir de conceptos y propiedades. Para lograr este cometido es fundamental basarse en estándares de la web semántica y relacionar las distintas fuentes utilizando estructuras conceptuales que luego puedan integrarse.

Wittman et. al. (2019) se plantea otro enfoque de las colecciones digitales, que es la de las propias colecciones como datos. Este enfoque busca extender el valor de las colecciones digitales desde un uso tradicional para ofrecer a los usuarios un acceso flexible que facilite el acceso a los datos de las colecciones de modo de facilitar la investigación en humanidades digitales. Las colecciones son entonces datos que se pueden manipular, analizar y visualizar con métodos computacionales, por ejemplo usando minería de datos, análisis de sentimientos, grafos de redes, visualización de datos, uso de información geográfica, entre otras estrategias.

Este enfoque es particularmente importante en las colecciones digitales que se crean a partir de recursos digitalizados. En la fase que plantea Tramullas (2007), estudio de las necesidades de información y definición de objetivos de la colección en la comunidad potencial de usuarios, surge la caracterización del proceso de digitalización a desarrollar. Las colecciones de textos que no sean sometidos a procesos de reconocimiento de caracteres de modo que sean realmente un texto y no una imagen, no servirán luego para realizar minería de datos, análisis lexicográfico u otras estrategias de este tipo.

Los aspectos de planificación de bibliotecas digitales, el establecimiento claro de los objetivos de la colección y los usuarios potenciales son clave para llevar a cabo estos emprendimientos de digitalización, ya que el objetivo no será solamente dar acceso a las colecciones digitales, sino que éstas sean útiles, recuperables y susceptibles de analizarse con herramientas digitales con el concepto de colecciones como datos.

Por otra parte en la fase de definir estándares, - v.g. metadatos, taxonomías, etc. - se estará suministrando información ampliada y contextualizada de los objetos y con ello se estarán habilitando operaciones posibles como el descubrimiento, la navegabilidad, la recuperación y la interoperabilidad, que también contribuyen al uso de las colecciones digitales como datos de investigación.

El profesional de la información debe estar siempre atento y participar de las iniciativas de creación de colecciones digitales aportando esta mirada estratégica y profesional que potencie la utilización de estas colecciones.

Por último queremos señalar que aparece una nueva área de objetos para las colecciones digitales, que son los datos propiamente dichos y de allí las colecciones digitales de datos que son un tipo de colección especializada y merece un estudio aparte.

De lo expuesto, surgen algunos elementos a considerar para la creación y gestión de las colecciones digitales, siempre en el marco de una planificación que tenga en cuenta las fases detalladas.

En cualquier sistema de información debe considerarse la arquitectura de la información, como se estructura, como se accede, como es el mapa que lo define, ya que será esta arquitectura la que nos hará decidir la solución informacional a implementar. El otro aspecto a considerar para las decisiones técnicas a tomar es como se produce la interacción en ese sistema de información, como opera esa CD en el contexto, en la comunidad o como puede proyectarse.

En lo que sigue nos enfocaremos en algunos de los elementos que se consideran buenas prácticas o elementos a considerar en la creación de las colecciones digitales.

2. Buenas prácticas

2.1 Estándares

Vamos a analizar dentro de este punto los estándares de todo tipo involucrados en las CD: administrativos (fecha de creación, lugar, instrucciones a los usuarios, identificadores internos, etc.), descriptivos (información del contenido, como título, resumen, etc.), estructurales (como se organizan las distintas partes del contenido), temáticos (asignación de un vocabulario, ontología, tesauro o similar), derechos (como es el licenciamiento, información de propiedades del derecho), interoperabilidad (adherencia a OAI-PMH), preservación (PREMIS), o de nicho como IPTC ¹entre otros.

Las colecciones digitales pueden provenir de instituciones diversas y en esos casos tendrán una cierta tendencia a heredar sus estándares o metadatos. La cuestión no es tanto que metadatos usar, sino que estos metadatos sean un estándar y que sea posible mapear a otros estándares y construir interoperabilidad.

Los estándares de archivo más utilizados en la región son: ISAD(G), ISAAR(CPF) , ISDIAH, ISDF y EAD, aunque también es posible que se esté experimentando con Records in Context (RIC). Los museos utilizan CIDOC-CRM, las bibliotecas alguna variante de MARC o RDA, y también los nichos tendrán sus propios estándares como IPTC para las fotos a modo de ejemplo.

El estándar más usado en colecciones digitales que es muy escueto pero que mapea a los antes detallados, es Dublin Core. Dublin Core constituye una lingua franca para los objetos digitales, es muy sucinto pero detalla los elementos más sustantivos y tiene correspondencia con los estándares mencionados.

Tiene también una característica muy importante ya que Dublin Core es intercambiado a través del protocolo OAI-PMH². Este protocolo es una herramienta de interoperabilidad porque permite compartir y agrupar objetos digitales a través del intercambio de metadatos. El protocolo proporciona dos roles, el de proveedor de servicios y el de proveedor de datos. El proveedor de servicios es un rol que recolecta metadatos de otros lugares con lo cual habilita una búsqueda unificada. Cuando un usuario encuentra algo que busca, el proveedor de servicios a través de los metadatos lo redirecciona al origen de los datos. El proveedor de datos proporciona sus datos o más bien, sus metadatos a un proveedor de servicios. Los dos roles se pueden implementar en un mismo sitio.

El otro aspecto importante de Dublin Core es que es extensible, es decir, tiene un núcleo con los metadatos principales, pero cada sitio puede adicionarle los metadatos específicos para su tipo de recurso digital que no estén contemplados. Dublin Core, con posibles extensiones, puede utilizarse para describir un recurso digital en forma sintética, salvo los casos especiales de Recursos educativos abiertos (REA) que tienen esquemas de metadatos más apropiados como Learning Object Metadata (LOM), o los mapas digitales que utilizan el estándar ISO 19115 o perfiles de este esquema.

La importancia de dar acceso a un recurso digital está fuertemente ligada a las condiciones de acceso y las posibilidades de reutilización y explotación. Los aspectos de recuperación de información, la manera por la cual este recurso digital puede ser descubierto y buscado, son primordiales. La descripción del recurso, la calidad de la misma, la exhaustividad, la pertinencia de los elementos descriptivos en un contexto, es fundamental para lograr que el recurso pueda ser buscado y accedido. Si los metadatos pueden adecuarse a la naturaleza de los recursos y al contexto, esto garantizará su descubrimiento y recuperación.

Sostiene Rieger (2019) que el contexto ha cambiado, de una escasez de información con canales de distribución limitados, a un panorama de abundancia informativa con múltiples canales. Esta situación conduce a que lo que se publique a través de una colección digital debe poder competir con todo tipo de oferta informativa. Es necesario que las CD ofrezcan facilidades de recuperación y descubrimiento y para lograrlo, una descripción de calidad es un imperativo insoslayable.

Como metadatos de preservación podemos mencionar a PREMIS ³ que es un esquema destinado a modelar la disponibilidad y la autenticidad en el contexto de preservación digital, incluyendo tanto los aspectos de derechos y permisos como los aspectos estructurales.

Un estándar muy importante en colecciones digitales de texto es TEI ⁴ (Text Encoding Initiative) , específicamente diseñado para el trabajo con textos de humanidades.

1 <https://iptc.org/std/photometadata/documentation/userguide/>

2 <http://travesia.mcu.es/portaln/jspui/html/10421/1823/page1.htm>

3 http://www.bne.es/es/Micrositios/Publicaciones/PREMIS/001_Introduccion/001_Antecedentes/002_Metadatosfundamentales/

4 <http://www.tei-c.org/index.xml>

Otro aspecto a considerar es la normalización temática: el uso de taxonomías, tesauros, ontologías o esquemas flexibles que refieran a la organización del conocimiento. En cada caso habrá que analizar a la luz de los objetivos de la colección digital, como desarrollar el tratamiento temático, a que estándares adherir, la creación de terminología ad hoc o la extensión de vocabularios. Se recomienda que estos vocabularios adhieran a Simple Knowledge Organization System (SKOS) que constituye una iniciativa de la W3C que proporciona un modelo para los esquemas conceptuales como listas de encabezamientos de materia, taxonomías, esquemas de clasificación, tesauros, ontologías y en general cualquier tipo de vocabulario controlado. La adherencia a SKOS permite la edificar la articulación temática.

No se mencionan específicamente los estándares informáticos pero es de importancia adherir a estándares abiertos en este campo de modo de garantizar la preservación y la interoperabilidad, así como el acceso.

2.2 Uso, reproducción de los materiales. accesibilidad

El licenciamiento de los objetos de la colección debe establecerse claramente: pueden ser recursos bajo dominio público, bajo licencias abiertas o protegidos con licencias. Deberá determinarse si los materiales pueden accederse, pueden descargarse, pueden reutilizarse y bajo que condiciones esto es posible.

Un recurso digital es abierto si se puede utilizar, reutilizar y redistribuir solamente con la mención de la autoría intelectual y manteniendo los mismos términos de la licencia original cuando se modifica el recurso. Esta sería la licencia de Creative Commons *compartir bajo la misma licencia*.

La práctica de crear colecciones digitales de recursos abiertos permite a las instituciones dar a conocer sus colecciones y facilita la cooperación y la reutilización de la información. El movimiento Open Glam⁵ recomienda la apertura de los objetos digitales.

El licenciamiento propietario puede ser un escollo para las estrategias de preservación digital sino se permiten las copias de respaldo o los cambios de formato para los distintos paquetes en el modelo OAIS.

Los objetos digitales (texto, gráfico, sonoro o multimedia) de una colección deberán estar accesibles para todos y hay que considerar que entre los usuarios puede haber personas que tengan alguna limitación física, sensorial o cognitiva, y encuentren problemas para poder acceder a toda la información. La accesibilidad debe garantizarse poniendo a disposición contenido en formatos legibles para distintos perfiles de usuarios con discapacidad: visual, auditiva u otra.

Las pautas de accesibilidad para el contenido web (WCAG)⁶ constituyen un estándar que satisface las necesidades de las personas con discapacidad y es compartido por organizaciones y gobiernos a nivel internacional. Las pautas tienen cuatro principios a cumplir: perceptible, operable, comprensible y robusto. De acuerdo a los criterios de conformidad se certifican los sitios con su nivel de cumplimiento como A, AA y AAA.

El texto digitalizado en formato imagen, no es legible para un dispositivo asistivo como un lector de pantalla que utilizará una persona ciega o con baja visión, de modo que ese texto no será accesible. Estos aspectos deben tenerse en cuenta al diseñar una colección digital y en los procesos de digitalización para crear la misma. Las imágenes en un texto tampoco son legibles y deben tener alguna adecuación, como un texto alternativo. Un video no es accesible para una persona sorda si no se le realiza la adecuación con un subtítulo o la incorporación de lenguaje de señas. El proyecto ESVIAL realizó una guía para elaborar contenido digital accesible:

http://www.esvial.org/wp-content/files/ESVIAL.LibroDigital_es_2015.pdf que indica como generar material digital accesible.

Respecto al licenciamiento propietario con respecto a la accesibilidad el tratado de Marrakesh⁷ brega por la anuencia para el cambio de formatos a los efectos de garantizar la misma a los distintos perfiles.

2.3 Datos semánticos

Las bibliotecas, archivos y museos junto a instituciones culturales estarán interesadas no solamente en crear colecciones digitales y dar acceso al patrimonio cultural propio sino en vincular los objetos culturales con los de otras instituciones patrimoniales. Un museo que posee algunas obras de un artista, puede vincularse con otros museos que poseen otras obras del mismo artista y con un archivo que posee documentos biográficos del artista, todo lo cual permite una

5 <https://openglam.org/>

6 <https://www.w3.org/WAI/standards-guidelines/es#intro>

7 https://es.wikipedia.org/wiki/Tratado_de_Marrakech

navegación extramuros que enriquece el conocimiento en cuanto a integrar otros elementos que no están en la institución.

La web semántica a través de un modelo de datos para la representación de información en la Web, Resource Description Framework, RDF⁸, permite aportar semántica a estos datos, ir más allá del dato, para representar un concepto: un pintor, una pintura, un país. Si el concepto tiene un identificador que permite ubicarlo, entonces puedo empezar a construir un mapa o un grafo que vincule conceptos, otros pintores del mismo país, otras pinturas del mismo pintor. Los conceptos se vinculan por propiedades que también pueden estandarizarse. Por ejemplo el cuadro *X utiliza una técnica*: el óleo, donde *utiliza una técnica* es una propiedad, y óleo es un objeto, la nominación estandarizada es una taxonomía de una técnica específica. Si se estandarizan los conceptos, las propiedades y los objetos, entonces se construye una herramienta normalizadora más poderosa que una taxonomía o un tesoro, una ontología.

Al sumar a toda esta parafernalia la apertura estamos en el terreno de los datos abiertos enlazados (Linked Open Data)⁹, que tejen un grafo de conceptos, propiedades y objetos, que es buscable, navegable, y en un entorno que va más allá de una biblioteca, un museo, un archivo y sus colecciones digitales. Un entorno más amplio permite ir enlazando la colección digital de un museo con otras colecciones digitales análogas o relacionadas de otros museos o bibliotecas, que contienen información complementaria. De esta forma se va armando un mundo de información semánticamente conectada, información que se vuelve más visible y puede ser reutilizada, enriquecida y utilizada en la web de maneras innovadoras, a través de servicios que explotan nuevos nichos.

Todos estos componentes han sido estandarizados en forma de normas, especificaciones y recomendaciones por la W3C.

Las colecciones digitales cada vez más dejarán de estar confinadas en instituciones para desarrollar la interoperabilidad semántica con otras colecciones digitales abiertas.

2.4. Gestión de colecciones digitales

Los recursos en las colecciones digitales atraviesan distintos procesos: ingreso, monitoreo, control, aseguramiento de la calidad y servicios. Estos procesos son generalmente implementados a través de aplicaciones: los sistemas de colecciones digitales. Los procesos de gestión tienen que ver con políticas institucionales y con todos los aspectos de planificación, monitoreo y ajuste, de modo que entallan cada proyecto.

Las aplicaciones automatizan los flujos de trabajo y la seguridad de la información que se establezcan. Se puede controlar el acceso verificando derechos apropiados y guardar las trazas que van dejando los procesos. Trataremos las aplicaciones para CD en la siguiente sección.

2.5 Servicios: Exposiciones virtuales de fondos archivísticos, museísticos, servicios de colecciones audiovisuales, servicios de valor agregado de todo tipo.

Los servicios de colecciones digitales según Dempsey (2014) constituyen una mezcla de servicios locales, externos y colaborativos contruídos alrededor de las necesidades y los comportamientos de los usuarios. Estos servicios se van orientado y ajustando cada vez más rápidamente.

Los servicios que se ofrecen sobre colecciones digitales son entonces muy variados. Se puede considerar por ejemplo un servicio, la capacidad de descargar los recursos digitales en formatos apropiados y en el caso de documentos audiovisuales con una resolución apropiada para su reutilización. Existen gestores de colecciones digitales que ofrecen la visualización de la imagen en formato miniatura o intermedio y luego la descarga en un formato con mayor resolución es un servicio pago. Estos servicios lindan con las aplicaciones de comercio electrónico pero en este caso los productos son documentos audiovisuales. Los bancos de imágenes son un tipo de CD de amplio uso y tienen la ventaja de que entre sus metadatos tienen claramente establecido el licenciamiento, de modo que su reutilización es sencilla.

Los servicios pueden estar orientados a los procesos institucionales y a otras aplicaciones de negocio de las cuales las colecciones digitales pueden ser un emergente. Por ejemplo, una aplicación de gestión de carreteras que controla las fallas en las rutas para hacer un plan de obras o reclamar a un proveedor, puede funcionar en conjunto con una colección digital de fotografías de las fallas en rutas - con los metadatos apropiados - que lleve el registro de auditoría y prueba para los reclamos con la prueba documental. Es fundamental en estos servicios digamos “de acople” lograr interoperabilidad a través de la adhesión a estándares abiertos y el análisis de metadatos apropiados.

8 <https://www.w3.org/RDF/>

9 https://www.w3.org/egov/wiki/Linked_Open_Data

Las exposiciones virtuales según Alcaraz Martínez (2014) son un instrumento de presentación, interpretación y difusión de las colecciones, objetos de interés patrimonial y conceptos socioculturales. La exposición es una narrativa de un tema, dispone los objetos y construye una interpretación sobre los mismos. Requiere una planificación y un diseño y es un servicio importante para las instituciones.

2.6 Preservación digital

Las instituciones deben abordar las cuestiones de la preservación digital y encaminarse en la definición de un plan de preservación digital que tendrá que incluir a los activos de sus colecciones digitales. El objetivo de la preservación digital es asegurar el acceso, la legibilidad y la autenticidad de los recursos digitales a lo largo del tiempo. Este plan abarcará la definición de políticas, estrategias, procesos, formatos, sistemas, que confluyan en el fin establecido.

El estándar 14721 (OAIS)¹⁰, los metadatos PREMIS¹¹ y las directrices para la preservación¹², en conjunto con el esquema METS¹³, constituyen el marco ideal para la gestión de un sistema de información (v.g. colección digital) para asegurar su interoperabilidad y dar preservación a sus contenidos.

Los aspectos de seguridad de la información tienen que ver con la garantía de acceso a un documento que no ha sido modificado y como contraparte el resguardo de acceso solo a los perfiles establecidos de acuerdo a la protección de datos por razones de privacidad o legales.

Los sistemas de colecciones digitales permiten la creación de perfiles de usuarios con permisos específicos de modo de controlar el acceso a los mismos y determinar las opciones en cuanto a la manipulación (lectura, escritura) y también se habilita el rastreo de las operaciones realizadas sobre los recursos a través de registros de log y auditoría.

2.7. Interactividad

Los aspectos de interacción de los usuarios con los contenidos digitales de la colección y a través de servicios de valor agregado son un elemento a planificarse y diseñarse con detalle y es un aspecto vinculado a los objetivos de la colección.

Dentro de los objetivos de la colección, está la definición de los usuarios actuales o potenciales y la manera como esta colección se irá entallando a las necesidades de los mismos. La CD se adecuará a los cambios de los objetos a lo largo de la vida de la colección y a cuales serán los canales de comunicación con los usuarios. La interacción puede ir desde la posibilidad de realizar comentarios, hasta la ingesta de objetos digitales pasando por formas variadas dependiendo de los objetivos de las CD.

Las CD que ofrecen servicios sociales y que se involucran con el entorno en forma importante seguramente diseñarán niveles de interacción mayor, ya que la colección actuará como un mediador en el contexto. Como un caso interesante se comenta la experiencia de rescate de la historia de un barrio de Montevideo, “Puntas de Manga”¹⁴, en este caso el diseño de los objetos y su vinculación es también un aspecto a destacar.

3. Aplicaciones

Existen muchas aplicaciones de software libre que pueden utilizarse en el marco de las colecciones digitales, que operan de forma diferente o son más apropiadas para algunos tipos de colecciones en especial, por lo que en cada caso habrá que considerar cual es la aplicación más apropiada.

Sin ninguna pretensión de ser exhaustivos, ni realizar comparativas, se relevan algunas de estas aplicaciones a modo de ejemplo de manera muy sucinta.

10 <https://www.dpconline.org/docs/technology-watch-reports/1359-dpctw14-02/file>

11 <https://www.dpconline.org/docs/technology-watch-reports/894-dpctw13-03/file>

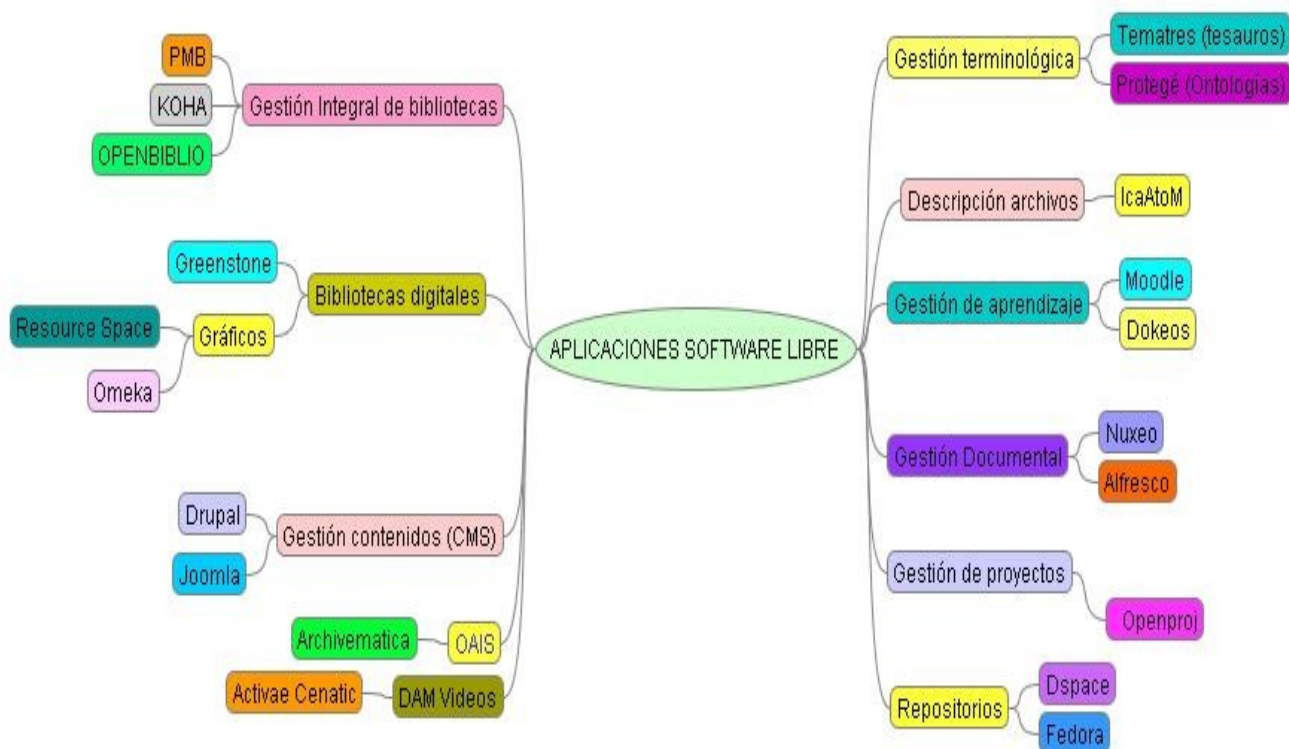
12 https://unesdoc.unesco.org/ark:/48223/pf0000130071_spa

13 https://www.loc.gov/standards/mets/METSOverview_spa.html

14 http://ji.fic.edu.uy/wp-content/uploads/2018/07/GT3_Rodriguez_Placeres_Rodriguez_Inchausti.pdf

En lo que sigue en la figura 2 se muestran algunas categorizaciones de sistemas, todos ellos de software libre.

Figura 2. Aplicaciones de software libre



3.1 Greenstone

Sitio: http://www.greenstone.org/index_es

Es un conjunto de programas para crear y distribuir colecciones de bibliotecas digitales. Es multilingüe. Lo produce la Universidad de Waikato (Nueva Zelanda) y ha sido desarrollado y distribuido en cooperación con la Unesco.

- Licencia: Software libre, Licencia GNU GPL
- Interoperabilidad: Soporta OAI-PMH, METS
- Metadatos: Los usuarios definen los metadatos en forma interactiva dentro de la interface del bibliotecario. Existen formatos predefinidos: Dublin Core (calificado y no calificado), RFC 1807, NZGLS (New Zealand Government Locator Service), AGLS (Australian Government Locator Service).
- Formatos de documentos: Tiene plugins para ingestar documentos:
- Texto: PDF, PostScript, Word, RTF, HTML, texto plano, Latex, archivos ZIP, Excel, PPT, Email (varios formatos)
- Multimedia: imágenes, cualquier formato incluyendo PNG, JPEG, JPEG 2000, GIF, TIFF
- Audio: MP3 audio, Ogg Vorbis audio, y un plugin genérico que puede ser configurado para formatos de audio: , MPEG, MIDI, etc.

3.2 Tainacan

Sitio: <https://tainacan.org/>

Se trata de un plugin para Wordpress que permite la gestión de colecciones digitales.

Es un proyecto desarrollado por la Universidad Federal de Goiás (UFG), Universidad de Brasilia (UnB) y el Instituto brasileiro de información en ciencia y tecnología (IBICT) desde el 2014 con el auspicio del Instituto brasileiro de Museos y la Fundación Nacional de las Artes.

Tiene más de 200 instalaciones activas reportadas.

<https://www.museus.gov.br/acessoainformacao/acoes-e-programas/projeto-tainacan/>

Licencia: software libre

Metadatos: Puede usarse metadatos estándar o aportar los metadatos específicos para describir los elementos de la colección.

Taxonomías: maneja vocabularios que pueden ser usados por todas las colecciones y se pueden configurar en forma jerárquica.

Interoperabilidad: Maneja una API RESTful (lectura y escritura) que permite a otras aplicaciones interactuar con las colecciones. Las colecciones pueden desplegarse en diferentes formatos, json, jsonLD y OAI-PMH entre otros.

Documentación: <https://wiki.tainacan.org/>

Se proyecta el desarrollo de servicios que integren colecciones museológicas y datos abiertos enlazados.

3.3 Omeka

Sitio: <http://omeka.org/>

Omeka es una herramienta libre para publicar en la web archivos digitales. El foco está en la publicación web y las exhibiciones.

La aplicación tiene dos partes, una trastienda donde se ingresan los objetos y su información y una vitrina en donde se pueden armar exhibiciones.

Omeka tiene muchos plugin que ofrecen funcionalidades diversas y uno en particular que ofrece las funcionalidades de OAI-PMH para que las colecciones puedan ser recolectadas.

Metadatos: Dublin Core, MODS y posibilidad de extensión.

3.4 Resourcespace

Resource space es una aplicación que cumple con prestaciones muy similares a OMEKA. Se denomina gestor de colecciones digitales y es software libre.

El recurso digital se enfoca en un elemento con todas las posibles variantes y los metadatos asociados en una jerarquía plana que puede vincularse a contenedores mayores, las colecciones en realidad son agrupaciones de marcas, no son contenedores.

Algunas colecciones se pueden caracterizar como estanterías o exhibiciones.

Los metadatos son fundamentales en la gestión y organización de las colecciones y existe mucha flexibilidad para la extensión de los mismos en esta aplicación.

Tiene muchos plugin sobre cuestiones diversas y tiene la posibilidad de actuar como comercio electrónico con la gestión de envío y pedidos de recursos.

El sitio es <https://www.resourcespace.com/>

Hay un caso de implementación de interés en el Museo de arte Walters ¹⁵.

3.5 Collective Access

CollectiveAccess es un conjunto de aplicaciones basado en la web que proporciona un marco para la catalogación y el descubrimiento y permite crear descripciones que se acerquen a necesidades muy variadas. Tiene dos componentes, una interfaz de descripción de objetos y una plataforma de publicación y descubrimiento. La publicación se puede hacer con otras aplicaciones como Drupal u Omeka.

15 https://www.resourcespace.com/downloads/casestudies/ResourceSpace_Walters-case_study.pdf

La interfaz de ingreso es altamente configurable y puede usarse para catalogar casi cualquier cosa. No hay campos codificados y todos los campos de todo el sistema se pueden personalizar mediante programación o mediante la interfaz de usuario. Agregar nuevos campos, eliminar campos e incluso reorganizar campos se puede hacer fácilmente en cualquier etapa de un proyecto.

Estos campos configurables pueden llevar una variedad de configuraciones y admitir muchos tipos de datos diferentes. Esto significa que la aplicación reconoce y normaliza los valores de datos y mejora las funciones de búsqueda y navegación.

Los esquemas de Collective Access incluyen muchos estándares preconfigurados como PBCore, DublinCore, VRA, DarwinCore, CDWA Lite, MARC y otros. También incluye una selección de esquemas personalizados específicos de uso, como un archivo fotográfico basado en la ubicación, un archivo documental o un archivo de exhibición.

Uno de los aspectos clave de esta aplicación es su capacidad para crear relaciones entre registros. Las relaciones comunes pueden incluir: Entidad relacionada con un Trabajo, Objetos relacionados con un Evento u Ocurrencia, entre otras.

Sitio: <https://www.collectiveaccess.org/>

Software libre

3.6 Atom

Sitio: <https://www.accesstomemory.org/es/>

Atom es una aplicación para descripción archivística, multilingüe y multirepositorio. Está disponible bajo la licencia de software libre A-GPL.

Fue promovido por la International Council o Archives (ICA) y la UNESCO.

Está basado en estándares: ISAD(G), ISAAR(CPF), ISDIAH, ISDF, SKOS y soporta Dublin Core, RAD, MODS y DACS.

Permite adosar los objetos digitales.

3.7 Tematres

Esta aplicación no es para el tratamiento de colecciones digitales sino como una aplicación complementaria para trabajar con las taxonomías de la colección.

La indización es un elemento vertebrador de informaciones dispersas, es un factor de interoperabilidad que nos permite descubrir y navegar por facetas de la información en formatos diversos, en lugares distintos, con tratamientos distintos. La planificación de la organización del conocimiento como un elemento subyacente debe pensarse desde la etapa del diseño. Desde allí considerar la arquitectura y la disposición de los servicios para construir los necesarios acoples entre mundos transitables que pueden ser conectados. Pensar esa interoperabilidad necesaria es utilizar los estándares de todo tipo que subyacen a su implementación: SKOS, XML, RDF, ontologías, entre otros.

TemaTres es una aplicación Web para la gestión de lenguajes documentales. Se encuentra orientada especialmente al desarrollo de tesauros jerárquicos, pero también puede utilizarse para desarrollar estructuras de navegación web, o como complemento para articular con un gestor de contenidos, una biblioteca digital, o inclusive en una biblioteca tradicional para gestionar los lenguajes documentales en uso.

TemaTres ofrece las siguientes funcionalidades:

- Relaciones entre términos (equivalencia, jerárquicas y asociativas)
- Soporte para vocabularios polijerárquicos
- Navegación sistemática y alfabética
- Exportación del tesoro en XML, Zthes, RDF, SKOS

Sitio:

<https://www.vocabularyserver.com/>

<https://r020.com.ar/tematres/manual/>

<https://sourceforge.net/projects/tematres/>

3.8 Publicación digital

Una revista electrónica, con una mirada flexible, puede considerarse una colección digital con características especiales y procesos similares a los de los repositorios.

Sin internarse en aspectos que tienen que ver con la socialización de la ciencia y la apertura de lo que se publica o en como estas CD transitan por los canales ortodoxos o por los atajos abiertos a machetazos en las estructuras de poder, puede establecerse que tienen como servicio primordial la publicación digital.

Este tipo de CD están en una etapa de grandes cambios, tanto en el servicio de publicación, como en los elementos que componen la colección que empiezan a remezclarse con formatos de todo tipo, construyendo una experiencia multisensorial y de trayecto personalizado. Un aspecto que puede señalarse es que en este tipo de CD los derechos y la licencia tienen un papel principal.

Open Journal Systems (OJS)

Sitio: <https://pkp.sfu.ca/ojs/>

OJS es un sistema de gestión y publicación de revistas electrónicas, que asiste en cada etapa del proceso de publicación referado, desde el envío, la revisión por pares, la maquetación y la publicación.

Es software libre y tiene el propósito de hacer que la publicación en abierto sea una opción viable que permita el acceso contribuyendo así al bien público.

Los objetos de estas colecciones digitales son normalmente textuales o mayoritariamente textuales con algunas imágenes. Con el tiempo los objetos de estas colecciones digitales empiezan a sufrir transformaciones y remezclas y los artículos empiezan a convertirse en un objeto que tiene video, imagen, fotografía, texto, mapas interactivos, hipervínculos, entre otros elementos, e incluso dejan de tener un guión secuencial y se convierten en objetos de navegación errática. Lo que Cortázar¹⁶ nos propone en *Rayuela* o *62 modelo para armar*, con sus trayectos a definir por el lector empieza a suceder en estas colecciones pero de una forma más intensiva en los formatos como en la propuesta de Scalar.me

Scalar.me

Sitio:

Es una plataforma de publicación semántica que permite publicar contenido enriquecido con multimedia.

La Alliance for Networking Visual Culture (ANVC)¹⁷ señala que busca expandir las prácticas visuales y reformatearlas con la cultura digital, creando contexto para el uso de medios digitales. En esta iniciativa tienen como asociados instituciones archivísticas, bibliotecas, centros culturales y prensa universitaria.

Esta propuesta busca acortar la distancia entre los archivos visuales digitales y la publicación académica, creando trayectos interpretativos a través de los materiales y permitiendo nuevas formas de análisis. Exploran la relación ente el análisis académico y las tipologías digitales o géneros emergentes. Se analiza la organización de CD para facilitar el uso explorando como pueden establecerse flujos entre la evidencia, la investigación y la publicación.

Este proyecto se relaciona con la academia, la prensa, las bibliotecas y los archivos y busca identificar un tipo amplio de comunicación académica emergente modelando publicaciones interactivas y multimodales.

Scalar es de fuente abierta.

Soporta una variedad muy amplia de formatos y lo que resulta muy interesante es la manera en que se estructura el contenido a través de los trayectos y las etiquetas como se muestra en la figura 3.

16 https://es.wikipedia.org/wiki/Julio_Cort%C3%A1zar

17 <https://scalar.me/anvc/about/>

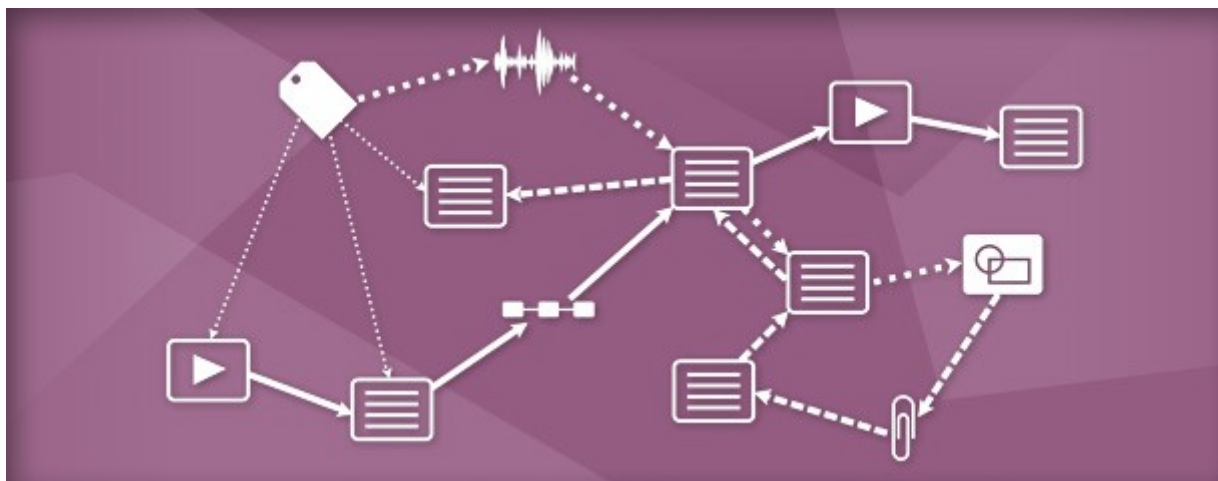


Figura 3. Trayectos en la lectura Extraída de https://scalar.me/anvc/wp-content/uploads/2012/07/feat_structure_banner.jpg

Los trayectos son secuencias lineales de contenido, como una sección, mientras que las etiquetas son agrupaciones de contenido no lineal, como descriptores. Lo que es interesante es la flexibilidad con que se puede transitar las colecciones.

La estructuración puede ser a la vez jerárquica o rizomática. La estructuración jerárquica es la que estamos acostumbrados en un modelo arbóreo, como las taxonomías, mientras que la aproximación rizomática no tiene subordinación, es un elemento heterogéneo que puede conectarse con cualquier otro.

Esto significa que cualquier pieza de contenido puede ser trayecto, etiqueta o ambos y que puede referenciar cualquier otro contenido: texto, video, audio, imágenes.

El modelo se basa en RDF, el estándar de la web semántica. y las relaciones se expresan usando ontologías y estándares como Dublin Core.

3.9 Publicación de datos. Colecciones de datos.

Los datos son un tipo de contenido no tradicional, pero cada vez más se señala la importancia de publicar los datos a nivel de gobierno electrónico como parte de la transparencia, el buen gobierno y la posibilidad de que terceros desarrollen servicios con estos datos.

Por otra parte los datos de investigación son sustantivos para avalar la investigación realizada o para permitir que se continúe investigando con nuevas iniciativas. A partir de lo ya estudiado se pueden proponer nuevos estudios reutilizando los datos de investigaciones previas. Permitir el acceso a los datos de investigación facilita el avance científico.

A partir de estas prácticas aparece la denominación de bibliotecario de datos o archivista de datos y también se protocolizan todos los procesos necesarios para gestionar un plan de datos en un proyecto de investigación, incluido su depósito en una colección de datos.

Como señala la Open Knowledge Foundation :

“Los datos abiertos son datos que pueden ser utilizados, reutilizados y redistribuidos libremente por cualquier persona, y que se encuentran sujetos, cuando más, al requerimiento de atribución y de compartirse de la misma manera en que aparecen.”¹⁸

Dataverse

Sitio: <https://dataverse.org/>

Dataverse es una aplicación de software libre para compartir, preservar, citar, explorar y analizar datos de investigación.

18 <https://opendatahandbook.org/guide/es/what-is-open-data/>

Publicar los datos de investigación hace que los investigadores y las publicaciones electrónicas reciban crédito académico y aumentan su visibilidad.

Un repositorio Dataverse aloja múltiples archivos que se denominan dataverse. Cada dataverse contiene conjuntos de datos y cada conjunto de datos se forma con los archivos de datos propiamente dichos y los metadatos descriptivos, incluyendo documentación y código que acompaña a los datos. Cada dataverse puede contener a su vez otros dataverses.

Los desarrolladores de este software forman parte de la Universidad de Harvard, el Instituto para la Ciencia Social Cuantitativa (IQSS) junto con la biblioteca de la Universidad con colaboraciones a lo largo del mundo. King (2007) recomienda a los autores de revistas electrónicas subir los datos y la documentación relacionada a un repositorio dataverse. La aplicación proporciona una cita formal a sus datos que incluye un identificador global. Recomienda además que la cita a los datos se incluyan en los curriculum vitae en una sección específica junto con las citas de artículos.

CKAN

Sitio: <https://ckan.org/>

CKAN es una aplicación de software libre para la gestión de datos, que proporciona acceso a los mismos y constituye una herramienta para publicación, compartir, buscar y usar datos abiertos.

Muchos portales de datos de gobierno están publicados en esta aplicación.

A cada conjunto de datos que se deposita se le incorporan una serie de metadatos, por los cuales luego serán buscables. Estos metadatos incluyen el título del conjunto de datos, un identificador único para el mismo, una descripción, licencia, etiquetas (descriptores), entre otros.

Se permite la ingesta de varios formatos de datos, aunque se recomienda el depósito en formatos abiertos que faciliten la preservación y la interoperabilidad.

Tim Berners Lee ¹⁹desarrolló la recomendación de las 5 estrellas de datos abiertos que va desde un primer nivel con formatos que no permiten la manipulación y la desagregación, un segundo nivel que ofrece datos estructurados, un tercer nivel que son los datos en formatos abiertos, el cuarto que incluye la semántica en los datos (identificador y propiedades) y el mayor nivel -las cinco estrellas – son datos semánticos que se enlazan con otros datos, dando contexto a los mismos.

Lo que se persigue es la utilización y reutilización de los datos, de modo que la idea de apertura es fundamental.

4. Conclusiones

Las colecciones digitales están en constante cambio y constituyen objetos fronterizos que congregan a distintos profesionales y disciplinas por lo que esta reflexión busca ofrecer un punto de vista que aglutine los aspectos que se consideran más destacados, sin la pretensión de dar por finalizada la cuestión. Este trabajo tiene el espíritu de una propuesta para el debate de lo planteado.

Lo que se pretende es dar un enfoque que permita abordar estos temas de forma amplia y desde una praxis que ofrece una línea de acción posible.

La diversidad de las colecciones digitales en cuanto a sus fines, sus objetos, sus formatos y los enfoques es muy amplia, pero en todos los casos existirán ciertos elementos de diseño con los cuales los profesionales de la información estamos familiarizados y podemos establecer una mirada contextualizada y flexible.

Existe una tendencia en la comunidad de la ciencia de la información que adhiere a una visión filosófica y ética en la apertura. Los servicios abiertos que promuevan la equidad y el acceso universal pueden ser sustentados por una amplia oferta de software libre para el tratamiento de las colecciones digitales.

Una hibridación de servicios, tratamientos y enfoques puede resultar muy apropiada para situaciones remezcladas. Las colecciones digitales están evolucionando todo el tiempo, tanto en sus elementos componentes como en sus prestaciones y en la forma de navegar y bucear en ellas. Lo importante es elegir adecuadamente los estándares que permitan la navegabilidad y la interoperabilidad y tener siempre en cuenta las necesidades de los usuarios.

Bibliografía

- Alcaraz Martínez, R. (2014). Exposiciones virtuales para difundir colecciones digitales. <http://eprints.rclis.org/23236/>
- Alvite Díez, M. L., Pena Sueiro, N. (2019). Colecciones digitales patrimoniales especializadas. Estudio de la Red ARACNE. 14º Congreso ISKO España (4º ISKO España-Portugal) http://eprints.rclis.org/38563/1/ARACNE_Preprint%20ISKO_Spain2019.pdf
- Anderson, T. (2007). Settings, arenas and boundary objects: socio-material framings of information practices. *Information Research*, 12(4). <http://www.informationr.net/ir/12-4/colis/colis10.html>
- Bowker, G. C.; Star, S. L. (1999). *Sorting Things Out: Classification and Its Consequences*, Cambridge, MA: MIT Press. ISBN 978-0-262-02461-7
- Calhoun, K. (2014) Exploring digital libraries: foundations, practice, prospects. Londres: Facet
- Chan, K.M., Guerry, A.D., P Balvanera, P., Klain, S, Satterfield, T., Basurto, X., Bostrom, A., Chuenpagdee, A., Gould, R., Halpern; B. (2012) Where are cultural and social in ecosystem services? A framework for constructive engagement. *Bioscience*, 62, pp. 744-756 <https://academic.oup.com/bioscience/article/62/8/744/244312>
- Dempsey, Lorcan (2014) "The Facilitated Collection," *OCLC*, January 2014, <http://orweblog.oclc.org/towards-the-facilitated-collection/>
- Ferrer, Antonia; Lloret, Nuria ; Peset, María y Moreno, María (2002). Presentación del proyecto: Guía metodológica para la implantación de una biblioteca Digital Universitaria. *Revista general de información y documentación*, 12(2), 479-494. <https://core.ac.uk/download/pdf/38822312.pdf>
- King, Gary. 2007. "An Introduction to the Dataverse Network as an Infrastructure for Data Sharing." *Sociological Methods and Research*, 36, Pp. 173–199. <http://j.mp/2owjuRr>
- Méndez Rodríguez, Eva (2001) Metadatos y recuperación de información: estándares y aplicabilidad en bibliotecas digitales. https://e-archivo.uc3m.es/bitstream/handle/10016/26863/tesis_eva_mendez_2001.pdf
- Rieger, Oya (2019). What's a Collection Anyway? Ithaka S+R Issue brief. <https://sr.ithaka.org/publications/whats-a-collection-anyway/>
- Sorli Rojo, A., Merlo Vega, J. A. (2000). Bibliotecas digitales (I): colecciones de libros de acceso público. <https://digital.csic.es/bitstream/10261/9468/1/redcbd1.pdf>
- Star, Susan; Griesemer, James (1989). "Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39". *Social Studies of Science*. 19 (3): 387–420. doi:10.1177/030631289019003001.
- Steger, C., Hirsch, S., Evers, C., Branoff, B., Petrova, M., Nielsen-Pincus, M., ... & Van Riper, C. J. (2018). Ecosystem services as boundary objects for transdisciplinary collaboration. *Ecological Economics*, 143, 153-160. <https://www.sciencedirect.com/science/article/pii/S0921800917304664>
- Tramullas, J. (2007). Bibliotecas digitales. <http://eprints.rclis.org/11304/1/oleirosbd.pdf>
- Tramullas, J. (2002). Propuestas de concepto y definición de la biblioteca digital. http://eprints.rclis.org/15118/1/04_2002.pdf
- Witten, I. H., Bainbridge, D., Nichols, D. M. (2009). *How to build a digital library*. Morgan Kaufmann.
- Wittmann, R., Neatrou, A., Cummings, R., & Myntti, J. (2019). From Digital Library to Open Datasets. *Information Technology and Libraries*, 38(4), 49-61. <https://ejournals.bc.edu/index.php/ital/article/view/11101/9837>