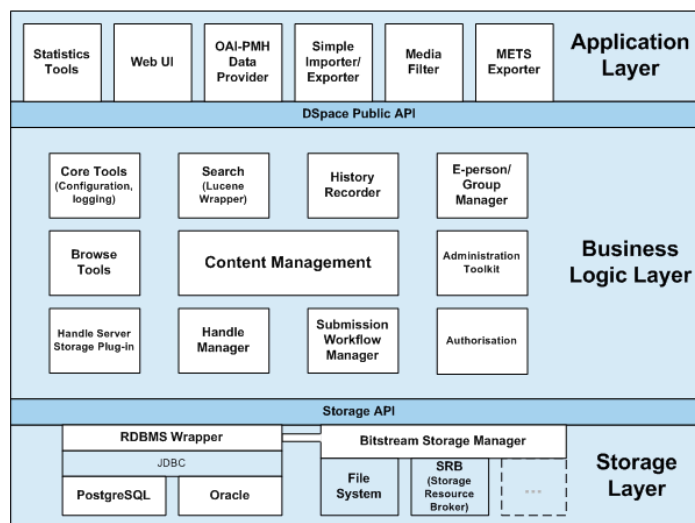




Clase 2. Aspectos tecnológicos

- 1) Defina concisamente las tres capas principales que componen la arquitectura de DSpace. Describa el modelo de contenidos completo de DSpace (colecciones, comunidades, items, etc.).

DSpace es un paquete de software de código abierto que se utiliza para la creación de repositorios destinados a la preservación y provisión de objetos digitales. La arquitectura de tres *layers* de DSpace se presenta en la siguiente imagen:



Disponible en: <https://wiki.lyrasis.org/display/DSDOC6x/Architecture>

Cada *layer* tiene “confianza” e invoca al inferior, muy resumidamente cada *layer* tiene la siguiente función:

<i>Application</i>	Responsable de atender y comunicarse con usuarios, aplicaciones, servicios externos a la instalación de DSpace. (por ejemplo: usuarios de la interface web, repositorios (mediante OIA-PMH), aplicaciones (mediante Sword).
<i>Business Logic</i>	Responsable de la configuración general de la lógica de funcionamiento de DSpace para una institución definida. Entre otras funciones se configuran y gestiona procesos de trabajo, autenticación, autorización de acciones, plugins, metadatos, e-personas (usuarios registrados), entre otros.
<i>Storage</i>	Responsable del almacenamiento físico del contenido y los metadatos.

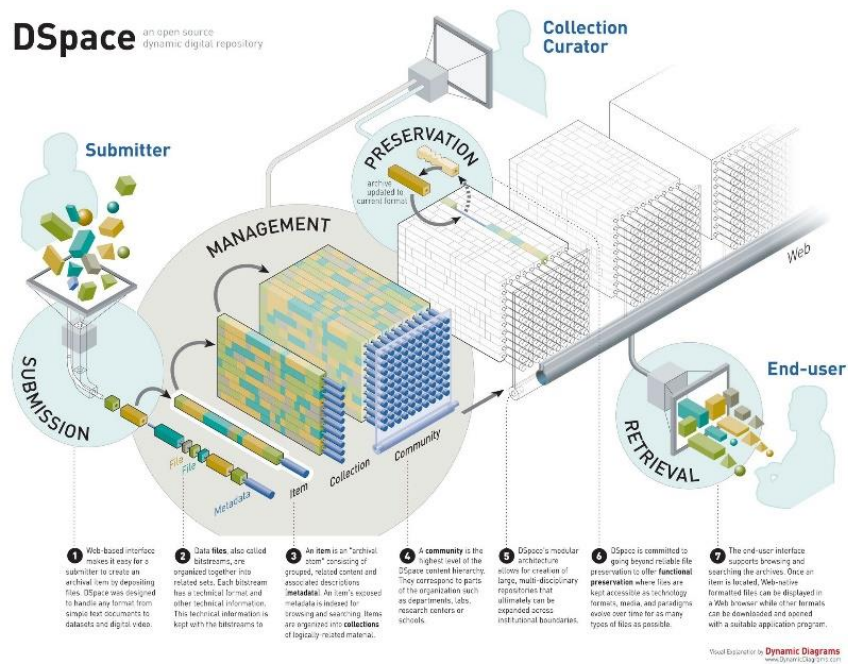
El modelo de contenidos de DSpace está formado por:

Comunidades y Subcomunidades	Conforman espacios de trabajos organizados jerárquicamente que permiten adecuar la instalación de DSpace al modelo organizativo y de gestión de la institución en la que está implementado (Departamentos, Facultades, Áreas, etc.). A estas se pueden asignar usuarios con permisos / perfiles. Cada comunidad contiene metadatos descriptivos de sí misma y las colecciones que contiene.
------------------------------	---



Colecciones	Agrupan y presentan contenido relacionado (por tema, área, grupo, etc.) perteneciente a una comunidad o subcomunidad (o a varias (colecciones compartidas)). Contienen metadatos descriptivos de sí misma y su contenido.
Items	Refiere al registro almacenado en el repositorio. Está formado por metadatos descriptivos del recurso y archivos para un contenido específico.
Bundles	Los bundles conforman los ítems. Agrupan bitstreams de acuerdo a criterios
Bitstreams	Los bitstreams son objetos digitales que representan la obra publicada o que se desea publicar (documentos, planillas, videos, audios, imágenes, software, etc.). Los bitstreams contienen información de validación, descripción y preservación de cada ítem.
Bitstore	Donde se alojan los bitstreams. Estos pueden ser locales (en el servidor o almacenamiento) o en cloud (Amazon)
Metadatos	Datos estructurados que permiten describir a otros datos.

La siguiente imagen presenta la relación conceptual de los componentes de la arquitectura de DSpace con los diversos perfiles de uso del repositorio:



Disponible en: https://library.rverson.ca/open/files/2017/08/DSpaceBackgrounder_July112017-2.pdf



- 2) Describa al menos 2 protocolos de interoperabilidad utilizados en DSpace. Basándose en el estándar del protocolo OAI-PMH (<https://www.openarchives.org/OAI/openarchivesprotocol.html>), especifique los tipos de proveedores existentes y la definición de Repositorio.

Protocolos de interoperabilidad:

Simple Web Service Offering Repository Deposit (SWORD): (<http://swordapp.org/>) es un protocolo de interoperabilidad desarrollado con el objetivo de “facilitar la tarea de depositar” en repositorios. Este permite a los autores depositar objeto digital desde una aplicación (cliente o web) en uno o más repositorios. Actualmente en su versión 3, permite apoyar el ciclo de vida completo del proceso de depósito (depositar, actualizar, reemplazar, eliminar). Se definen los siguientes casos de uso:

- Editor a Repositorio
- Sistema de información de la investigación a Repositorio
- Escritorio a Repositorio
- Repositorio a Repositorio
- Interface de usuario especializada a Repositorio
- Equipo de laboratorio a Repositorio
- Ingestión a granel en el Repositorio
- Autoría colaborativa

OAI-PMH (<https://www.openarchives.org/pmh/>) es el protocolo de interoperabilidad entre repositorios de acceso abierto. Es parte de los estándares para la implementación de RI, adhiere el esquema de metadatos Dublin Core sin calificar (DC) y utiliza transacciones HTTP (Hyper Text Transfer Protocol) para la transferencia de contenidos Web.:

Definición de repositorio:

De acuerdo al estándar del protocolo OAI-PMH un repositorio se define como: un servidor accesible en red que puede procesar las 6 solicitudes OAI-PMH de la manera descrita en el estándar. En un repositorio un proveedor de datos expone los metadatos (de los objetos depositados) a los recolectores.

Proveedores:

- *Data providers*: sistemas que soportan OAI-PMH y exponen metadatos a,
- *Service providers*: sistemas que utilizan los metadatos cosechados por OAI-PMH para crear servicios de valor agregado.
 - Harvesters: son operados por *service providers* para la cosecha de metadatos vía OAI-PMH



3) Describa la relación existente entre un metadato y una autoridad vinculada. Enumere las ventajas de utilizar metadatos controlados por autoridad.

Algunos conceptos y definiciones:

- Una autoridad es un conjunto de valores controlados para un dominio determinado, estando cada valor único identificado por una clave (clave de autoridad).
- Un registro de autoridad es la información asociada con cada uno de los valores en una autoridad (incluyendo variaciones de deletreo, valores equivalentes y/o alternativos, etc). Un registro de autoridad (de nombres de autor, por ejemplo), contiene la forma autorizada del nombre del autor definida por la institución. Puede definir también las formas variantes del nombre y nombres relacionados.¹
- La Clave de autoridad de un identificador persistente asociado a un registro de autoridad.

Las autoridades permiten el uso de vocabularios (nombres, instituciones, etc.) controlados y gestionados entre sistemas (p.e.: repositorios). Se utilizan para realizar la carga de metadatos mediante “listas u opciones” predefinidas. Pueden ser cerrados (restringidos) o abiertos (permite agregar otros valores a los definidos en la lista).

La relación entre un metadato y una autoridad vinculada facilita la carga de los metadatos al brindar opciones para su carga a partir de valores (listas) o consultas (*Choice management*). Facilita la recuperación de la información mediante *Facets* de búsqueda utilizados por el usuario al realizar búsquedas.

1. Choice mangement

Autor Institucional:
Institución creadora del contenido intelectual de la obra

- Universidad de Salamanca (USAL)
- AGH University of Science and Technology (AGH)
- Benemérita Universidad Autónoma de Puebla (BUAP)
- Cadi Ayyad University (UCA)
- Centro de Investigación en Economía Política y Comunicación (CIEPYC)
- Editorial de la Universidad Nacional de La Plata (Edulp)
- Facultad de Periodismo y Comunicación Social
- Hanoi University of Mining and Geology (HUMG)
- Ibn zohr University (UIZ)
- Kyoto University (KYOTOU)

2. Facets de búsqueda

Materia
Ciencias Físicas (84)
Óptica, Acústica (40)
Física (23)
Ingeniería de los Materiales (12)
Ciencias Químicas (9)
Diseño Arquitectónico (8)
Ingeniería de los Materiales (7)
Física Atómica, Molecular y Química (3)
Ingeniería Eléctrica y Electrónica (3)
Físico-Química, Ciencia de los Polímeros, Electroquímica (2)
... ver más

Entre las ventajas de utilizar metadatos controlados por autoridad, podemos citar:

- Dar consistencia e integridad a los metadatos de los objetos digitales, mejorando su calidad.
- Mejorar la precisión en la recuperación de la información disponible en RI y otros sistemas integrados.
- Facilitar el intercambio de información bibliográfica al compartir autoridades.

¹ Extraído de: <http://eprints.rclis.org/20560/1/Ponencia%2021-277-1-PB.pdf>



- Facilita (reduce las decisiones y el tiempo, evita imprecisiones) la carga de metadatos a los objetos digitales.
- Permite probar que dos valores son idénticos al compararlos con la clave de autoridad.
- Mejora la interoperabilidad al compartir nombres de autoridades con otras aplicaciones.

4) Describa brevemente qué es un identificador persistente, explique las ventajas de su uso. Investigue 3 proveedores de identificadores persistentes, indique el tipo de recursos sobre los que se utilizan (artículos, imágenes, personas, etc).

Un identificador persistente (*Persistent Identifier*, PID) es un identificador construido e implementado de manera que el recurso identificado continúe siendo el mismo independientemente de la ubicación de su representación, así como del hecho de que varias copias estén disponibles en diversas ubicaciones (dominios / *path*). Estos se pueden aplicar a documentos de texto, datos de investigación, personas.

Entre las ventajas y características de los identificadores persistentes:

- Unicidad, por la que el esquema de nombres deberá ser único en el contexto de los recursos digitales de que se trate, incluso desde una perspectiva global.
- Compromiso con la persistencia, es decir, con el mantenimiento de la asociación del recurso identificado con su ubicación. Por ejemplo: ante cambios en su URL.
- Simplicidad, lo hará que el sistema sea fácil de comprender y aplicar, prestándose a las citas o referencias breves y sencillas.
- Un recurso puede disponer de más de un identificador persistente (p.e.: uno de publicación y otro al momento del depósito en el repositorio).
- Se utilizan al momento de catalogar y publicar recursos, en citas y menciones de sitios web, blogs, etc.

Entre los proveedores se pueden mencionar los siguientes:

- **Handle System:** (<https://www.handle.net/>) Proporciona un sistema eficiente, extensible y seguro de PID sobre Internet. Los handle no tienen una representación URI nativa ni dependen del protocolo DNS y HTTP. Cada handle es una cadena de caracteres compuesta de: la autoridad de nombres (naming authority) y un nombre local único bajo dicha autoridad (local name) bajo el modelo: {naming_authority} "/" {local_name} por ejemplo: 20.500.1234.
- **Digital Object Identifier (DOI):** (<https://www.doi.org/>) Sistema gestionado por la International DOI Foundation (IDF), se utiliza lo habitualmente para identificar y localizar publicaciones digitales. La norma ISO 26324:2012 especifica su sintaxis, modelo y componentes. IDF garantiza la existencia de un sistema de resolución de DOI utilizando el protocolo HTTP bajo el modelo: http://doi.org/{naming_authority}/{local_name}, por ejemplo: <https://doi.org/10.1109/5.771073>.
- **ORCID:** (<https://orcid.org/>) Provee identificadores persistentes a personas e instituciones. En el caso de personas, permite relacionar su ID con información profesional. El ID se puede compartir con otras personas o sistemas.