



# DSpace:

## Manejo de repositorios abiertos

Presenta:

García Rojas Alan 201765944

Liévana Poy Erick 201742006

Lima Estrada Efraín 201705754

Reyes Reyes Julián 201643331

Benemérita Universidad Autónoma de Puebla  
Facultad de Ciencias de la Computación

October 5, 2021

## Índice

## 1 Resumen

## 2 Introducción

- ¿Qué es DSpace?
- ¿Para qué se usa DSpace?
- Historia
- Logros

### 3 Características

- Arquitectura de la Aplicación
- Motor de Búsqueda Integrado
- Reconoce todos los Tipos de Archivos

- Metadatos
- Herramientas y Plugins
- Seguridad
- Recuperación
- Compilación
- Base de Datos
- Manejo de Archivos

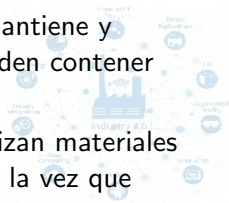
#### 4 Funcionamiento

- Propuesta de Información
- Estructura

## 5 Referencias

# Resumen

- Un repositorio es un espacio centralizado donde se almacena, organiza, mantiene y difunde información digital, habitualmente archivos informáticos, que pueden contener trabajos científicos, conjuntos de datos o software
- Un repositorio abierto son sistemas de información que preservan y organizan materiales científicos y académicos como apoyo a la investigación y el aprendizaje, a la vez que garantizan el acceso a la información
- Los repositorios abiertos tienen sus inicios en los años 90, en el área de la física y las matemáticas, donde los académicos aprovecharon la red para compartir sus investigaciones con otros colegas. Este proceso era valioso porque aceleraba el ciclo científico de publicación y revisión de resultados



# Introducción

# ¿Qué es DSpace?

- Es un software de código abierto que provee herramientas para la administración de colecciones digitales a través de repositorios abiertos.
- Responde a la necesidad específica como sistema de archivos digitales centrado en el almacenamiento, acceso y preservación a largo plazo de contenido digital.



# DSPACE

# ¿Para qué se usa DSpace?

- Comunmente usada para la administración de colecciones digitales o repositorio bibliográfico
- Soporta una gran variedad de datos entre los que destacan:
  - Libros
  - Tesis
  - Fotografías
  - Filmes
  - Videos
  - Datos de Investigación



# Historia

- Fue desarrollado en 2002 por Hewlett-Packard(HP) y el Instituto de Tecnología de Massachusetts(MIT)
- Actualmente es desarrollado y mantenido por DuraSpace
- DSpace esta disponible en mas de 20 lenguajes



**Massachusetts  
Institute of  
Technology**

# Logros

- Dspace es usado por mas de 2500 instituciones al lo largo del mundo, entre las que destacan:
  - El Banco Mundial
  - Universidad de Cambridge
  - Universidad de Harvard
  - Instituto de Tecnología de Massachusetts
  - Imperial College London
  - La Organización Mundial de la Salud





# Características

## Arquitectura de la Aplicación

- Software de Código libre y abierto
- Manejo del FrontEnd y BackEnd
- Base de Datos
- Motor de Búsqueda
- Uso de Metadatos

# Motor de Búsqueda Integrado

- Tiene integrado Apache Solr, un motor de búsqueda que permite:
  - Búsqueda y Recuperación
  - Filtro de resultados
  - El texto completo de los archivos es indexado
  - Búsqueda a travez de metadatos
- Las interfaces de busquedas son personalizables



## Reconoce todos los Tipos de Archivos

- DSpace puede almacenar cualquier tipo de archivos
- Automaticamente reconoce los tipos de archivos mas comunes:
  - DOC
  - PDF
  - XLS
  - PPT
  - JPEG
  - MPEG
  - TIFF

## Metadatos

-

## Herramientas y Plugins

- DSpace provee un conjunto de herramientas para la administración, búsqueda, edición, importación, exportación, etc. de los contenidos del repositorio.
- De igual forma DSpace permite el uso de plugins comerciales



# Seguridad

- DSpace tiene integrado su propio sistema de autenticación y autorización
- De igual forma permite el uso de otros sistemas de autenticación como LDAP o Shibboleth
- Permite la asignación de permisos para leer/escribir, en todo el sistema, por comunidad o por colección, por ítem o por archivo.
- Igual permite la asignación de permisos administrativos por comunidad o por colección



## Recuperación

-



# Compilación

- DSpace compila de acuerdo a los protocolos estandar, asi como con las mejores prácticas para acceder, importar y exportar
- Uso de los protocolos:
  - Open Archives Initiative Protocol for Metadata Harvesting(OAI-PMH)
  - SWORD V2
  - OpenAIRE
  - WebDAV
  - OpenSearch
  - OpenURL
  - RSS
  - ATOM



# Base de Datos

- DSpace permite elegir entre las Bases de Datos:
  - PostgreSQL
  - Oracle
- Esta base puede manejar los archivos y su metadata.

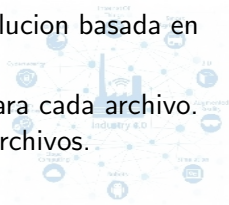


El logo de Oracle, con la palabra 'ORACLE' en una tipografía roja, sans-serif, en mayúsculas.



# Manejo de Archivos

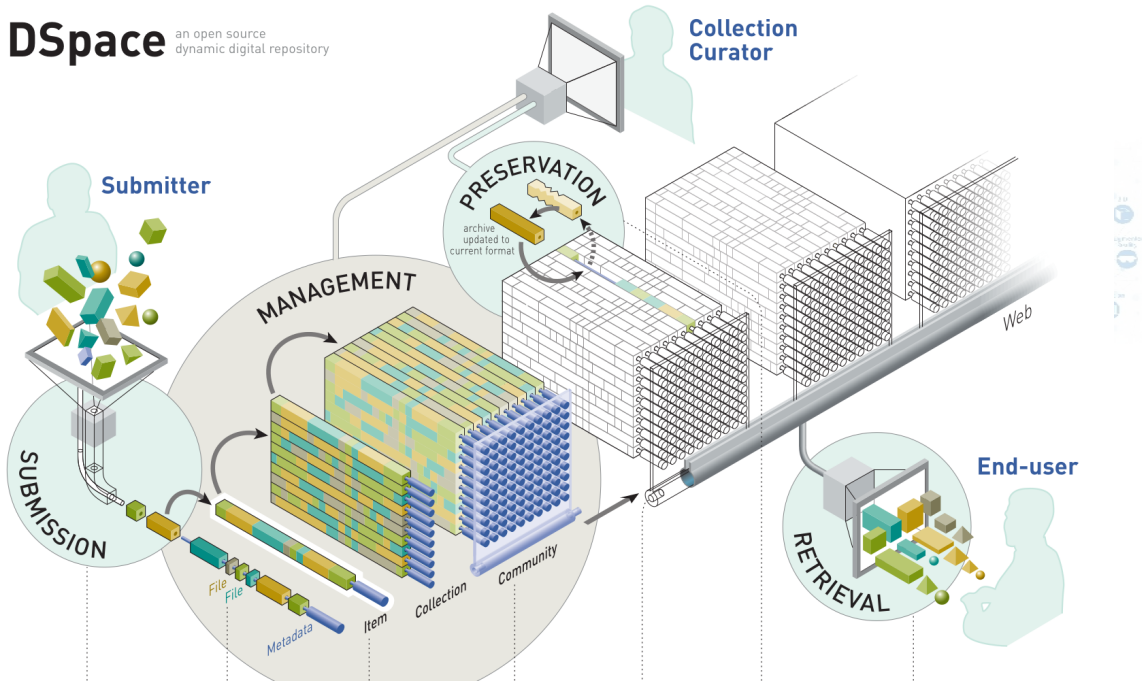
- El repositorio de DSpace puede ser guardado de manera local o en una solución basada en la nube como Amazon S3
- Al recibir un archivo DSpace calcula y guarda una suma de verificación para cada archivo. DSpace usa estas sumas de verificación para validar la integridad de los archivos.



# Funcionamiento

# DSpace

an open source  
dynamic digital repository



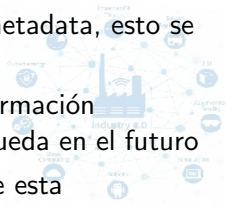
# Propuesta de Información

- El usuario envía su propuesta de información. En esta propuesta se envían los archivos junto con su metadata.
- Esto se realiza a través de la interfaz web del repositorio.
- Esta propuesta es revisada por el curador de la colección, quien puede aceptar o rechazar la propuesta
  - Si es rechazada se informa al usuario a través de un correo la razón del rechazo, el usuario puede hacer las modificaciones necesarias para reiniciar el proceso
  - Si es aceptada, entonces se inicia el siguiente paso en la metodología de trabajo. De no haber otro paso la información es guardada en la colección
- La información ya almacenada en el repositorio, puede ser buscada y recuperada de manera sencilla a partir de la interfaz web del repositorio. Puede buscarse por su contenido o por su metadata



# Estructura

- La información del usuario puede componerse de 1 o mas archivos y su metadata, esto se une en un bloque unitario llamado *Item*
- En DSpace un item es la unidad atómica del repositorio. Consiste de información relacionada por su metadata, la cual es usada para indexada para su búsqueda en el futuro
- Un conjunto de *Items* es una colección. Esta es formada por material que esta relacionado de manera lógica.



## Estructura

-





- Página DSpace: <https://duraspace.org/dspace/>
- Especificaciones DSpace:  
<https://duraspace.org/dspace/resources/technical-specifications/>
- Diagrama DSpace: [https://duraspace.org/wp-content/uploads/dspace-files/DSpace\\_Diagram.pdf](https://duraspace.org/wp-content/uploads/dspace-files/DSpace_Diagram.pdf)
- Smith M., Barton M., Bass M., Branschofsky M., McClellan G., Stuve D., Tansley R. & Harford J.. (2003). DSpace An Open Source Dynamic Digital Repository. Septiembre 2021, de D-Lib Magazine Sitio web:  
<http://www.dlib.org/dlib/january03/smith/01smith.html>

