ENCUENTRO LATINOAMERICANO DE INNOVACIÓN EN EDUCACIÓN SUPERIOR

10-11 Octubre 2019 Bogotá D.C.

¿y tu repositorio institucional está certificado?

Parte 4

Nuestra propuesta

Octubre de 2019

Dra. Marisa R. De Giusti

PREBI-SEDICI- UNLP

CESGI-CIC

UNIVERSIDAD NACIONAL DE LA PLATA

CESGI Centro de Servicios en Gestión de Información cesgi.cic.gba.gob.ar

PREBI prebi.unlp.edu.ar SEDICI sedici.unlp.edu.ar

COMISIÓN DE INVESTIGACIONES CIENTÍFICAS

Ministerio de Ciencia, Tecnología e Innovación

Personal

Personal

- Presenta un equipo de trabajo multidisciplinario con conocimientos de catalogación, gestión, informática diseño y comunicación.
- El personal se nuclea dentro del repositorio.
- Se mantienen redes de colaboración y asistencia con otras instituciones.
- Prevalece la adopción de estándares por sobre las decisiones locales.
- Hay formación contínua del personal.

Visibilidad, promoción y apertura

Interoperabilidad

- Servidor OAI-PMH
  - Exposición de recursos en Dublin Core Element Set (DC simple).
  - Exposición de recursos en conformidad con las directrices 2015 del SNRD.
  - Ofrece SETs lo que permite a los agregadores cosechar los recursos que interesan.
  - Mantiene registro de ítems borrados (borrado persistente).
  - Openaire 4: DC +DATACITE
- Exportación a otros formatos más elaborados: RDF, Refworks, JSON, etc..
- Soporta ingesta vía SWORD.
- Permite la recuperación de datos a partir de OpenSearch y/o Feeds RSS.
- Interopera de forma transparente con otros sistemas institucionales.

Funcionamiento global del repositorio

- Asocia identificadores persistentes a todas las obras (handle, DOI u otros).
- El procedimiento de carga varía en función de los permisos del usuario (la experiencia determina perfiles).
- Utiliza uno o más esquemas de metadatos estandarizados.
- El perfil de metadatos incluye metadatos técnicos y de preservación.

- Hay trazabilidad del objeto digital y es clara y hay metadatos!
- Se utilizan vocabularios controlados para la descripción.
- Se dispone de un módulo de embargo para bloquear temporalmente la publicación de recursos.

Políticas y documentación

Cuenta con documentación sobre:

- Los objetivos del repositorio: misión, visión, etcétera.
- El procedimiento de autoarchivo y su finalidad.
- Políticas generales del repositorio: datos, metadatos, preservación, reúso…
- Aspectos legales: licencias, versiones de las obras.

Cuenta con documentación sobre:

- Procedimientos escritos, por ejemplo procedimiento de digitalización
- Plan estratégico de conservación que defina el enfoque a largo plazo.
- Plan de continuidad y contingencia.

Además, se debe validar que cada recurso posee:

- Una licencia de distribución (que incluye la responsabilidad del autor y la declaración de autoría) y una licencia de uso, preferentemente en un metadato específico.

Infraestructura Tecnológica

El software que sostiene al repositorio

- Es un proyecto de código abierto, no propietario.
- Tiene una comunidad que da soporte.
- Está desarrollado usando lenguajes y librerías ampliamente difundidas.
- Recibe actualizaciones regulares.

El frontend público

- Ofrece una performance aceptable para los usuarios.
- Es usable desde diversos dispositivos (ejemplo: móviles).
- Tiene un módulo de búsqueda y exploración de documentos.
- Está optimizado para ser navegado y analizado por motores de búsqueda (SEO).
- Ofrece acceso a estadísticas públicas.

Recomendaciones generales

- Hardware:
  - Memorias con detección y corrección de errores (ECC RAM)
  - Sistema de alimentación ininterrumpida de energía (UPS)
  - Fuente de alimentación redundante
  - Esquemas redundantes de datos (raid 1+)
- Sistema operativo:
  - Filesystem robusto con mecanismos de comprobación y recuperación de errores.
  - Se aplican automáticamente actualizaciones de seguridad del fabricante
  - Regularmente se aplican otras actualizaciones disponibles
  - Se utiliza un Sistema Operativo para Servidores
- No se usa el servidor para otros propósitos que no sean del repositorio.
- Se configuran los recursos asignados a cada servicio en función del uso (memoria / buffers, cantidad de procesos y/o accesos simultáneos)

Seguridad

- Uso de reglas de firewall para controlar accesos en función de su orígen, destino, puertos entrantes y salientes.
- Uso de mecanismos de detección de ataques por fuerza bruta.
- Revisión periódica de usuarios con permisos de acceso al servidor.
- Revisión de ingresos autorizados (logs).
- Uso de conexiones seguras para administración remota (ej. VPN).
- Se obliga al uso de contraseñas con longitud y complejidad “aceptable” para evitar ataques por diccionario de palabras.

Servidor - Monitoreo

- Control recursos mínimos disponibles (Watchdog).
- Se Implementa un mecanismo de control de logs (ej. logrotate) que
  - rote y descarte logs grandes, antiguos y/o innecesarios,
  - rote y comprima logs de accesos.
- Se monitorea automáticamente la disponibilidad de servicios desde el exterior del repositorio: performance, uptime, conectividad, etcétera.

Servidor - Monitoreo

- Se detectan y bloquean IPs con comportamiento abusivo (demasiado frecuente o malintencionado).
- Se realizan sumas de comprobación (checksum) y validación de ficheros.
- Se revisa la correcta ejecución de las tareas programadas.

Resguardo de datos

- Se utiliza versionado del código fuente y configuraciones.
- Se generan backups de:
  - configuración del servidor: asignación de recursos, permisos, tareas programadas, etc.
  - software y sus configuraciones
  - archivos binarios / assetstore
  - logs de acceso
  - bases de datos
  - datos complementarios para funcionamiento de servicios (Solr)
  - registros estadísticos
  - sistemas complementarios (ej. de autoridades)

Resguardo de datos

- Hay definido un esquema de rotación de backups en función de su importancia y frecuencia de cambios.
- En caso de usar un entorno de virtualización, se realizan snapshots frecuentes e imágenes completas periódicamente.