Repositorio Institucional UNISON

"El saber de mis hijos hará mi grandeza"

Política de accesibilidad

Hermosillo, Sonora 2018

Contenido

Introducción

La presente política considera los dos grandes rubros que deberán asegurar la accesibilidad de la información para los usuarios, tanto en el corto como en el largo plazo.

Con el objetivo de asegurar la accesibilidad a la información, es necesario tomar en cuenta dos grandes aspectos, el primero, es asegurar que el contenido se encuentre íntegro y pueda ser localizado, interpretado por las computadoras y presentado de forma correcta en pantalla para que pueda ser comprendido por el lector.

Por otro lado, es necesario asegurar que el sistema cuente con la información que describe a los contenidos debidamente estructurados y estandarizados, de forma que pueda ser consultada por diversos sistemas y que esta pueda ser correctamente explotada, asegurando así la localización de los contenidos por parte de los lectores.

Políticas de accesibilidad

Para cubrir estos dos aspectos, es necesario tomar en cuenta los siguientes rubros.

Consideraciones de preservación de la información

El repositorio digital deberá establecer mecanismos para la preservación a perpetuidad de los recursos de información contenidos. Es necesario que los procesos de trabajo del equipo encargado de operar el repositorio incluyan cuando menos las siguientes actividades.

Respaldo de la información

Para garantizar la preservación de la información, el repositorio digital realizará la tarea de extraer los recursos completos de los sistemas externos y almacenarlos en su propia infraestructura, esto con el objetivo de asegurar tanto el acceso como la preservación de los contenidos en el largo plazo.

Para respaldar la información almacenada en el repositorio digital se llevará a cabo un plan de actividades de respaldo que cubrirá como mínimo los siguientes aspectos:

1) La realización de un respaldo integral semestral del repositorio digital hacia medios físicos que deberán ser almacenados en una ubicación geográfica distinta a la ubicación donde se encuentren los servidores, tomando en cuenta que este respaldo integral debe incluir:

a) Respaldo completo de la aplicación instalada en el ambiente de producción.

b) Respaldo completo de los códigos fuente de la versión de la aplicación instalada en producción, incluyendo documentación técnica.

c) Respaldo completo de la base de datos de producción en el formato que maneje el servidor de bases de datos.

d) Respaldo completo de archivos y documentos.

e) Respaldo de la base de datos en formato de texto o SQL estándar.

2) La integración de un modelo de preservación digital que incluya:

a) La generación de una cadena de bits única para cada RIACTI, de acuerdo con las buenas prácticas internacionales relacionadas con la generación de suma de verificación o checksum.

b) La generación de un reporte automatizado que notifique al administrador de cualquier posible alteración en el contenido o estructura informática de los RIACTIs.

c) La generación de múltiples copias de los RIACTIs en al menos dos ubicaciones adicionales al RI, siendo una de ellas en un sistema en línea externo a la institución. Estas copias de respaldo deberán ir acompañadas de un archivo en formato XML que incluya los metadatos que describen al RIACTI en el RI.

d) La implementación de un sistema automatizado de inventario, que no permita la generación de copias de respaldo cuando haya alteraciones en el checksum de los RIACTIs y que permita, además, restaurar copias sin error para reemplazar copias defectuosas.

Responsable de preservación digital

El coordinador del repositorio digital será el encargado de asignar cada periodo de 24 meses, a un encargado de preservación digital. El cual será responsable de cumplir las especificaciones descritas en esta política.

Consideraciones de interoperabilidad

La interoperabilidad es la capacidad de las tecnologías, los servicios y las organizaciones para trabajar juntos de un modo transparente hacia objetivos comunes o diferentes. En el ámbito técnico de los repositorios digitales, se apoya en los estándares abiertos para la comunicación entre sistemas informáticos, permitiendo el intercambio de datos de forma estructurada y controlada.

En estas políticas la interoperabilidad se considera principalmente en el contexto de la interacción que el repositorio digital mantendrá con sistemas externos que deseen obtener información, así como en la interacción con sistemas que deseen aportar información hacia el repositorio digital.

Información sujeta a interoperabilidad

1. Podrá ser entregada únicamente aquella información relacionada con los RIACTIs, que de acuerdo con la configuración del repositorio digital, pueda ser accedida libremente por el público en general a través de su interface de búsqueda.
2. Solamente se entregarán registros que tengan configurada una licencia de uso que permita su libre acceso y difusión de acuerdo con las especificaciones definidas en el sitio web de la organización mundial Creative Commons (www.creativecommons.org).
3. El repositorio digital únicamente recopilará información desde sistemas externos cuando esta cumpla con las características de los dos puntos anteriores.

4. El repositorio digital no compartirá información que, de acuerdo con las leyes nacionales o internacionales, o por solicitud expresa de su autor, se pueda clasificar como restringida, sensible, personal o privada.

Protocolo a utilizar

El repositorio digital utiliza como estándar de interoperabilidad el protocolo denominado OAI-PMH (Open Archives Initiative Protocolo for Metadata Harvesting).

OAI - Open Archives Initiative

La Open Archives Initiative (OAI) es una iniciativa que desarrolla y promueve normas de interoperabilidad que tienen por objeto facilitar la difusión eficaz de contenidos. Esta iniciativa cuenta con fondos y respaldo de fundaciones, universidades y organizaciones de investigación, y surgió como un esfuerzo para mejorar el acceso a los recursos digitales que las organizaciones educativas y de investigación producen y publican, principalmente en acceso abierto.

Definición de OAI-PMH

Open Archives Initiative - Protocol for Metadata Harvesting es un protocolo para el intercambio de metadatos a través de Internet. Este protocolo define un mecanismo de recolección de registros que contienen los metadatos que describen la información alojada en los repositorios.

Su funcionamiento opera a través de transacciones sobre el protocolo HTTP en las que un sistema recolector de contenidos, por ejemplo, el Repositorio Nacional, pide a un repositorio de contenidos, por ejemplo, el RI, que proporcione los metadatos de los documentos que almacena y en caso de contar con los permisos suficientes, solicita el envío de los documentos completos.

En respuesta, el repositorio de contenidos devuelve un conjunto de registros en formato XML según los criterios determinados, como pudiera ser la fecha de creación de los registros, la tipología de los documentos o las áreas temáticas a las que pertenecen.

Para poder operar correctamente, el protocolo OAI-PMH requiere que los metadatos solicitados a los nodos sean devueltos en formato Dublin Core Cualificado de acuerdo con las especificaciones establecidas en estas políticas.

Por qué utilizar OAI-PMH

El uso del protocolo OAI-PMH permitirá recolectar los metadatos y contenidos digitales para ser integrados en el RI, garantizando así la compatibilidad con los principales sistemas de información científica en el mundo.

Uso de metadatos en OAI-PMH

En el uso general del protocolo OAI-PMH se permite la utilización de distintos esquemas de metadatos, sin embargo, el RI únicamente permitirá el uso de esquemas fundamentados en Dublin Core.

Existen diversos parámetros que deberán configurarse dentro del servicio de OAI-PMH con la intención de que la recolecta de información entre sistemas sea ágil y con el objetivo de cometer los menos errores posibles en la recolección.

Para obtener información del RI

Cuando sistemas externos deseen obtener información del repositorio digital, deberán apegarse a la estructura de metadatos que se entregarán vía OAI, en el contexto denominado OpenAIRE, en el contexto denominado Driver o en el contexto denominado Default.

Para que el repositorio digital obtenga información de sistemas externos

A continuación, se describen algunos parámetros que deberán ser tomados en cuenta e indicados en los servicios de OAI-PMH de los sistemas externos para asegurar una recolecta exitosa.

Solicitud de recolecta de datos

El administrador del sistema externo deberá realizar una solicitud para que el repositorio digital pueda recolectar su información, indicando:

1. Nombre del sistema a recolectar.
2. Tipo de información que contiene.
3. Justificación de la necesidad de integración de la información hacia el repositorio digital.
4. Cantidad de registros que deberán ser recolectados la primera vez.
5. Pronóstico de crecimiento de la colección a uno, tres y cinco años.
6. Colección del repositorio digital a la que se pretende afectar con la inserción de registros.
7. URL pública del sistema externo.
8. URL del servicio OAI del sistema externo.
    a. El servicio de OAI del sistema externo, deberá contar con un contexto previamente configurado, que pueda entregar únicamente los registros indicados en la solicitud y en el que, sus registros cumplan con las especificaciones de metadatos descritas en estas políticas.

Revisión y pruebas

Una vez recibida la solicitud, el equipo administrativo del repositorio digital realizará una revisión previa del cumplimiento de especificaciones del sistema externo. Si esta revisión es aprobada, se realizarán pruebas de recolección (en adelante, cosecha) de la información.

Si las pruebas resultan exitosas, se procederá a la cosecha de información de forma periódica.

El repositorio digital mantendrá el derecho de cancelar la cosecha e incluso eliminar la colección en caso de que el sistema externo incumpla con las especificaciones

descritas en estas políticas o que se encuentre fuera de línea por más de dos semanas sin comunicación por parte de sus administradores.

Sincronización de bases de datos

Para que las tareas que se describen a continuación se realicen de forma correcta, los sistemas deberán reportar la fecha bajo el estándar ISO 8601 que señala el formato AAAA-MM-DD.

Marca de fecha: Este parámetro se muestra en el encabezado de cada registro y contiene la fecha de creación, modificación o eliminación del registro y que a su vez será tomado en cuenta por el RI para permitir la recolección selectiva de acuerdo con las marcas que se describen a continuación.

Modificación: La respuesta debe incluir registros que correspondan al argumento metadataPrefix y que hubieran cambiado en los límites de los argumentos “from” y “until”.

Creación: La respuesta debe incluir registros que correspondan al argumento metadataPrefix y que hayan pasado a estar disponibles dentro de los límites de los argumentos “from” y “until”.

Eliminación: La respuesta debe incluir registros que correspondan al argumento metadataPrefix que se hayan extraído del repositorio local dentro de los límites de los argumentos “from” y “until”. El estado de eliminación del registro se indicará en el header o “encabezado” del registro y no se incluirá ningún metadato.

Reporte de registros eliminados

Si un registro deja de estar disponible para su consulta dentro del repositorio, éste se debe considerar como eliminado. Cualquier repositorio deberá reportar en su plataforma local el nivel de eliminación de un registro en el elemento deletedRecord de la respuesta de Identify, por lo que es importante que se declare alguno de los siguientes 2 niveles que soportan los registros eliminados en el elemento DeletedRecord de la respuesta de Identify:

Uso del nivel transitorio: El sistema no garantiza el mantenimiento permanente de su lista de eliminaciones. Vigencia: 60 días.

Uso del nivel permanente: El sistema mantendrá la lista de eliminaciones indefinidamente.

Uso del testigo de reanudación

Un testigo de reanudación es un valor que envía el sistema que entrega la información hacia el sistema que recibe la información cuando éste último se encuentra haciendo la recolecta de registros y metadatos. Su objetivo principal es permitir al sistema que recibe la información recuperarse de posibles errores de red o de otro tipo a fin de que no sea necesario reanudar la secuencia de solicitudes de recolección desde un inicio.

Vigencia del testigo de reanudación

El tiempo medio de vida de un testigo de reanudación es el tiempo durante el cual el repositorio guarda en memoria el testigo junto con la información de reanudación.

Los repositorios deberán conservar activo el testigo de reanudación durante un mínimo de 48 horas a fin de dar tiempo suficiente para reanudar la recolecta de datos.

Uso del set de metadatos (contexto)

Un set o contexto, es una agrupación de ítems dentro de un repositorio que permite una partición lógica de ítems para una recolección selectiva de metadatos. Los Sets definen grupos de metadatos en un repositorio, y los metadatos se pueden agrupar por cualquier característica que proporcione una partición razonable para una recolección selectiva.

Los sistemas externos deberán agrupar en un set o contexto los recursos textuales que podrán ser recolectados por el repositorio digital mediante el protocolo OAI-PMH.

Este Set o Contexto, deberá ser denominado: UNISON.

No se deberán incluir dentro del set aquellos elementos que no cuenten con textos completos disponibles.

Uso de identificadores persistentes

A continuación, se presenta una descripción y la metodología a utilizar para que se pueda asignar un identificador persistente a cada recurso almacenado en el RI.

¿Qué son las direcciones persistentes?

Los sistemas de direcciones persistentes son herramientas basadas en URN (Nombres de Recursos Uniformes) cuyo objetivo es solucionar los problemas que surgen al cambiar la ubicación o nombre de algún archivo disponible a través de Internet. Su función primordial es direccionar a los documentos, no importando si estos cambiaron de ubicación dentro de un servidor lo que garantiza que los documentos digitales siempre se encontrarán disponibles para su consulta y/o descarga.

A diferencia de una URL, los sistemas de direcciones persistentes permiten manipular recursos digitales especificando su nombre en lugar de la dirección electrónica en dónde el recurso se encuentra alojado.

¿Por qué utilizar identificadores únicos y persistentes?

Uno de los problemas recurrentes de los usuarios al consultar información digital es encontrarse con enlaces rotos que les impiden la consulta o descarga de algún documento digital. Estos problemas se presentan por que el documento digital se ha movido de servidor o bien por que el servidor en donde está depositado no se encuentra disponible. Al utilizar un sistema de direcciones persistentes dentro de los repositorios locales, un recurso digital podrá ser movido en un futuro sin necesidad de modificar la dirección persistente que lo referencia.

Por lo anterior es importante que el RI mantenga direcciones persistentes para cada uno de los documentos que aloja a fin de asegurar a los usuarios de la información el acceso a los documentos digitales, así como su correcta citación.

Estándar IEFT RFC1737

El Grupo de Trabajo de Ingeniería en Internet (IETF Internet Engineering Task Force) es una entidad internacional de normalización cuyo objetivo es contribuir a la arquitectura de Internet desarrollando normas encaminadas al transporte, procesamiento de datos y seguridad. El IEFT es reconocido internacionalmente por ser la organización que regula las propuestas y estándares en Internet conocidos como RFC.

El estándar RFC 1737 dicta los Requerimientos Funcionales para Nombres de Recursos Uniformes respecto a sus capacidades funcionales y la forma en cómo se codifican. Entre estos requerimientos destacan los dos siguientes.

Unicidad: El URN asignado a un recurso digital nunca podrá asignarse a otro recurso; es decir un mismo URN nunca podrá ser asignado a dos recursos digitales distintos.

Persistencia: El periodo de vida de un URN es permanente, es decir se usará como referencia a un recurso aun cuando el mismo hubiera caducado. Por lo que un identificador permanente no podrá volver a ser asignado a ningún otro recurso una vez que ya fue utilizado.

Metodología de asignación de identificadores persistentes

Con el objetivo de garantizar la disponibilidad total de los documentos alojados en el repositorio digital, se hará uso de sistemas formales de identificación persistente de los recursos, ya sea a través del servicio DOI o del servicio HANDLE, según lo defina la coordinación del sistema.

Es muy importante tener en cuenta que el repositorio digital hará uso de identificadores persistentes en todos sus registros.

Asignación del identificador persistente generado por el repositorio digital

Se establece como lineamiento que todos los recursos contenidos en los Sets de OAI que serán entregados para cosecha, deberán utilizar en sus metadatos el identificador persistente asignado por el repositorio digital, es decir que los sistemas externos no

deberán asignar direcciones persistentes, ya que estas no serán utilizadas, teniendo como única excepción a las revistas científicas institucionales que ya cuenten con una versión digital operando con el identificador DOI, en cuyo caso, el RI respetará la dirección existente y direccionará a los usuarios hacia el texto completo directamente en el sitio de la revista.

"El saber de mis hijos hará mi grandeza"