



PONTIFICIA
**UNIVERSIDAD
CATÓLICA**
DEL PERÚ



Conferencia Internacional
BIREDIAL - ISTE
22-23-24-25 octubre 2018

Ciencia abierta, Interoperabilidad,
Preservación digital, Visibilidad
académica, Gobierno abierto

VIII Conferencia Internacional sobre
Bibliotecas y Repositorios Digitales
de América Latina (BIREDIAL-ISTEC)

Evaluación de la calidad de los repositorios

Dra. Marisa R. De Giusti

Mg. Malgorzata Lisowska



UNIVERSIDAD
NACIONAL
DE LA PLATA



CESGI
Centro de Servicios en
Gestión de Información
cesgi@cc.gba.gov.ar



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



CIC COMISIÓN DE
INVESTIGACIONES CIENTÍFICAS
Ministerio de Ciencia, Tecnología e Innovación



Universidad del
Rosario



CRAI
Centro de Recursos para el
Aprendizaje y la Investigación



Esta obra está bajo una [Licencia Creative Commons Atribución-NoComercial-CompartirIgual 4.0 Internacional](#).

Hoja de ruta

- ★ Los repositorios institucionales en el marco de la iniciativa de AA.
- ★ Concepto de la calidad y medición de la calidad en Repositorios Institucionales
- ★ Revisión de propuestas internacionales en la evaluación de repositorios de acceso abierto.
- ★ Organizaciones que brindan certificaciones y sellos de calidad.
- ★ Propuesta de evaluación.
- ★ ¿Posibilidad de acuerdo latinoamericano? CIRI

Los repositorios institucionales en el marco de la iniciativa de AA

Una de las dos vías de materialización de acceso abierto, recomendada por la Declaración de Budapest (2002)

Debe ser interoperable con otros repositorios para participar en redes

Da cuenta de la actividad académica y científica de una institución



Los repositorios institucionales de Acceso Abierto

- ❖ No es suficiente con la implementación de un repositorio institucional en una dada tecnología (Dspace, Eprint etc).
- ❖ Se debe cumplir con ciertos criterios de calidad.

Definición de calidad

- La definición de la Norma ISO 9000 no deja dudas: *“**Calidad:** grado en el que un conjunto de características inherentes a un objeto (producto, servicio, proceso, persona, organización, sistema o recurso) cumple con los requisitos.”*
- El requisito es un problema porque es la expectativa del usuario, del cliente.
 - ¿De qué modo “se mide” esa calidad en Repositorios Institucionales?



Medición de la calidad en Repositorios Institucionales

- Evaluar la calidad de los **repositorios en sí mismos y en relación a los servicios que ofrecen a los usuarios.**
- Analizar cómo **los contenidos y los servicios creados satisfacen las necesidades** de investigadores, principales colaboradores y comunidad.
- Utilizar los indicadores adecuados para medir el **cumplimiento de los objetivos** para los que fueron creados.



Algunas consideraciones

El rendimiento del repositorio, pero ¿entendido en términos de qué?: por ejemplo la gestión del repositorio.

Al principio la evaluación se orientó a lograr aumentar contenidos, después se pensó en otros aspectos: tecnológicos, procesos, personal, económicos, difusión, formación de usuarios.

Al evolucionar los repositorios empiezan a aparecer otros requerimientos como el cumplimiento con la legislación de AA.



Sobre el contenido: consideraciones que hacen a la calidad

- Cantidad de materiales: ¡en acceso abierto y a texto completo!
- Tipos de licencia: cuestiones a reflexionar sobre el tipo de licencia CC más conveniente.
- Datos: ¿qué datos? ¿en qué sistemas?

Requisitos

Entendidos como necesidades o expectativas de grupos distintos:



- Las diversas comunidades de usuarios del repositorio: desde la administración hasta los usuarios finales (clientes).
- Otros sistemas, otras tecnologías con las que hay que asegurar una interoperabilidad que no siempre es sencilla.

Revisión de propuestas internacionales de evaluación



- ❖ **Serrano, Melero y Abadal (2014)** recogen 10 años de estudios de evaluación y destacan algunos indicadores propuestos por distintos autores, para luego presentar su propia propuesta.
 - Westell (2006): políticas, apoyo institucional, modelo de financiación, digitalización retrospectiva, interoperabilidad, promoción, preservación...
 - Kim, Kim (2008): contenido, gestión y políticas, sistema, usos Y usuarios.
 - Cassella (2010): indicadores del rendimiento int. y ext.
 - Swan (2012): contenidos, usuarios, flujos de trabajo y financiación.

- La Encuesta del Primary Research Group (2012) definió un conjunto de criterios: promoción, acceso y estadísticas, economía, cooperación, impacto, catalogación, derechos financiación y contenido.
- Vierkant (2013): Censo sobre repositorios institucionales alemanes (8 indicadores).
- Guía ReBIUN, Fecyt y Recolecta (2010 y 2014): visibilidad, políticas, aspectos legales y metadatos, interoperabilidad, estadísticas y seguridad, autenticidad e integridad de los datos.

Conclusiones del trabajo

La comparación de los indicadores propuestos en la bibliografía revisada han sido clasificados en 6 categorías:

- Marketing
- Económicos
- Interoperabilidad y Cooperación,
- Políticas
- Contenidos (sólo aparece la propuesta de Swan)
- Servicios de valor añadido

Propuesta de Serrano, Melero y Abadal

Consideran dos aspectos:

- **Factores técnicos o internos del repositorio y Factores institucionales.**

Establecen **32** indicadores separados en 5 categorías: tecnologías, procedimientos, contenidos, marketing y personal.

El trabajo cuenta con una segunda parte de recolección de datos en función de encuestas (a gestores de repositorios e investigadores) y revisión de contenidos, la totalidad del trabajo se consolida con la Tesis Doctoral de Rocío Serrano Vicente de este año 2017 titulada “Evaluación de los repositorios institucionales de Acceso Abierto en España”.

Organizaciones que brindan certificaciones, sellos y normas de calidad



Repository Audit and Certification DSA–WDS Partnership WG

Research Data Alliance (RDA) es una organización comunitaria de investigación iniciada en 2013 por la Comisión Europea, la United States National Science Foundation and National Institute of Standards and Technology, y el Department of Innovation del gobierno de Australia con el objetivo de construir la infraestructura técnica y social para permitir la distribución de datos en abierto.

The screenshot shows the RDA website interface. At the top, there is a green navigation bar with the tagline "Building the social and technical bridges to enable open data sharing" and links for "RDA EU", "RDA US", "CONTACT US", "LOGIN", and "REGISTRATION". Social media icons for RSS, YouTube, LinkedIn, and Twitter are also present. Below the navigation bar, the RDA logo is displayed on the left. The main content area features two columns: "MEMBERSHIP" with 7388 members and "WORKING GROUPS" with 94 groups. A secondary navigation bar includes links for "ABOUT RDA", "GET INVOLVED", "GROUPS", "RECOMMENDATIONS & OUTPUTS", "RDA FOR DISCIPLINES", "PLENARIES & EVENTS", and "NEWS & MEDIA". The main heading is "Repository Audit and Certification DSA–WDS Partnership WG", with a breadcrumb trail: "Home » Working And Interest Groups » Historical Group » Repository Audit And Certification DSA–WDS Partnership WG". The group details section includes a "HG" icon, the group name, status "Completed", chair(s) "Lesley Rickards, Mary Vardigan, Rorie Edmunds", secretariat liaison "Contact Enquiries email", and a case statement download link. A note states that the group has finished its term and delivered its outputs. For those interested in certification, it directs them to the parent Interest Group: "RDA/WDS Certification of Digital Repositories IG". A public access notice and a "Join Group" link are also visible.

Building the social and technical bridges to enable open data sharing

RDA EU RDA US CONTACT US LOGIN REGISTRATION

RSS YouTube LinkedIn Twitter

RDA
RESEARCH DATA ALLIANCE

MEMBERSHIP Members: 7388

Becoming a member of RDA is simple and open to both individuals and organizations
[Register now](#)

WORKING GROUPS Groups: 94

Discover what RDA Working and Interest Groups and all other Groups are up to and find out how to join them. [Explore Groups](#)

ABOUT RDA GET INVOLVED GROUPS RECOMMENDATIONS & OUTPUTS RDA FOR DISCIPLINES PLENARIES & EVENTS NEWS & MEDIA

Repository Audit and Certification DSA–WDS Partnership WG

Home » Working And Interest Groups » Historical Group » Repository Audit And Certification DSA–WDS Partnership WG

HG **Group details**

Status: Completed
Chair (s): Lesley Rickards, Mary Vardigan, Rorie Edmunds
Secretariat Liaison: Contact Enquiries email
Case Statement: Download

Please note that the Repository Audit and Certification DSA–WDS Partnership WG has now finished its term under RDA and has delivered its **outputs**.

For those with an interest in the topic of certification, please visit (and join) our parent Interest Group: **RDA/WDS Certification of Digital Repositories IG**

To ensure the quality and usability of shared data, the long-term preservation of these data in

Repository Audit and Certification DSA–WDS Partnership WG

Status: Completed

Secretariat Liaison: Contact Enquiries email

Public - accessible to all site users

[Join Group](#)

Repository Audit and Certification DSA–WDS Partnership WG

Este grupo de la Alianza está dedicado a crear procedimientos armonizados para la certificación de repositorios en un nivel básico. Establecieron 16 requisitos organizados en 5 criterios:

- Los datos se deben encontrar en Internet
- Los datos deben ser accesibles (llevar licencia y tener derechos explícitos).
- Los datos deben estar en un formato utilizable
- Los datos deben ser confiables.
- Los datos deben tener un identificador único y permanente.

Autoevaluación y revisión por pares para otorgar el sello.





DEUTSCHE INITIATIVE
FÜR NETZWERKINFORMATION E.V.

Certificado DINI

DINI (2006) se funda en las universidades alemanas con el fin de mejorar los servicios de información. Crean un certificado que es aplicable sobre repositorios de AA con 8 dimensiones y dentro de cada una aspectos obligatorios (M) y otros recomendados (R):

- Visibilidad del servicio
- Políticas,
- Soporte para autores y editores
- Aspectos legales
- Seguridad de la información
- Indexación e interfases
- Acceso a estadísticas
- Disponibilidad a largo plazo



¿Cómo obtenerlo?

- ❑ Se paga un pequeño canon.
- ❑ El gestor del repositorio solicita la certificación DINI tras completar un formulario en línea: <https://dini.de/dini-zertifikat/fragebogen/>
- ❑ Una vez completado el formulario y enviada, la solicitud, los datos serán verificados - generalmente se nombran dos revisores.
- ❑ Debe darse acceso a los servicios a ser certificados a los revisores.
- ❑ Los responsables del servicio deben responder a las preguntas de los revisores.
- ❑ Para obtener el certificado no basta con cumplir los criterios obligatorios, hay que cumplir con los recomendados.

ISO 16363

El propósito principal de esta norma es definir una Práctica Recomendada en la que basar un proceso de auditoría y certificación para evaluar la confianza de los repositorios digitales.

El desarrollo del Modelo de Referencia de Sistema Abierto de Información de Archivo (OAIS) consensuó lo que se requiere para que un repositorio digital provea conservación a largo plazo.

Las instituciones comenzaron a declararse a sí mismas “Conforme a OAIS” para marcar lo confiable de sus repositorios. ¿Cómo justificar esto?

Se hizo vital establecer criterios capaces de dar cuenta si un repositorio digital es confiable o no. En 2002 RLG y OCLC establecieron los criterios TRAC.

ISO 16363

El documento sigue los criterios TRAC para generarse como norma. Las secciones 1 y 2 son informativas y las secciones 3-5 proporcionan las métricas:

Las que abarcan la infraestructura organizacional con 25 requisitos.

- Viabilidad de la institución, estructura organizativa y personal, sostenibilidad económica, contratos y licencias, etc.

Las que abarcan la Gestión de objetos digitales con 42 requisitos: sigue la 14721.

Las que abarcan la Gestión de Riesgos de Seguridad y de Estructura/Infraestructura y gestión de seguridad con 23 requisitos:

- Riesgos de las infraestructuras físicas, gestión de seguridad, asignación de roles.

Corolario: distintos aspectos, pero mucho en común

1. Contenidos
2. Catalogación y metadatos
3. Interoperabilidad
4. Visibilidad
5. Preservación
6. Servicios
7. Estadísticas
8. Seguridad y actualizaciones
9. Comunicación y marketing

Nuestra propuesta



Consideraciones generales

- Es necesario proponer algunas alternativas para evaluar su calidad de los repositorios de la región.
- Pensar en métodos de auditoría con los cuales se compruebe si el repositorio cumple con un conjunto de indicadores.
- Observar el grado de cumplimiento con esos indicadores para mejorar a futuro lo que sea necesario.
- Lo precedente introduce el concepto de “alcance de la auditoría” y puede permitir etapas para el logro de los objetivos institucionales planteados.

Consideraciones generales

Depende de la institución y sus objetivos, ¿qué se evalúa?

- ❑ Servicios, valor añadido.
- ❑ Objetivos cumplidos, grado de colaboración con otros repositorios, estado de desarrollo, personal...

Premisas:

- ❑ Es distinto evaluar un repositorio en sus inicios que cuando ya está desarrollado.
- ❑ No hay una única forma de mirar la calidad de un repositorio.



Alternativas de los indicadores

- Enfocados en aspectos tecnológicos.
- Enfocados en la promoción.
- Enfocados en los servicios.



Nuestra propuesta técnica



Aspectos a auditar

- Personal
- Visibilidad, promoción y apertura
- Interoperabilidad
- Funcionamiento global del repositorio
- Políticas y documentación
- Infraestructura Tecnológica



Personal

Aspectos a auditar

- Cuenta con un equipo de trabajo multidisciplinario con conocimientos de catalogación, gestión, informática y comunicación.
- El personal se nuclea dentro del repositorio, es decir, su lugar de trabajo es en el repositorio.
- Se mantienen redes de colaboración y asistencia con otras instituciones.
- Prevalece la adopción de estándares por sobre las decisiones locales.
- Hay formación continua del personal.



Visibilidad, promoción y apertura

Aspectos a auditar

- Existe un enlace al repositorio desde el portal principal de la institución.
- El repositorio está inscripto en los principales directorios de repositorios, p.e. OpenDoar
- El repositorio es cosechado/indexado por recolectores nacionales e internacionales: ej SNRD, BASE, Google Scholar, etcétera.
- La URL del repositorio (dominio + ruta) contiene el nombre propio del repositorio pero no la tecnología con la que está implementado. Ej: /xmlui, /jspui , /dspace, etc.
- Posee perfiles activos en las redes sociales: Facebook, Twitter, entre otros.
- Cuenta con un blog institucional y/o del personal, sobre el que se realizan actualizaciones periódicas.
- Realiza promociones de diverso tipo y dirigidas a distintos públicos: cursos, seminarios, conferencias, talleres, folletos.
- Se incentiva el autoarchivo.
- Se expone el porcentaje de recursos abiertos de acuerdo a su tipología documental.



Interoperabilidad

Aspectos a auditar

- Servidor OAI-PMH
 - Exposición de recursos en Dublin Core Element Set (DC simple).
 - Exposición de recursos en conformidad con Directrices internacionales como las de [OpenAire](#) o las existentes a nivel país (p.e. Argentina: [directrices 2015 del SNRD](#)).
 - Ofrece SETs lo que permite a los agregadores cosechar los recursos que interesan.
 - Mantiene registro de ítems borrados (borrado persistente).
- 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 internos de la universidad: portales, portales de revistas, de congresos.



Funcionamiento global del repositorio

Aspectos a auditar

- Utiliza uno o más esquemas de metadatos estandarizados.
- El perfil de metadatos incluye metadatos técnicos y de preservación.
- Se utilizan vocabularios controlados para la descripción.
- Se dispone de un módulo de embargo para bloquear temporalmente la publicación de recursos.
- Asocia identificadores persistentes a todas las obras (handle, DOI u otros).
- El procedimiento de carga varía en función de los permisos del usuario.



Políticas y documentación

Aspectos a auditar

- 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.
- 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

Aspectos a auditar

- 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
- 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.



IS.

Aspectos a auditar

- 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.
 - Actualizaciones automáticas 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 origen, 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).
- Uso obligatorio de contraseñas con longitud y complejidad “aceptable” para evitar ataques por diccionario de palabras.

Servidor - Monitoreo

- Control recursos mínimos disponibles (Watchdog).
- Implementación de un mecanismo de control de logs (ej. logrotate) que
 - rote y descarte logs grandes, antiguos y/o innecesarios,
 - rote y comprima logs de accesos.
- Monitoreo automático sobre disponibilidad de servicios desde el exterior del repositorio: testear performance, uptime, conectividad, etcétera.
- Detección y bloqueo de IPs con comportamiento abusivo (demasiado frecuente o malintencionado).
- Sumas de comprobación (checksum) y validación de ficheros.
- Chequeo de la correcta ejecución de tareas programadas.

Resguardo de datos

- Utilizar versionado del código fuente y configuraciones.
- Generar 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)
- Definir un esquema de rotación de backups en función de su importancia y frecuencia de cambios.
- En caso de usar un entorno de virtualización, realizar snapshots frecuentes e imágenes completas periódicamente.

To be continued 2019



OCIRI | Organismo internacional de Certificación de Repositorios Institucionales



OCIRI

Busca evaluar los repositorios de la región a través de auditoría de calidad.





Por consultas: marisa.degiusti@sedici.unlp.edu.ar
margarita.lisowska@urosario.edu.co

La presentación está disponible en la colección: <http://sedici.unlp.edu.ar/handle/10915/25295>



Esta obra está bajo una [Licencia Creative Commons Atribución-NoComercial-CompartirIgual 4.0 Internacional](https://creativecommons.org/licenses/by-nc-sa/4.0/).

<http://sedici.unlp.edu.ar>

<http://digital.cic.gba.gob.ar/>

<http://cesgi.cic.gba.gob.ar/>

<http://prebi.unlp.edu.ar>

<http://www.istec.org/liblink/>

<http://revistas.unlp.edu.ar/cientificas/>

<http://revistas.unlp.edu.ar>

<http://congresos.unlp.edu.ar>

<http://ibros.unlp.edu.ar>