



**Universidade do Minho**  
Serviços de Documentação

# Calidad, interoperabilidad, evaluación y certificación de repositorios – el presente y el futuro

Eloy Rodrigues

[eloy@sdum.uminho.pt](mailto:eloy@sdum.uminho.pt)





# Agenda

- *Introducción*
- *Interoperabilidade y directrices para repositorios*
- *Evaluación y certificación de los repositorios*
- *Los retos de la nueva generación de repositorios*





# INTRODUÇÃO

## CALIDAD DE LOS REPOSITARIOS DE QUE HABLAMOS?





# Que es la Calidad?

- Conjunto de propiedades inherentes a un objeto que le confieren capacidad para satisfacer necesidades implícitas o explícitas.
- La calidad de un producto o servicio es la percepción que el cliente tiene del mismo y la capacidad del mismo para satisfacer sus necesidades
- Walter A. Shewhart: “La calidad como resultado de la interacción de dos dimensiones: dimensión subjetiva (lo que el cliente quiere) y dimensión objetiva (lo que se ofrece)”.





# Calidad

- Según el modelo de la norma ISO 9000, la calidad es el

“grado en el que un conjunto de características inherentes cumple con los requisitos”, entendiéndose por requisito “necesidad o expectativa establecida, generalmente implícita u obligatoria”





# Calidad

- ISO 9001:2015
- “La calidad de los productos y servicios de una organización está determinada por la capacidad para satisfacer a los clientes, y por el impacto previsto y el no previsto sobre las partes interesadas pertinentes.”
- “La calidad de los productos y servicios incluye no sólo su función y desempeño previstos, sino también su valor percibido y el beneficio para el cliente.”





# Calidad de repositorios

- Capacidad para satisfacer las necesidades de la(s) comunidad(es)/institución(es) que sirven
  - Múltiples “clientes” de los repositorios
- Capacidad de interoperar con otros repositorios y sistemas de información relevantes
  - Múltiples sistemas relevantes





# Calidad de repositorios

- Depende /relaciona con varias dimensiones:
  - Infraestructuras
  - Políticas y organización
  - Normas y procedimientos
  - Recursos humanos y capacitación
  - Respaldo institucional/comunidad, financiación y sostenibilidad





# Calidad de repositorios

- Depende /relaciona con varias dimensiones:
- ... y esta muy relacionada con la interoperabilidad





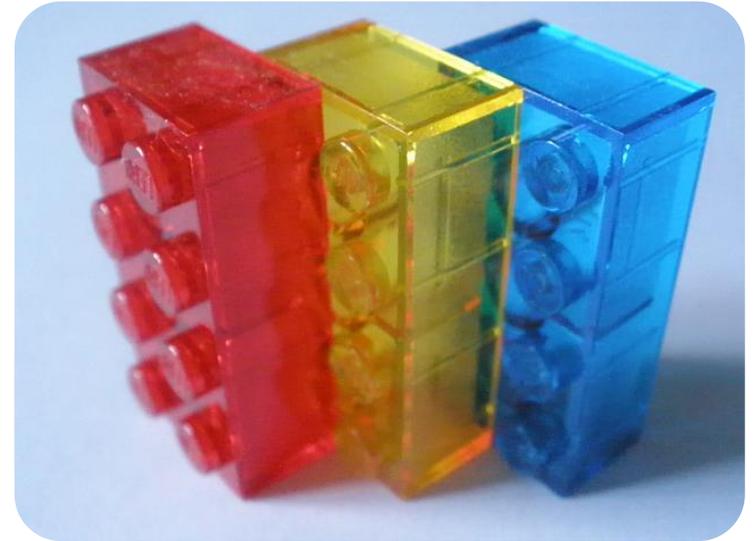
# LA INTEROPERABILIDAD EN LOS REPOSITARIOS





# INTEROPERABILIDAD

Capacidad de trabajar, comunicar y de interactuar entre diversos sistemas.



Capacidad para pasar información, metadatos y objetos digitales de ida y vuelta entre sistemas en un formato utilizable



Nuestros repositorios pueden ser (o contener) tesoros de nuestra institución...



Universidade do Minho  
Serviços de Documentação





Pero no pueden ser islas del tesoro...

Universidade do Minho  
Serviços de Documentação





# El valor de los repositorios

- Cada repositorio individualmente puede ser valioso para su institución o comunidad;
- Pero cada repositorio tiene un valor muy limitado para la ciencia y la investigación...





# El valor de los repositorios

El valor real de los repositorios recae en el potencial de interconexión para crear una red de repositorios, una red que puede proveer un acceso unificado a los resultados de la investigación y son (re-) utilizados por las máquinas y los investigadores.

(El caso de Interoperabilidad para Repositorios de Acceso Abierto - COAR Briefing Paper)





# El valor de los repositorios

Para lograr este potencial de crear un cuerpo unificado de resultados de la investigación necesitamos...

## Contenidos OA

...y el “adhesivo técnico” que hace esta integración posible ...

## Interoperabilidad





Los repositorios OA en todo el mundo, tienen que trabajar con políticas, términos y condiciones, esquemas de datos, protocolos, etc., similares para construir un recurso (transparente) de contenidos basado en una red mundial de repositorios de acceso abierto





# INTEROPERABILIDAD

## PARA QUÉ?





## CONECTAR

- Diferentes repositórios
- Repositórios con outros sistemas de informação
- Objetos y recursos digitais distribuidos





# Interoperabilidade para...

## CREAR NUEVOS SERVICIOS y HERRAMIENTAS

- “Por en cima” de los repositorios, con los contenidos agregados de los repositorios...
- ...utilizando la capacidad computacional que hoy tenemos para procesarlos y conectarlos, generando nueva información y nuevo conocimiento





# Cuestiones y Retos Actuales

## *Retos Técnicos:*

- Nuevos tipos de contenido, nuevos protocolos, identificadores persistentes, vocabularios, idiomas, etc.

## *Retos Organizacionales:*

- Contexto Global/realidades locales, apoyo a la adopción y implementación de las normas, tecnologías y directrices, desarrollo y operación de servicios globales





# ***DIRECTRICES PARA LA INTEROPERABILIDAD***





# ***DIRECTRICES OPENAIRE***





OpenAIRE

# Directrices OpenAIRE

Universidade do Minho  
Serviços de Documentação

## História de una serie de directrices

- DRIVER Guidelines v2 ( 2008 )
  - Normalización del uso de elementos DC y de del protocolo OAI-PMH
  - OAI-Set "driver" -> registros de publicaciones en acceso abierto
  - Introducción de vocabularios (uso del namespace info:eu-repo/semantics)
- OpenAIRE Guildelines v1 ( 2010 )
  - OAI set "ec\_fundedresources"
  - extensión de vocabularios (access rights, embargo period)
  - Esquema de identificación para proyectos (limitado a recursos financiados por EC)
- OpenAIRE Guidelines v2 ( 2012 )
  - Soporte a agregadores OAI
  - Esquema de identificación para proyectos que se puede aplicar a cualquiera financiador, esquema o proyecto



# Data providers: 3 directrices OpenAIRE

**1** Repositorios de literatura

Dublin Core (DRIVER)

**2** Repositorios de datos

Datacite

**3** Sistemas CRIS

CERIF-XML





# Directrices OpenAIRE para repositorios de literatura

- Continúan las directrices DRIVER v2
- Se establecen en OAI-PMH
- Formato de metadatos: oai\_dc
- Application Profile: info:eu-repo
- Con nuevas propiedades para describir:
  - **Funder / Project information**
  - **Embargo Period / Embargo End Date**
  - **Open Access Status / Access Rights**
  - Alternative Identifier
  - Referenced Dataset(s)
  - Referenced Publication(s)



# Directrices OpenAIRE para repositorios de datos

- **Basado en el esquema de metadatos de DataCite**
  - Norma probada para fuentes heterogéneas de datos y archivos interdisciplinarios
  - Mantenido por una organización de confianza y sostenible
  - Soporte para citación de datos
- **Formato de Matadatos: oai\_datacite**
- **OAI-Set: openaire\_data**
  
- **OpenAIRE ha adoptado el DataCite Metadata Schema v3.1 con algunos ajustes menores.**

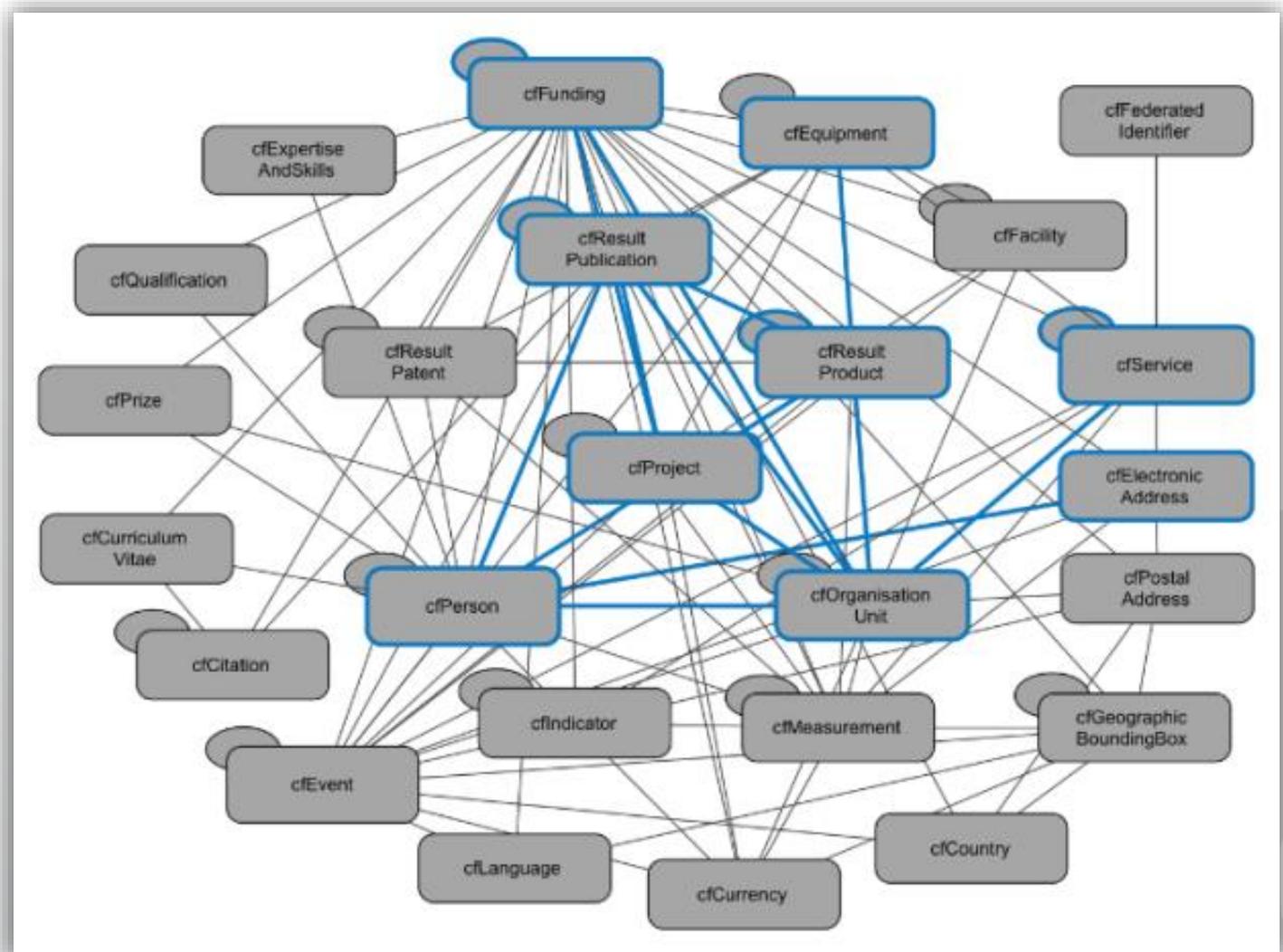
# Directrices para Sistemas CRIS

- **CERIF como estándar de facto para modelo de datos CRIS**
  - Entidades para la organización, persona, proyecto, output, datos..
- **CERIF-XML como estándar de intercambio de datos (En el mundo CRIS)**
- **Mantenido por euroCRIS una organización sostenible**

# Directrices para Sistemas CRIS

- **Adaptaciones para OpenAIRE:**
  - Subconjunto de entidades CERIF relevantes para OpenAIRE
- **Interoperabilidad semántica**
  - Reutilización de vocabularios semánticos CERIF para la clasificación y definición de relaciones entre entidades
  - Coherencia con las otras directrices de OpenAIRE
  - P.ej. tipos de recursos, derechos de acceso

# Entidades CERIF relevantes para OpenAIRE



# Directrices OpenAIRE

## OpenAIRE Guidelines

3.0

OpenAIRE Guidelines for Literature Repositories

OpenAIRE Guidelines for Data Archives

OpenAIRE Guidelines for CRIS Managers based on CERIF-XML

Contributing

How the Horizon 2020 Open Access requirements are met

[Docs](#) » OpenAIRE Guidelines

[View page source](#)

## OpenAIRE Guidelines

Welcome to the OpenAIRE Guidelines. The intention of this is to provide a public space to share OpenAIREs work on interoperability and to engage with the community. The OpenAIRE Guidelines helps repository managers expose publications, datasets and CRIS metadata via the OAI-PMH protocol in order to integrate with OpenAIRE infrastructure.



OpenAIRE Guidelines consists of three guidelines for publication repositories, data archives and CRIS systems respectively:

- [OpenAIRE Guidelines for Literature Repositories](#)
- [OpenAIRE Guidelines for Data Archives](#)
- [OpenAIRE Guidelines for CRIS Managers based on CERIF-XML](#)

The guidelines specifically provide guidance on how to specify:

- Access right
- Funding information

<https://guidelines.openaire.eu>



# Nueva version de las Directrices OpenAIRE Guidelines v4.0

OpenAIRE Guidelines for Literature Repository Managers  
latest

Search docs

Introduction

Application Profile Overview

Read the Docs v: latest

Docs » Application Profile Overview

[Edit on GitHub](#)

## Application Profile Overview

This documentation uses the following namespace abbreviation:

- `dc` : <http://purl.org/dc/elements/1.1/>
- `datacite` : <http://datacite.org/schema/kernel-4>
- `rdf` : <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
- `oaire` : <http://namespaceopenaire.eu/schema/oaire/>

OpenAIRE-Field	Metadata Element	Refinement by vocabulary
1. Title (M)	dc:title	
2. Creator (M)	datacite:creator	
3. Contributor (R)	datacite:contributor	DataCite contributor type
4. FundingReference (MA)	datacite:fundingReference	DataCite funderIdentifier type
5. AlternateIdentifier (R)	datacite:alternateIdentifier	
6. RelatedIdentifier (R)	datacite:relatedIdentifier	DataCite relatedIdentifier type
7. Embargo Period Date (M)	datacite:date	DataCite dateType
8. Language (R)	dc:language	IETF BCP 47. ISO 639-1:2.3

# OpenAIRE guidelines v4.0

## Objetivos

Las Directrices de OpenAIRE versión 4.0 se actualizan exhaustivamente para responder a retos como

- servicios más sofisticados requieren metadatos con más granularidad, precisión y normalización
- soporte de vocabularios controlados donde los conceptos se asignan con URL accionables
- localización adecuada de los archivos de texto completo
- distinción entre derechos de acceso y condiciones de licencia
- descripción de la referencia bibliográfica con más detalle
- Apoyo a la alineación de metadatos con otras infraestructuras de comunicación académica

# OpenAIRE guidelines v4.0

## Estrategía

- actualización gradual mediante la reutilización de estándares de metadatos bien establecidos
- combinando elementos del kernel de metadatos de Dublin Core y DataCite en un perfil de aplicación

# OpenAIRE guidelines v4.0

## Cambios en comparación con las Directrices anteriores (v3)

- provisión de un esquema XML que implementa el perfil de aplicación y permite una validación más fácil
- Sustitución de info: vocabularios eu-repo por otras iniciativas de vocabulario, como COAR Controlled Vocabularies para tipos de recursos y derechos de acceso, términos controlados utilizados en el núcleo de metadatos de DataCite
- apoyo a esquemas de identificación para autores, organizaciones, financiadores, recursos académicos
  - p.ej. conectando con ORCID, ISNI, GRID, Funder Data un esfuerzo de la comunidad internacional para apoyar

# Otras directrices y recomendaciones



- **Metadatos y Políticas de Cosecha de La Referencia (2015)**
- **Directrices SNRD:** Directrices para proveedores de contenido del Sistema Nacional de Repositorios Digitales (2013 y 2015)
- ... y documentos similares en Colombia, Perú, México...





# EVALUACIÓN Y CERTIFICACIÓN DE REPOSITARIOS



# Evaluación y certificación de repositorios



Universidade do Minho  
Serviços de Documentação

- **Porqué?**
  - Para promover a melhoria, la sostenibilidad y la calidad de los repositorios
  - Para aumentar la confianza de los usuarios y otros interesados (stakeholders) en los repositorios
  - Para facilitar la interoperabilidad, la cooperación y la comparabilidad (benchmarking) con otros repositorios



# Evaluación y certificación de repositorios



Universidade do Minho  
Serviços de Documentação

## Como?

- Utilizando diversos instrumentos como guías, recomendaciones, certificados y normas.
  - Guia de Recolecta
  - DINI Certificate
  - Data Seal of Approval y Core Trust Seal
  - ISO 16363
  - ....
- Muy diferentes en complejidad, dimensión y esfuerzo necesario - desde los 16 criterios del Data Seal of Approval, hasta mas de 100 métricas en ISO 16363





# Guía de Re colecta

- *Guía para la evaluación de repositorios institucionales de investigación*
  - primera edición fue publicada por RECOLECTA en Diciembre de 2010
  - Segunda versión publicada por REBIUN/FECYT en 2014





# Guía de Recolecta

- Los repositorios deben cumplir con requisitos mínimos:
  - ***Recopilar, gestionar, difundir y preservar la producción científica digital de la institución.*** Los documentos almacenados en el repositorio deben haber sido creados por la institución o alguno de sus miembros como producto de las funciones de investigación.
  - ***Colección organizada.*** Un repositorio no puede ser un mero depósito de documentos. Estos deberán estar descritos utilizando un número suficiente de metadatos basados en estándares internacionales y mínimamente normalizados.





# Guía de Recolecta

- Los repositorios deben cumplir con requisitos mínimos:
  - **De acceso abierto.** El objetivo de un repositorio debe ser aumentar la visibilidad e impacto de la investigación que se realiza en una institución. Por lo tanto, la situación ideal sería aquella en la que el 100% de los contenidos depositados son de libre acceso.
  - **Interoperable.** Es fundamental la interoperabilidad del repositorio a través de la utilización de protocolos como el OAI-PMH.





# Guía de Recolecta

- Siete secciones y 53 criterios de evaluación
  1. Visibilidad
  2. Políticas
  3. Aspectos legales
  4. Metadatos
  5. Interoperabilidad
  6. Logs y estadísticas
  7. Seguridad, autenticidad e integridad de los datos





# Guía de Re colecta

## **1. Visibilidad**

**1.1 Mención del repositorio en la página principal de la institución**

**1.2 en directorios nacionales e internacionales**

**1.3 Presencia en recolectores nacionales e internacionales**

**1.4 Existencia de un nombre normalizado en todos ellos**

**1.5 Existencia de una URL amigable**

**1.6 Existencia de iniciativas para fomentar la visibilidad del repositorio dentro de la propia institución**

**1.7 Al menos el 75% de los recursos textuales de investigación que ofrece el repositorio se encuentran en acceso abierto**

**1.8 La Institución se ha adherido a alguna de las declaraciones open access (Budapest, Berlín o Bethesda)**





## 2. Políticas

2.1.- Existe una declaración sobre la misión y objetivos del repositorio

2.2.- Documento de acceso público sobre el archivo en el repositorio, donde se establecen al menos los siguientes puntos: quién puede depositar, qué se puede depositar y en qué formatos

2.3.- Documento de acceso público sobre preservación de los contenidos

2.4.- Documento de acceso público sobre reutilización de metadatos

2.5.- Existe una oferta de contacto y asesoramiento visible

2.6.- Política institucional sobre acceso abierto





## **3. Aspectos legales**

**3.1.- El autor reconoce que al depositar no está infringiendo ningún derecho de propiedad intelectual. El repositorio facilitará al autor el cumplimiento con la normativa vigente**

**3.2.- Existe una autorización por el autor o el titular de los derechos que permite la distribución de contenidos**

**3.3.- Existe documentación disponible que ayude al autor a decidir si puede o no archivar el documento**

**3.4.- Inclusión de la información sobre los derechos de autor en los metadatos exportados por el repositorio**





## 4. Metadatos

- 4.1. Se utiliza el formato de metadatos OAI\_DC
- 4.2.- Todos los registros contienen el campo título (dc:title)
- 4.3.- Todos los registros contienen el campo descripción (dc:description)
- 4.4.- Todos los registros contienen el campo tipo de publicación (dc:type)
- 4.5.- El campo tipo de publicación (dc:type) se asigna según los tipos de documentos OpenAIRE
- 4.6.- El campo tipo de publicación (dc:type) se encuentra conforme a vocabulario establecido por OpenAIRE indicando la versión del documento
- 4.7.- Todos los registros contienen un campo de fecha de publicación (dc:date)
- 4.8.- El campo fecha (dc:date) se encuentra conforme al formato establecido
- 4.9.- Todos los registros contienen un campo de derechos de autor (dc:rights)
- 4.10.- Todos los registros contienen información del nivel de acceso (dc:rights)
- 4.11.- Todos los registros contienen el campo autor (dc:creator).
- (...)
- 4.19.- Se utiliza algún formato de metadatos técnicos y/o de conservación





## 5. Interoperabilidade

5.1.- Se proveen los datos a través del protocolo OAI-PMH

5.2.- Se identifican los recursos de investigación a través de uno o varios sets

5.3.- Se identifican las publicaciones financiadas por la Unión Europea o por otros financiadores a través de un set denominado 'openaire'.

5.4.- Se marcan los registros eliminados

5.5.- El tiempo de vida del testigo de reanudación es de un mínimo de veinticuatro horas

5.6.- El correo electrónico del administrador del repositorio está disponible en la etiqueta AdminEmail dentro de la respuesta a una orden Identify

5.7.- Existe una declaración de Description en la respuesta a una orden Identify

5.8.- La entrega de registros a través del protocolo OAI-PMH es progresiva a través de lotes

5.9.- El tamaño de los lotes para la entrega de registros está dentro del rango de 100-500 registros.

5.10.- Coincide el formato de la fecha expresado en la orden Identify con el campo datestamp de los registros.





## **6. Logs y estadísticas**

**6.1.- Los logs del servidor web donde está alojado el repositorio se archivan de forma permanente**

**6.2.- El repositorio proporciona un servicio de estadísticas sobre el uso de los documentos almacenados**

**6.3.- Se realiza un filtrado de accesos de los robots o motores de búsqueda**

**6.4.- Se realiza un filtrado de doble clics**





## **7. Seguridad, autenticidade e integridade de los datos**

**7.1.- Existe un procedimiento establecido sobre la elaboración de copias de seguridad, tanto del software sobre el que funciona el repositorio, los metadatos y los documentos propiamente dichos**

**7.2.- El repositorio utiliza identificadores persistentes para sus contenidos (DOIs, Handles, URNs, etc.)**





# DINI CERTIFICATE

- Desarrollado y atribuido por la German Initiative for Network Information – DINI (Deutsche Initiative für Netzwerkinformation)
- El certificado DINI se estableció en 2004, con ediciones en 2004, 2007, 2010, 2013 y 2016





# DINI CERTIFICATE

## DINI Certificate 2016 „Open Access Repositories and Publishing Services“

- Se aplica a :
  - Repositorios institucionales
  - Repositorios disciplinares/temáticos
  - Revistas de acceso abierto





# DINI CERTIFICATE

- Se atribuye sobretodo a servicios (repositorios y revistas) individuales
- Pero desde DINI Certificate 2013 hay una herramienta adicional para servicios de alojamiento: DINI-ready.
- Los servicios de alojamiento (hosting services) pueden ser certificados de que cumplen ciertos requisitos mínimos del DINI Certificate para todos los servicios que operan en su plataforma.





# DINI CERTIFICATE

- El DINI Certificate tiene 8 criterios y además unas “guidelines for the OAI interface”
- Cada criterio individual está dividido en dos secciones:
  - La primera sección especifica requisitos obligatorios (marcado como M), indispensables para la certificación.
  - La segunda es de recomendaciones (marcadas como R).





# DINI CERTIFICATE

## *Los 8 criterios del DINI Certificate*

- 1 – *Visibility of the Service* (3M y 15R)
- 2 – *Policy* (8M y 2R)
- 3 – *Support of Authors and Publishers* (4M y 5R)
- 4 – *Legal Aspects* (11M y 6R )
- 5 – *Information Security* (10M y 1R)
- 6 – *Indexing and Interfaces* (6M y 7R)
- 7 – *Access Statistics* (4M y 4R)
- 8 – *Long-Term Availability* (3M y 2R)





# DINI CERTIFICATE

- **Como se obtiene?**
  - El gestor del repositorio solicita certificación DINI rellenando un formulario online en el website de DINI
  - Una vez completado el formulario y enviada, la solicitud, los datos serán verificados - generalmente se nombrarán dos revisores.
  - El acceso a los servicios a ser certificados debe ser permitido a los revisores y los responsables del servicio deben estar preparados para responder a las preguntas de los revisores.





# DINI CERTIFICATE

- **Como se obtiene?**
- Se aplica una tasa para el DINI Certificate:
  - Organizaciones sin ánimo de lucro
    - Miembros DINI 50.00 €
    - otros 100.00 €
  - Organizaciones lucrativas
    - Miembros DINI 150.00 €
    - otros 250.00 €





# DINI CERTIFICATE

- A fee is charged after application for the DINI Certificate:
  - 1. Non-profit organizations
    - • DINI members 50.00 €
    - • others 100.00 €
  - 2. Profit organizations
    - • DINI members 150.00 €
    - • others 250.00 €





# DATA SEAL OF APPROVAL

[Home](#)

[Assessment](#)

[Community](#)

[News & Events](#)

[TOOL LOG IN](#)



## *Towards sustainable and trusted data repositories*

There are 16 guidelines that together determine whether your data repository qualifies for the Data Seal of Approval.





# DATA SEAL OF APPROVAL

- Desarrollado por DANS – Data Archiving and Networked Services en 2008
- Las directrices de calidad del Data Seal of Approval son de interés para los productores de datos e instituciones que crean datos digitales, organizaciones que archivan datos y consumidores de datos.
- Los objetivos del DSA son salvaguardar los datos, asegurando una alta calidad y guiando una gestión fiable de los datos para el futuro





# DATA SEAL OF APPROVAL

- Los 16 requisitos de DSA - Core Trustworthy Data Repository Requirements – se basan en los cinco criterios siguientes:
  - Los datos se pueden encontrar en Internet
  - Los datos son accesibles (derechos claros y licencias)
  - Los datos están en un formato utilizable
  - Los datos son fiables
  - Los datos se identifican de una manera única y persistente para que puedan ser referenciados





# DATA SEAL OF APPROVAL

## The Core Trustworthy Data Repository Requirements

1. The repository has an explicit mission to provide access to and preserve data in its domain.
2. The repository maintains all applicable licenses covering data access and use and monitors compliance.
3. The repository has a continuity plan to ensure ongoing access to and preservation of its holdings.





# DATA SEAL OF APPROVAL

## The Core Trustworthy Data Repository Requirements

4. The repository ensures, to the extent possible, that data are created, curated, accessed, and used in compliance with disciplinary and ethical norms.
5. The repository has adequate funding and sufficient numbers of qualified staff managed through a clear system of governance to effectively carry out the mission.
6. The repository adopts mechanism(s) to secure ongoing expert guidance and feedback (either in-house, or external, including scientific guidance, if relevant).





# DATA SEAL OF APPROVAL

## The Core Trustworthy Data Repository Requirements

7. The repository guarantees the integrity and authenticity of the data.
8. The repository accepts data and metadata based on defined criteria to ensure relevance and understandability for data users.
9. The repository applies documented processes and procedures in managing archival storage of the data.
10. The repository assumes responsibility for long-term preservation and manages this function in a planned and documented way.





# DATA SEAL OF APPROVAL

## The Core Trustworthy Data Repository Requirements

11. The repository has appropriate expertise to address technical data and metadata quality and ensures that sufficient information is available for end users to make quality-related evaluations.
12. Archiving takes place according to defined workflows from ingest to dissemination.
13. The repository enables users to discover the data and refer to them in a persistent way through proper citation.





# DATA SEAL OF APPROVAL

## The Core Trustworthy Data Repository Requirements

14. The repository enables reuse of the data over time, ensuring that appropriate metadata are available to support the understanding and use of the data.

15. The repository functions on well-supported operating systems and other core infrastructural software and is using hardware and software technologies appropriate to the services it provides to its Designated Community.

16. The technical infrastructure of the repository provides for protection of the facility and its data, products, services, and users.





# DATA SEAL OF APPROVAL

- Como se obtiene?
  - Completar una autoevaluación en la herramienta online de DSA. La herramienta lleva a través de los 16 requisitos y proporciona apoyo.
  - Una vez completada la autoevaluación, se puede enviar para su revisión por pares. El revisor revisará las respuestas y documentación y otorgará el Sello de Aprobación de Datos (DSA) cuando se haya cumplido con los requisitos.
  - La autoevaluación y revisión no serán públicas hasta que se otorgue el sello de aprobación de datos





# DATA SEAL OF APPROVAL

- Como se obtiene?
  - En la autoevaluación el solicitante debe indicar un nivel de cumplimiento para cada uno de los requisitos:
    - 0 - No aplicable
    - 1 - El repositorio no ha considerado esto todavía
    - 2 - El repositorio tiene un concepto teórico
    - 3 - El repositorio está en la fase de implementación
    - 4 - La directriz se ha aplicado plenamente en el repositorio





# DATA SEAL OF APPROVAL

- Los niveles de cumplimiento constituyen una parte útil del proceso de autoevaluación, pero todos los solicitantes serán evaluados de acuerdo con declaraciones respaldadas por pruebas apropiadas; no con los niveles de cumplimiento autoevaluados.
- Los niveles de cumplimiento 1 y 2 pueden ser válidos para las autoevaluaciones internas, mientras que la certificación puede ser otorgada si se considera que algunas directrices son de nivel 3 -en la fase de implementación- ya que los Requisitos incluyen una asunción de la mejora continua del repositorio.





# CORE TRUST SEAL



## CORETRUSTSEAL CERTIFIED DATA REPOSITORIES

Broad disciplinary and geographic coverage

[Browse Map and List](#)





# CORE TRUST SEAL

- Iniciativa de ICSU World Data System (WDS) y de Data Seal of Approval (DSA) para crear una nueva organización de certificación: CoreTrustSeal.
- CoreTrustSeal estará basado en los requisitos y procedimientos de DSA-WDS Core Trustworthy Data Repositories Requirements. Esto es el resultado de un esfuerzo cooperativo entre DSA and WDS en el marco de Research Data Alliance para fundir sus certificaciones de repositorios de datos y les va a substituir.
- La herramienta online de CoreTrustSeal Data Repository estará disponible antes del final de 2017





# ISO 16363

- **ISO 16363:2012** - Audit and certification of trustworthy digital repositories
  - Define una práctica recomendada para evaluar la confiabilidad de los repositorios digitales.
  - Es aplicable a toda la gama de repositorios digitales.
  - ISO 16363: 2012 puede utilizarse como base para la certificación.





# ISO 16363

## 3 Partes:

- Estructura organizacional (25 requisitos)
  - Viabilidad de la institución, estructura organizativa y personal, sostenibilidad económica, contratos y licencias, etc.
- Gestión de objetos digitales (42 requisitos)
  - Modelo OAIS
- Infraestructura y gestión da seguridad (23 requisitos)
  - Riesgos de las infraestructuras físicas, gestión de seguridad, asignación de responsabilidades



Auditoria ISO 16363 a Repositórios

# UN EJEMPLO DE APLICACIÓN EN PORTUGAL

## Plano geral (duração 1 ano)

Fase	Designação	Descrição	Data
1	Diagnóstico	Auditoria preliminar a todos os repositórios	Nov. 2013 - Fev. 2014
2	Plano de Ações	Identificação das ações para melhorar a conformidade do repositório com a norma	Maio 2014
3	Implementação	Gestores de Repositórios devem implementar as sugestões de melhoria	Maio - Nov. 2014
4	Auditoria Final	Auditoria final para validar o novo nível de conformidade	Nov-Dez 2014

## Auditoria de Diagnóstico

- O responsável avalia o repositório para cada critério do normativo
- O responsável indica evidências para cada critério
- A equipa de auditoria avalia o repositório com base nas evidências
- A equipa de auditoria fornece sugestões de melhoria para determinados critérios

## Níveis de Maturidade (ECM3)

Nível	Nome	Descrição
1	Inexistente	O repositório não implementa quaisquer processos.
2	Incipiente	O repositório está consciente da necessidade de existirem processos para suprir o requisito, porém estes não se encontram devidamente formalizados ou são realizados de forma ad-hoc.
3	Em formação	O repositório possui processos definidos que satisfazem o requisito normativo, porém estes ainda não se encontram totalmente implementados e/ou disseminados.
4	Operacional	Existem políticas, procedimentos e processos implementados que satisfazem as exigências do requisito normativo.
5	Pró-ativo	Existem políticas, procedimentos e processos devidamente enquadrados num sistema de gestão que visa a monitorização e a melhoria contínua tendo por base um plano estratégico assente em factos, i.e. objetivos, metas e indicadores.

Auditoria ISO 16363 a Repositórios

# RESULTADOS DA AUDITORIA PRELIMINAR

Resultados da Auditoria Preliminar (dos Gestores de Repositórios)

# ESTRUTURA ORGANIZACIONAL

## Resultados da Auditoria (Secção 1)

- A maturidade percebida indicada pelo gestor de repositório foi de **2.7**.

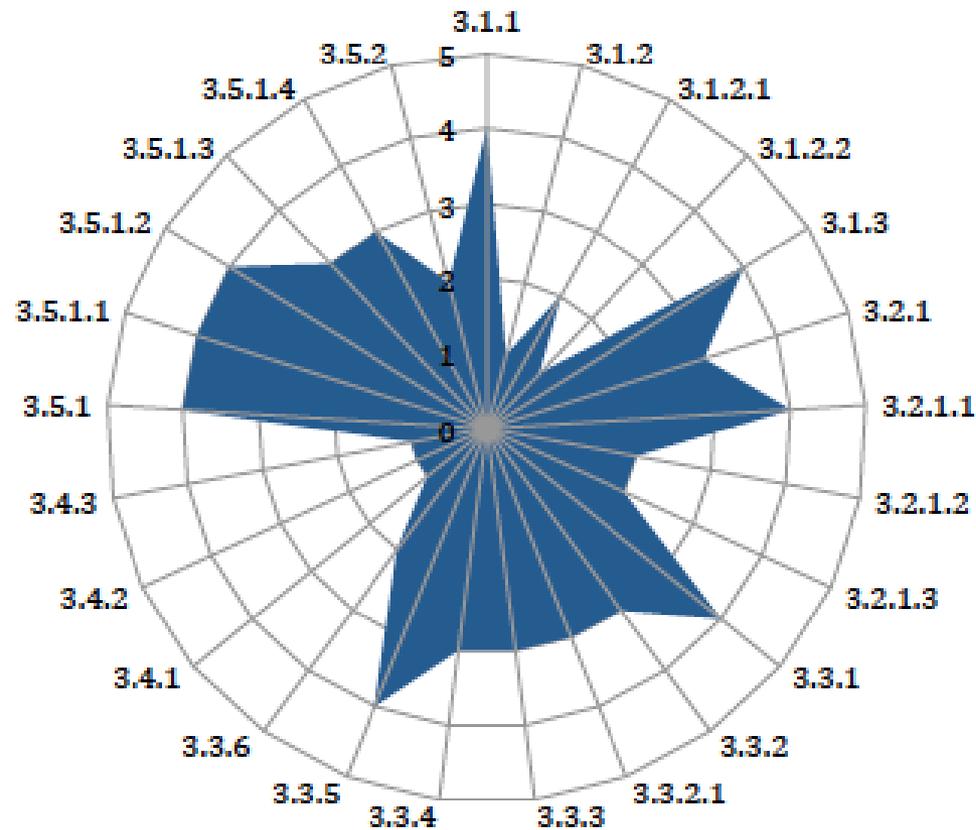
## Resultados da Auditoria (Secção 1)

- O nível médio de maturidade dos 26 repositórios na secção de “Estrutura Organizacional” foi de **2.1** numa escala de 1 a 5.

## Resultados da Auditoria (Secção 1)

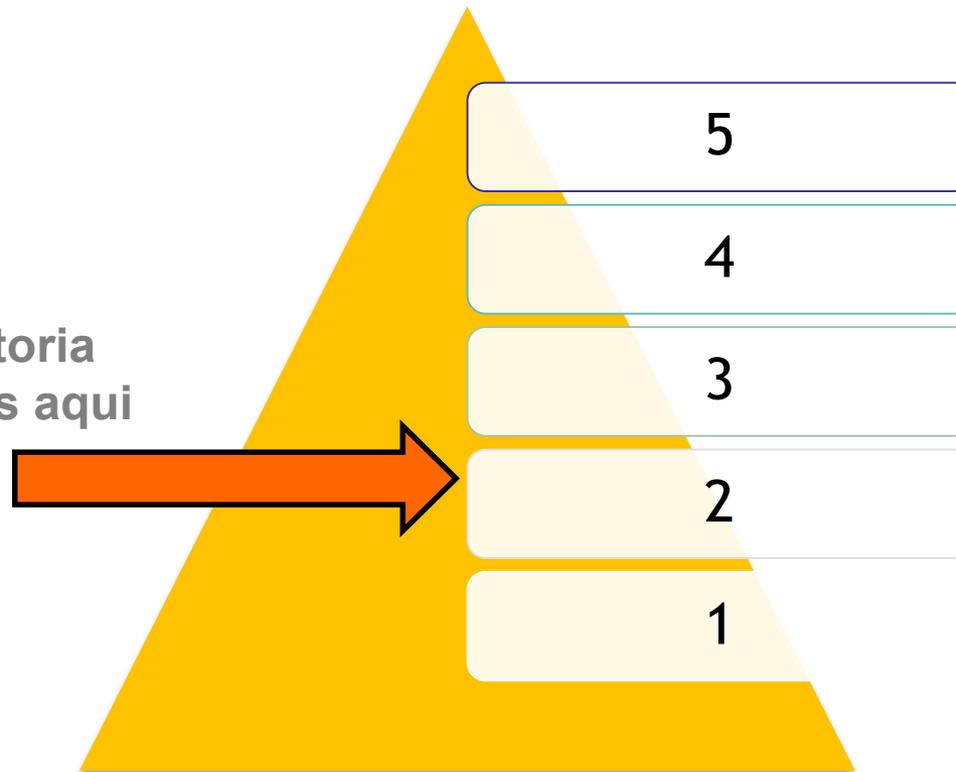
- O melhor repositório obteve um nível de maturidade de **3.0** enquanto que o pior obteve um nível de maturidade de **1.1**.

# Exemplo de Avaliação de Repositório



# Onde estamos?

Com base na auditoria preliminar estamos aqui



## Resultados da Auditoria (Secção 1)

- Dificuldade na obtenção de evidências da prática financeira.
- Se ignorados os 3 requisitos relacionados com a sustentabilidade financeira, o nível médio de maturidade sobe para **2.2**.

## Principais dificuldades

- Inexistência de um plano de sucessão
- Ausência de uma política de preservação digital
- Desenvolver a política de preservação
- Prever auditorias internas
- Garantir os direitos de propriedade intelectual

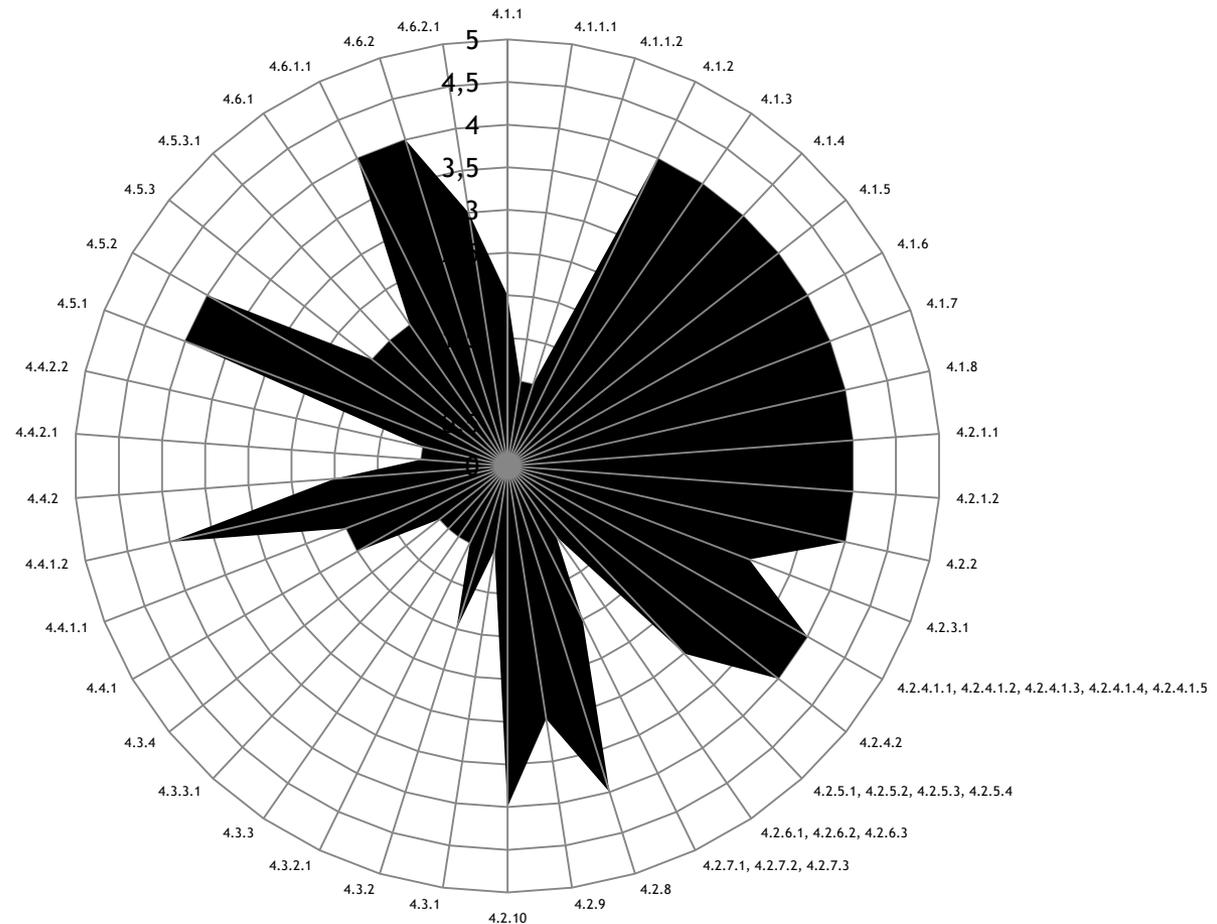
Resultados da Auditoria Preliminar

# GESTÃO DE OBJETOS DIGITAIS

## Resultados da Auditoria (Secção 2)

- O responsável pelo serviço de alojamento indicou um nível de maturidade percebido de **3.2**.
- A equipa de auditoria indicou a sua capacidade de gerir objetos digitais em **2.8**.

# Gestão de Objetos Digitais



## Principais Dificuldades

- Registo de mais informação sobre os conteúdos existentes e a forma como são alterados, transformados ou eliminados.
- Definição de processos documentados
- Falta de ferramentas / processos de Preservação Digital (gestão de AIPs)

Resultados da Auditoria Preliminar

# INFRAESTRUTURA E GESTÃO DA SEGURANÇA

## Resultados da Auditoria (Secção 3)

- Os responsáveis pela “Infraestrutura e Gestão da Segurança” do serviço de alojamento avaliaram a sua maturidade em **3.1**, enquanto a equipa de auditoria definiu a maturidade em **2.6**.



## Principais Dificuldades

- Definir e documentar processos para, por exemplo, atualizar as aplicações/software ou substituição de hardware (discos, memória, etc.)
- Falta de evidências de financiamento para atualização de hardware, software.

Auditoria ISO 16363 a Repositórios

# AUDITORIA FINAL & PRÓXIMOS PASSOS

Auditoria Final

# INFRAESTRUTURA ORGANIZACIONAL

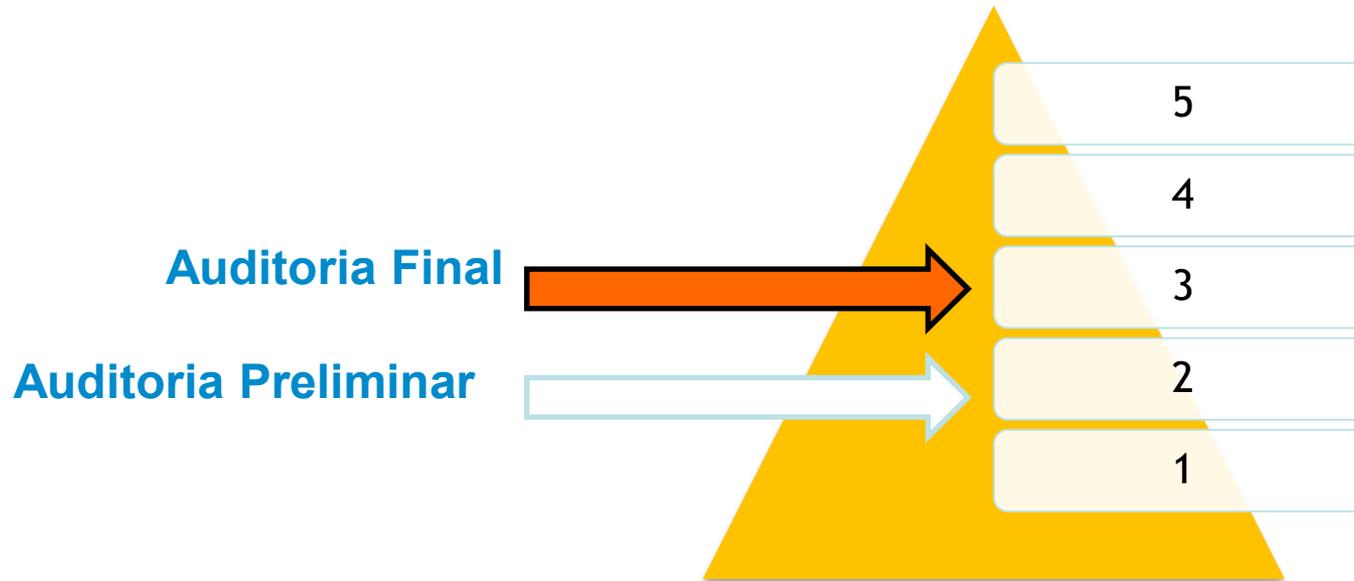
## Resultados da Auditoria Final

- Na auditoria final, a “Infraestrutura Organizacional” obteve um valor médio de **3.0**.
- 6 repositórios não desenvolveram melhorias (se ignorados, a média seria de **3.4**).

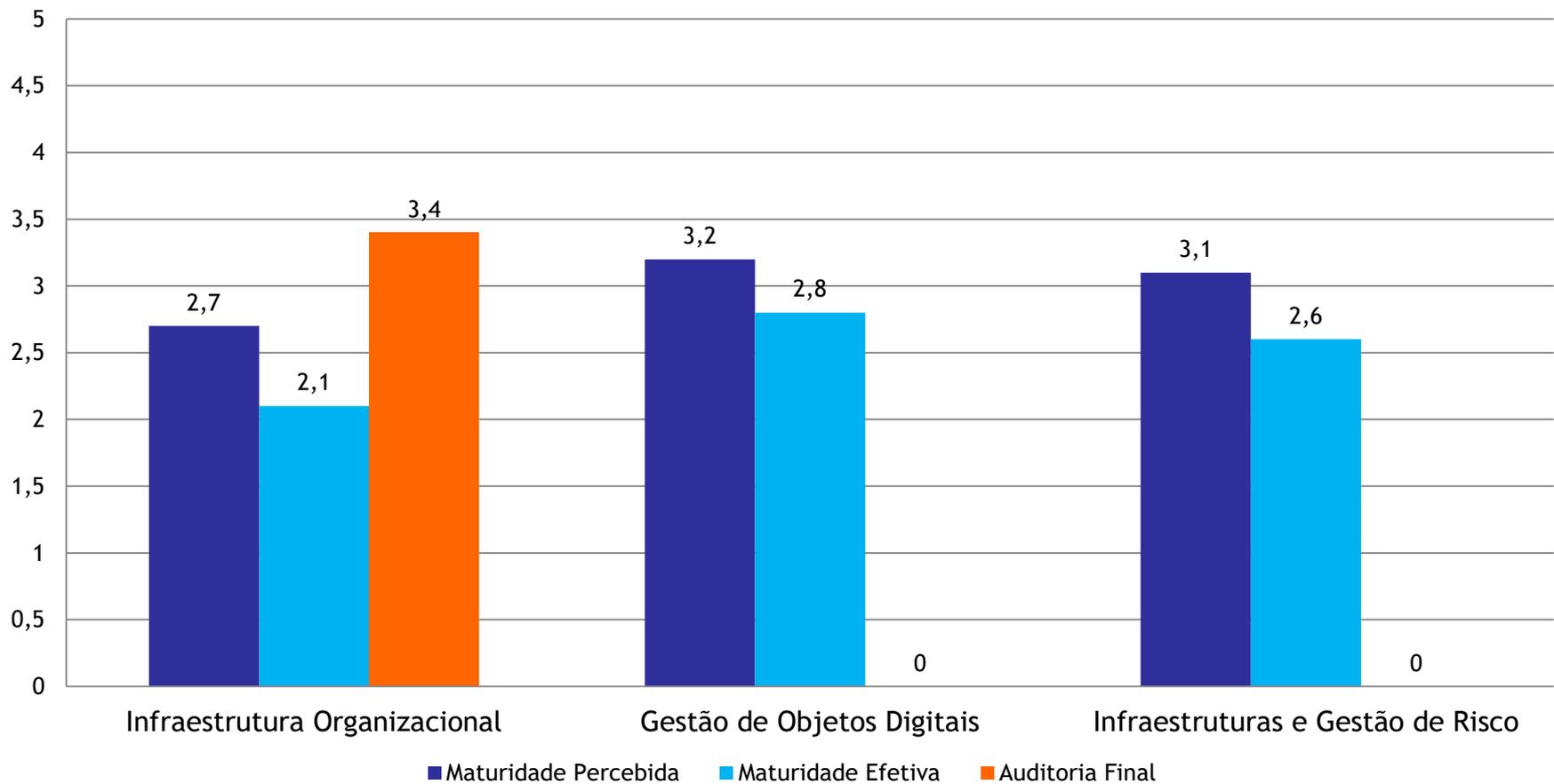
## Resultados da Auditoria Final

- 6 repositórios não desenvolveram as melhorias devido a alterações na organização e falta de recursos humanos,...
- Dos 25 repositórios avaliados na auditoria final, 15 podem ser submetidos a uma certificação.

# Onde estamos? (infraestrutura organizacional)



## Resultados da Auditoria





# Los retos de la nueva generación de repositorios



Trabajando para una

## Nueva Generación de Repositorios



# La Vision de COAR



Un espacio global de conocimiento abierto basado en la red de repositorios de acceso abierto



# Quien és COAR?



- Una asociación internacional establecida en 2009
- Miembros & Asociados: mas de 120 instituciones de 35 países en África, América del Norte y del Sur, Asia, Australasia y Europa,

## Objetivos:

- Voz estratégica para los repositorios
- Interoperabilidad y alineamiento entre las regiones
- *Capacity building*
- Apoyar el desarrollo de servicios de valor añadido



## El problema:

Los repositorios aún no realizarán todo su potencial y funcionan principalmente como recipientes pasivos y aislados de las versiones finales de los resultados de investigación publicados

Los repositorios siguen usando tecnologías y protocolos diseñados 20 años atrás, antes del boom de la Web, de Google, web semántica, redes sociales y dispositivos móviles ubicuos



## La Visión

Posicionar los **repositorios** como la **base** para una **infraestructura** de de comunicación académica y científica **distribuida y globalmente interconectada/en red...**



# Grupo de Trabajo



Eloy Rodrigues, chair (COAR, Portugal)  
Andrea Bollini (CINECA, Italy)  
Alberto Cabezas (LA Referencia, Chile)  
Donatella Castelli (OpenAIRE/CNR, Italy)  
Les Carr (Southampton University, UK)  
Leslie Chan (University of Toronto at Scarborough, Canada)  
Rick Johnson (SHARE/University of Notre Dame, US)  
Paolo Manghi (CNR, Italy)  
Lazarus Matizirofa (NRF, South Africa)  
Pandelis Perakakis (Open Scholar, Spain)  
Jochen Schirrwagen (University of Bielefeld, Germany)  
Daisy Selematsela (NRF, South Africa)  
Kathleen Shearer (COAR, Canada)  
Tim Smith (CERN, Switzerland)  
Herbert Van de Sompel (Los Alamos National Laboratory, US)  
Paul Walk (EDINA, UK)  
David Wilcox (Duraspace/Fedora, Canada)  
Kazu Yamaji (National Institute of Informatics, Japan)



# Nueva Generación de Repositorios



## Tres ideas centrales

Ser de, y no solamente estar en, La Web

Los usuarios son personas y máquinas

Repositorios pro-activos



## Metodología

1. Identificar los principales casos de uso
2. Determinar funcionalidades / comportamientos
3. Desarrollar modelos conceptuales
4. Definir tecnologías y arquitecturas
5. Publicar recomendaciones
6. Apoyar la adopción e implementación



# Next generation repositories



## Resultados iniciales

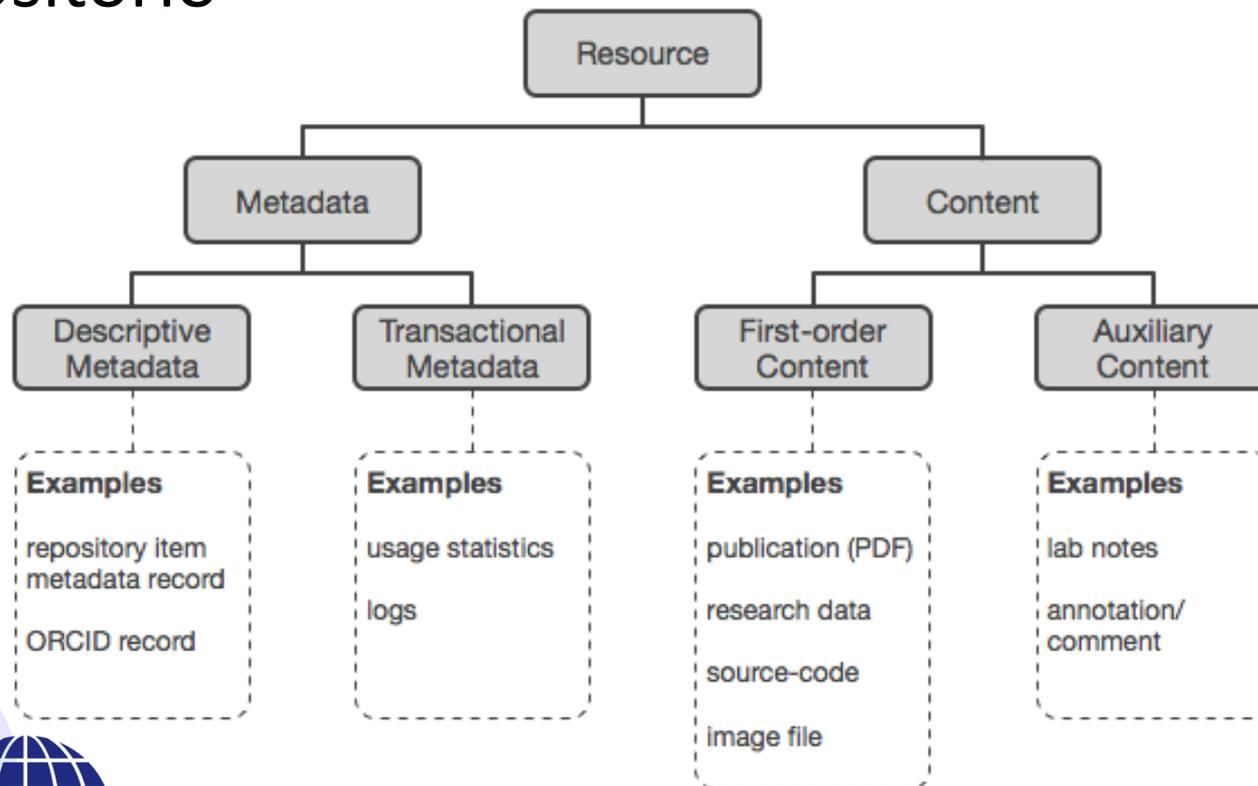
12 historias de usuario disponibles para comentarios del público del 7 de febrero al 3 de marzo de 2017

- Más de 60 comentarios recibidos
- Recomendaciones técnicas que se están desarrollando en base a las historias de usuarios



# ¿Cómo se ve un NGR?

» Centrado en los recursos en lugar de centrado en el repositorio



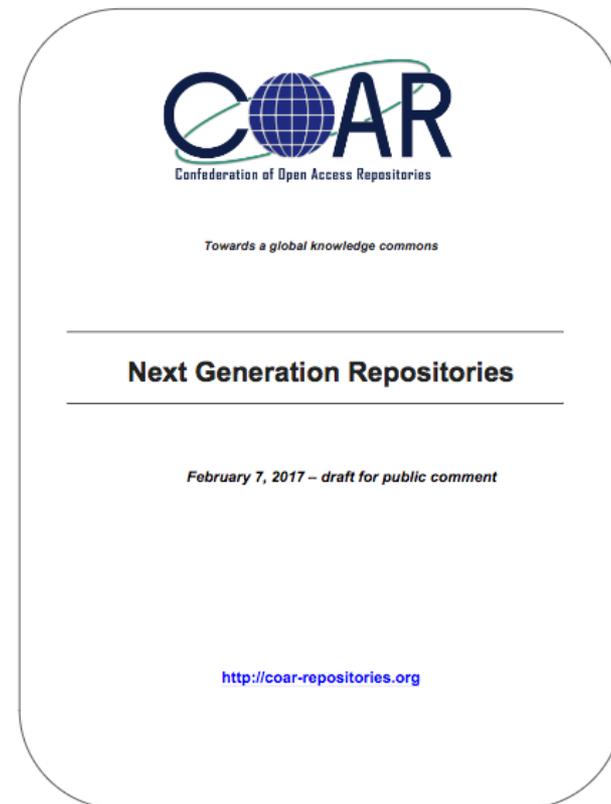
# ¿Cómo se ve un NGR?

- » Recursos disponibles en la Web
- » Abierto y *machine friendly*
- » Colabora con otros servicios
- » No es un silo para la representación del registro, porque el registro puede ser distribuido
- » MetadatoCC0 s por defecto
- » Valor añadido para el repositorio por ser parte de la red
- » Hace uso de datos inferidos de otras fuentes (por ejemplo, identificadores, TDM, datos de interacción), porque los metadatos generados por humanos son caros, requieren mucho de recursos y son propensos a errores..



# Historias de usuario

- Data mining
- Discovering metadata that describe a scholarly resource
- Discovering the identifier of a scholarly resource
- Discovering usage rights
- Resource syncing and notification
- Recognizing the user
- Commenting & annotating
- Providing a social notification feed
- Recommender systems for repositories
- Preservation
- Peer-review
- Comparing usage



<https://www.coar-repositories.org/files/COAR-Next-Generation-Repositories-February-7-2017.pdf>



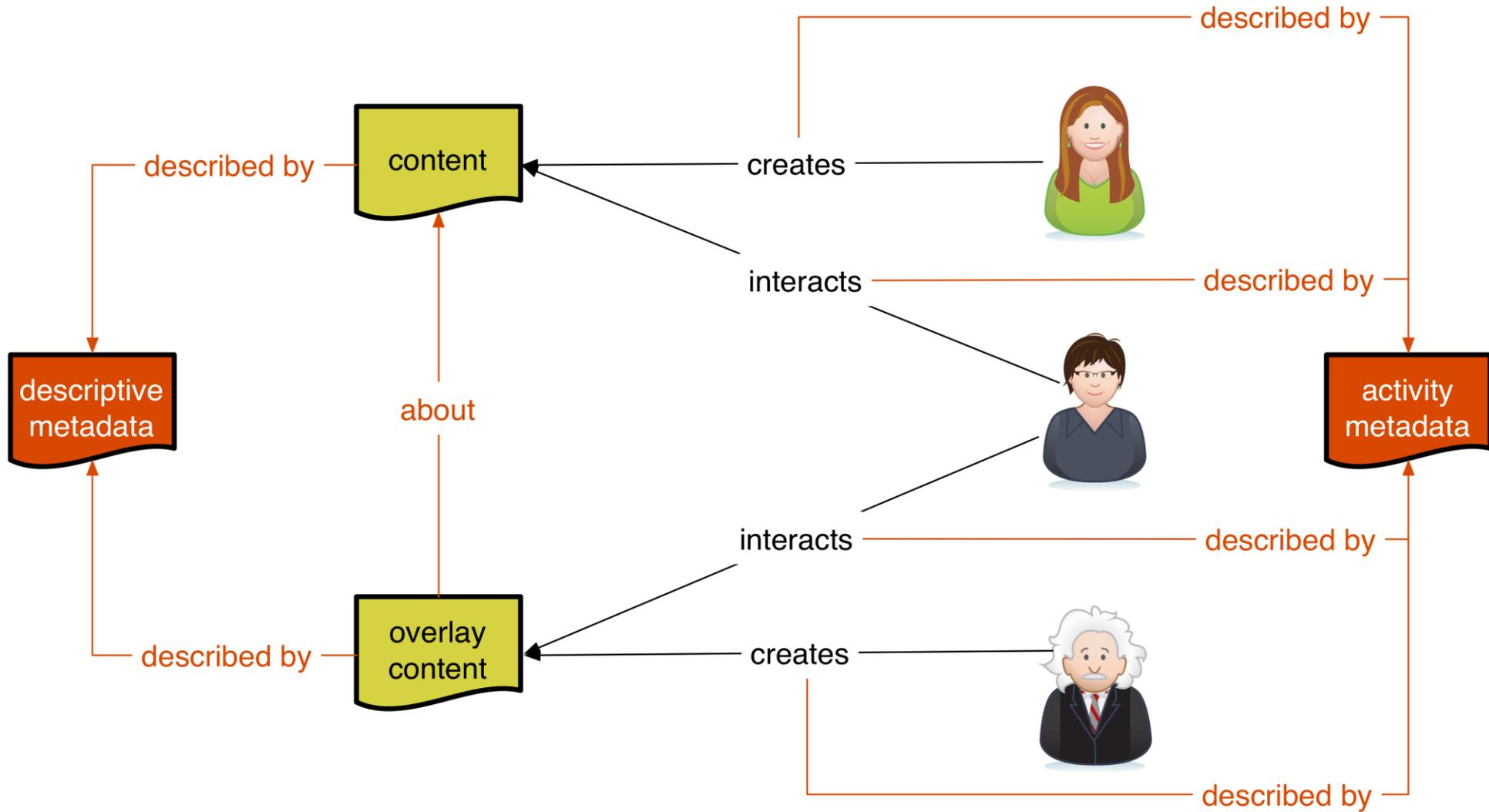
# Más allá del registro de metadatos



- » ¡Contenido! (manuscritos, datos)
- » Enlaces (citas, citas de datos, relaciones, versiones, etc.)
- » Datos de interacción de usuario
- » Comentarios
- » Mensajes
- » Revisiones hechas por colegas
- » Anotaciones



# Datos de actividad



from Herbert van de Sompel

# Tres mecanismos verticales de descubierta

- » Batch - Transferencia de datos en lote
- » Navegación - Ayudar a los robots a encontrar recursos en repositorios mediante navegación
- » Notificación: permite que los robots se suscriban a cambios en los repositorios



- » Guimarães, Portugal - Julio 2017
- » Evaluación de la tecnología:
  1. Qué tecnologías se pueden utilizar para soportar nuevas capacidades de repositorio
  2. Buy, sell, hold decisiones sobre las tecnologías identificadas
  3. Identificación de infraestructura que falta, p. Ej. formatos de datos para la revisión por pares, o camino hacia el resultado poco claro.
- » Objetivo de tener recomendaciones para finales de octubre de 2017



# Ejemplo de recomendaciones

Discovery and exposing resources	Batch	<ul style="list-style-type: none"><li>• Adopt ResourceSync</li><li>• Hold OAI-PMH</li> <li>• Adopt Signposting</li> <li>• Adopt Message Queue</li></ul>
	Navigation	
	Notification	





Synchronization framework for the web consisting of various capabilities that allow third-party systems to remain synchronized with a server's evolving resources.



# Conclusiones

- » Necesitamos que los repositories mejoren la interoperabilidad
- » Necesitamos de repositories mas *machine friendly*
- » Compartiendo una amplia gama de datos (no sólo metadatos de publicaciones)
- » Vale la pena ser parte de la red
- » Adopción de nuevas tecnologías



# Mantengan-se informados

- Visiten el COAR Website: [www.coar-repositories.org](http://www.coar-repositories.org)
- Sigán-nos en Twitter: **#COAR\_eV**
- Sigán-nos en Facebook:  
<https://www.facebook.com/COAReV>

