

Título	Análisis de la calidad de metadatos en el Repositorio Nacional Digital de Ciencia, Tecnología e Innovación de Acceso Abierto – ALICIA
Responsable del documento	Mg. Francisco Talavera Chocano Consultor especialista en control de autoridades y calidad de metadatos
Fecha	2019.04.12
Tema	Gestión y calidad de información
Estado	Concluido
Objetivos	Realizar un análisis de calidad de los metadatos de ALICIA y proponer indicadores y niveles aceptables de calidad para dichos metadatos
Descripción	En este documento se realiza un análisis de la calidad de los metadatos del Repositorio Nacional ALICIA y se proponen criterios e indicadores para medir la calidad de dichos metadatos
Unidad responsable	CONCYTEC / DEGC / SDGIC
Fuentes	<ul style="list-style-type: none"> - Documentos de trabajo del CONCYTEC - OpenAIRE Guidelines for Literature Repository Managers v.4 - Literatura científica sobre calidad de metadatos (ver referencias)
Difusión	<ul style="list-style-type: none"> - Personal técnico de la Dirección de Evaluación y Gestión del Conocimiento del Concytec - Encargados o gestores de repositorios institucionales integrados en ALICIA
Idioma	Español
Documentos relacionados	Directrices de ALICIA
Período de validez	Hasta la próxima versión
Depositado en el repositorio abierto	Repositorio institucional del Concytec
DOI	No definido

Versiones

Número de versión	Fecha	Modificado por	Comentarios
1.0	2019.04.12	FTC	

ANÁLISIS DE LA CALIDAD DE METADATOS EN EL REPOSITORIO NACIONAL DIGITAL DE CIENCIA, TECNOLOGÍA E INNOVACIÓN DE ACCESO ABIERTO - ALICIA

Contenido

I.	Introducción	3
II.	Metodología	4
III.	Repositorios digitales: metadatos, estándares y calidad	5
IV.	Categorías, criterios e indicadores de calidad para los metadatos de ALICIA.	7
V.	Resultados del análisis de los metadatos cosechados en ALICIA.....	13
VI.	Conclusiones y recomendaciones	16
VII.	Referencias.....	18
VIII.	Anexo.....	21

I. Introducción

La Ley Nº 30035, Ley que Regula el Repositorio Nacional Digital de Ciencia, Tecnología e Innovación en Acceso Abierto, y su Reglamento, definen a ALICIA como el “sitio centralizado donde se mantiene información digital resultado de la producción en ciencia, tecnología e innovación (libros, publicaciones, artículos de revistas especializadas, trabajos técnico-científicos, programas informáticos, datos procesados y estadísticas de monitoreo, tesis académicas y similares) (...) de acceso abierto (...) disponible para leer, descargar, reproducir, distribuir, imprimir, buscar o enlazar textos completos...”

El alcance de la Ley y su Reglamento especifican que ALICIA debe contener obligatoriamente el resultado de las investigaciones de las entidades del sector público, tales como universidades, institutos o empresas, entre otros, sean o no miembros del Sistema Nacional de Ciencia, Tecnología e Innovación Tecnológica (SINACYT); las entidades privadas o personas naturales cuya producción intelectual haya sido financiada por el Estado; adicionalmente, podrían incorporarse las entidades del sector privado o personas naturales que deseen compartir voluntariamente su información, siempre y cuando cumplan con las restricciones técnicas y académicas que se especifican en la misma Ley.

Asimismo, establecen que ALICIA es administrado por el Consejo Nacional de Ciencia, Tecnología e Innovación – Concytec, organismo que implementa, integra, estandariza, almacena, preserva y gestiona la información. Concytec, establece políticas para su seguridad y sostenibilidad, dispone estándares de interoperabilidad y promueve el uso y aprovechamiento de dicha información.

Adicionalmente, el Concytec rige la Red Nacional de Repositorios Digitales de Ciencia, Tecnología e Innovación de Acceso Abierto (RENARE) con el fin de propiciar el trabajo colaborativo entre sus miembros para promover al acceso abierto, la utilización y preservación de la información y el conocimiento en materia de ciencia, tecnología e innovación.

Concytec mantiene además un acuerdo de cooperación con la Red Federada de Repositorios Institucionales de Publicaciones Científicas (LA Referencia), de alcance regional, integrada por 9 países: Argentina, Brasil, Chile, Colombia, Ecuador, El Salvador, México, Perú y Venezuela. Su objetivo es facilitar el acceso equitativo y dar visibilidad a la producción científica de centros de educación superior y centros de investigación de los Estados participantes.

De acuerdo a este marco, ALICIA cuenta con las “Directrices para el procesamiento de información en los repositorios institucionales”, que incluyen las indicaciones para el ingreso de datos y los detalles técnicos a considerar para la recolección de metadatos desde los repositorios institucionales que la integran.

Las actuales Directrices de ALICIA están basadas en las “OpenAIRE Guidelines for Literature Repository Managers v. 3.0”, sucesoras de las directrices DRIVER 2.0, el esquema de metadatos Dublin Core y el protocolo de intercambio OAI-PMH. Se establecen un total de 39

metadatos que deben ser considerados para el registro de información en los repositorios institucionales de origen, de los cuales 23 son obligatorios¹, 9 recomendados y 7 opcionales.

Actualmente, ALICIA cosecha los datos de 161 instituciones del país, entre las que se encuentran universidades, escuelas y centros de enseñanza superior, ministerios y otros organismos públicos, institutos de investigación y de salud, ONGs, entre otras. Cuenta con cerca de 215 mil registro de: tesis de grado, tesis de maestría, tesis de doctorado artículos de revistas, reportes, libros, objetos de conferencia, documentos técnicos, entre otros².

Es indudable la importancia que tiene el Repositorio Nacional ALICIA para el desarrollo, acceso y visibilidad de la ciencia, tecnología e innovación tanto para el país como para la región. Por ello, es necesario que los datos cosechados de los repositorios institucionales cumplan con niveles óptimos de calidad que permitan validar la estructura semántica, sintáctica y los valores de los propios datos, para así cumplir con los objetivos que la Ley le asigna.

Con respecto a la calidad de los datos, la propia Ley N° 30035, en su artículo 5: “Lineamientos fundamentales”, acápite 5.2, inciso “c”, menciona que ALICIA debe garantizar la seguridad y calidad de la información. Por otro lado, el Concytec, en su Política Nacional para el Desarrollo de la Ciencia, la Tecnología y la Innovación Tecnológica – CTI, señala como uno de sus objetivos estratégicos (objetivo N° 5): “Generar información de calidad sobre el desempeño de los actores que conforman el SINACYT”, y para su cumplimiento plantea, en el lineamiento 5.1., “Desarrollar y mejorar la calidad de los sistemas de información en CTI considerando la generación, recopilación, gestión, difusión y uso del conocimiento”.

Dicho esto, en el presente documento se realiza un análisis de los datos contenidos en ALICIA para proponer una serie de categorías y criterios necesarios para garantizar su calidad, alineados a sus objetivos y requisitos funcionales. Esta propuesta, dado que el Concytec es el ente rector en la materia, pretende alcanzar la mejora de los metadatos en los repositorios institucionales que son cosechados por ALICIA .

II. Metodología

Para determinar la calidad de los metadatos del Repositorio Nacional ALICIA, se consultaron 15 elementos Dublin Core sin calificar, con una muestra de aproximadamente el 40% de datos cosechados en cada elemento. Adicionalmente, se cotejaron las “Directrices para el procesamiento de información en los repositorios institucionales” del Concytec, así como otros documentos referenciales como: “OpenAIRE Guidelines for Literature Repositories v3” de OpenAIRE, “Metadatos y políticas de cosecha” de LA Referencia y “Dublin Core Metadata Element Set” de la DCMI.

Para determinar los posibles errores en los metadatos se utilizó la aplicación del motor de búsqueda de código abierto Apache Solr; así también una muestra de los metadatos

¹ Algunos metadatos son obligatorios dependiendo del tipo de documento.

² Datos actualizados al 2019.02.13.

cosechados del repositorio de tesis de la Universidad Nacional Mayor de San Marcos (8789 registros).

Las únicas limitaciones encontradas en el análisis fueron: (1) el Solr permite la visualización de los datos por cada elemento sin calificadores; sin embargo, por los valores de algunos de los elementos era previsible determinar a qué elemento calificado pertenece cada valor; (2) a través del Solr, no fue posible visualizar los datos correspondientes a los metadatos para tesis “thesis.degree.*”.

III. Repositorios digitales: metadatos, estándares y calidad

Los repositorios digitales buscan organizar, almacenar, preservar, difundir y visibilizar la información científica y académica en acceso abierto (OA).

El OA es definido por la UNESCO como “el acceso gratuito a la información y al uso sin restricciones de los recursos digitales por parte de todas las personas”. La Iniciativa de Budapest para el Acceso Abierto establece su propósito e importancia: acelerar la investigación, enriquecer la educación, compartir el aprendizaje y la búsqueda continua del conocimiento (Open Society Institute, 2002).

Para la gestión e interoperabilidad de un repositorio digital es necesario cumplir con ciertas normas y estándares. Partimos de la definición que da la norma ISO 23081-1:2006 a los metadatos: “datos que describen el contexto, contenido y estructura de los documentos, así como su gestión a lo largo del tiempo (...) son información estructurada o semiestructurada que posibilita la creación, registro, clasificación, acceso, conservación y disposición de los documentos a lo largo del tiempo y dentro de un mismo dominio o entre dominios diferentes” (p. 276).

Los metadatos requieren de una organización dentro de un esquema o perfil de aplicación; así, el estándar más extendido en los repositorios digitales científicos y académicos es el Dublin Core Metadata Element Set (DCMES), mantenido por la Dublin Core Metadata Initiative (DCMI), que establecen un conjunto de 15 elementos base, su estructura semántica, reglas de contenido y sintaxis para la descripción de recursos de información.

Los repositorios digitales requieren ser interoperables, es decir, que la información contenida en ellos pueda ser accesible, intercambiada y reutilizada; para ello deberán cumplir con normas que permitan este enlace. Para la transmisión de metadatos codificados en Dublin Core, un repositorio digital utiliza el Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH).

Adicionalmente a las normas y estándares, los repositorios digitales requieren garantizar la calidad de sus datos, que a su vez repercute en la calidad de los recolectores nacionales o internacionales como ALICIA o LA Referencia.

En este sentido, Medrano, Figuerola y Berrocal (2012), citando a Guy, Powell y Day, definen a los metadatos de calidad como aquellos que “respaldan los requerimientos funcionales del

sistema que esté diseñado a soportar (...) la calidad está relacionada con la aptitud para el propósito” (p. 110).

Por su parte, Stivilia (2004) menciona que la alta calidad de los metadatos puede llevar a una mayor satisfacción del usuario y un mayor uso de la plataforma; mientras que Palavitsinis, Manouselis y Sanchez-Alonso (2014), anotan que la calidad se refiere a “datos aptos para el uso por parte de los consumidores de datos” (p. 1202).

Son muchos los estudios realizados para medir la calidad de los metadatos enfocados a repositorios digitales. De ellos, podemos identificar, en primer lugar, a Bruce & Hillmann (2004) que establecen 7 dimensiones o categorías para la calidad: completitud, exactitud, procedencia, conformidad con las expectativas, consistencia lógica y coherencia, actualidad y accesibilidad.

En segundo lugar, mencionamos a Azorín Millaruelo et al. (2017), quienes se refieren a los criterios considerados por RECOLECTA³ para la evaluación de repositorios institucionales de investigación. Estos son: visibilidad, políticas, aspectos legales, metadatos descriptivos de la publicación, interoperabilidad de los metadatos descriptivos de la publicación, logs y estadísticas, seguridad, autenticidad e integridad de los datos, y servicios y funcionalidades de valor añadido.

En tercer lugar, autores como Stivilia (2004) y Myrseth, Stang y Dalberg (2011), enumeran 3 categorías para asegurar la calidad de los metadatos: (1) intrínsecas o sintácticas, que miden los valores o la validación de los elementos de información en sí mismos con relación al esquema predefinido y al conjunto de reglas consideradas en él; (2) relacional / contextual o semántica, que considera las relaciones entre la información y aspectos sobre su contexto de uso, es decir, su conformidad con relación al objeto físico del mundo real que representa; y, (3) reputacional o pragmática, que mide la posición de un objeto de información en la estructura cultural o de actividades a la que pertenece, denota la percepción de calidad por parte de los usuarios.

Para efectos de este análisis, de todas las categorías de calidad mencionadas en los párrafos anteriores se han tomado aquellas que se consideran relevantes en concordancia con los requisitos funcionales y el propósito del Repositorio Nacional ALICIA; para proponer criterios e indicadores para su medición. Cabe resaltar que el Repositorio Nacional ALICIA, tiene como principales propósitos: conservar, preservar, compartir y ofrecer acceso abierto al patrimonio intelectual resultado de la producción en materia de CTI del país, difundir y visibilizar los resultados de investigación científica, apoyar las estrategias de OA mediante la interoperabilidad, aumentar la proyección de los investigadores a escala mundial y el impacto de su investigación, y formar parte de redes más amplias de repositorios en OA regionales y mundiales.

³ RECOLECTA: Recolector de Ciencia Abierta: plataforma que agrupa a repositorios científicos nacionales. Es fruto de la colaboración entre La Fundación Española para la Ciencia y la Tecnología (FECYT) y la Red de Bibliotecas Universitarias (REBIUN).

IV. Categorías, criterios e indicadores de calidad para los metadatos de ALICIA.

En este apartado se proponen 9 categorías de calidad que pueden ser consideradas para los datos recopilados por el Repositorio Nacional ALICIA; y por ende a los repositorios institucionales de los cuales se cosecha información. A continuación se enumeran y precisan:

4.1. Completitud e integridad

Este nivel de calidad evalúa si los datos de cada elemento requerido están completos, si contienen la información necesaria para la representación del objeto de información descrito y si éstos están presentes en la instancia del elemento del esquema que le corresponde. Así mismo, que el conjunto de elementos utilizados describan los objetos de información de manera íntegra y económica. Además, identifica aquellos elementos dentro del esquema que no son utilizados por no aplicarse a dichos objetos de información.

4.2. Exactitud y precisión

Los metadatos deben describir los objetos a través de una declaración sin controversia, es decir, que la información proporcionada debe ser correcta y objetiva, sin errores tipográficos, normalizada, utilizando apropiadamente los valores controlados asignados y las especificaciones que el esquema indica para el llenado del elemento (p.e. texto libre o formato de fecha correcto).

4.3. Procedencia y confiabilidad

Es importante considerar la procedencia de los datos para la calidad de la información, los valores deben provenir de fuentes con estándares de contenido establecidos. También se debe conocer las transformaciones que se han aplicado en los metadatos desde su creación y recopilación.

4.4. Conformidad con las expectativas

Este nivel mide el grado en que los perfiles de aplicación y los metadatos cumplen con las expectativas de la comunidad de usuarios que sirve. En otras palabras, no se deben considerar elementos que no sean útiles o que sean superfluos, irrelevantes o difíciles de implementar. Se debe contemplar, por ejemplo, que los vocabularios controlados utilizados sean significativos para los usuarios o que los metadatos cumplan las funciones que el usuario espera (búsqueda, evaluación, integración, etc.).

4.5. Consistencia lógica y coherencia

La consistencia lógica y coherencia busca garantizar que los elementos se conciben de manera razonable o congruente con las definiciones y conceptos del estándar utilizado y que se presenten de manera coherente. Implica también que dos o más valores no entren en conflicto entre sí o que los datos de los elementos que describen al objeto de información no sean contradictorios o inconsistentes. Por otro lado, alude a que un mismo elemento sea utilizado para varios tipos de datos similares. Una forma de ilustrar este criterio es a través de la capacidad de búsqueda de los usuarios, quienes esperan poder buscar objetos similares utilizando los mismos criterios.

4.6. Actualidad

La actualidad se puede definir como la utilidad del metadato a lo largo del tiempo, ya que los valores de los datos pueden variar en el tiempo al punto que no representen o difieran del objeto original que describen. Surgen también problemas de calidad cuando el dato “envejece” o cambia en el mundo real y pierde validez, o cuando el objeto de destino cambia pero los datos no. Ejemplos de este criterio pueden presentarse cuando los URI están dañados, los vocabularios controlados se vuelven obsoletos o los objetos de información se mueven, ya sea en estantes, sitios web o mapas conceptuales de una disciplina intelectual y los metadatos pierden la sincronización con esos movimientos.

4.7. Visibilidad y accesibilidad

La visibilidad se refiere a la presencia del repositorio en directorios o recolectores tanto nacionales como internacionales. Por su parte, la accesibilidad mide el grado en que se pueden acceder a los datos. Se define en dos niveles: accesibilidad cognitiva y accesibilidad física / lógica.

La accesibilidad cognitiva se refiere a la comprensión de la información contenida en los metadatos por parte del usuario; es decir, la dificultad para entender el texto o el dato. Por otro lado, la accesibilidad física o lógica se entiende como la facilidad para encontrar la información en el repositorio, la asociación o vinculación al objeto representado y las relaciones entre los propios datos y otros elementos. También se considera el uso de formatos obsoletos, inusuales o propietarios, que solo pueden leerse con equipos o softwares especiales o a través de un costo adicional.

4.8. Políticas

Se refiere al cumplimiento de la misión y objetivos de la plataforma, a las directrices y estándares internacionales para su implementación, la preservación de contenidos, reutilización de los metadatos, acceso abierto, entre otros aspectos formales que garanticen la calidad de la información.

4.9. Seguridad de los datos

Aspecto relacionado con la protección de los datos, tanto del software sobre el que funciona la plataforma como de los objetos digitales o datos contenidos, para evitar así su pérdida o modificaciones no autorizadas y garantizar su confidencialidad, autenticidad o integridad.

Una vez definidas estas categorías generales, se muestran sus respectivos criterios de calidad e indicadores para cada caso, en el cuadro siguiente:

Categorías de calidad	Criterios de calidad		Indicadores de calidad
Compleitud e integridad	1	Campos obligatorios completados.	Porcentaje de campos obligatorios completados por registro.
	2	Campos con datos incompletos ⁴ .	Porcentaje de campos con datos incompletos.
	3	Campos con datos que no le corresponden.	Porcentaje de campos con datos que no le corresponden.
Exactitud y precisión	4	Campos con datos duplicados.	Porcentaje de datos considerados duplicados.
	5	Campos con datos normalizados (control de autoridades).	Porcentaje de datos normalizados.
	6	Campos conteniendo más de un dato.	Porcentaje de campos (preferentemente aquellos no repetibles) con más de dos datos.
	7	Los datos cumplen con el uso de esquemas o vocabularios controlados especificados en las Directrices.	Porcentaje de datos (de ser el caso) que cumplen con el esquema o vocabulario controlado requerido.
Procedencia y confiabilidad	8	Se cuenta con un responsable (s) para la creación, extracción o transformación de los metadatos	Cumple / no cumple
	9	Cuenta con una metodología para la creación, extracción o transformación de los metadatos.	Cumple / no cumple
	10	Se cuenta con datos acerca de las transformaciones que han sufrido los metadatos a lo largo del tiempo.	Cumple / no cumple
Conformidad con las expectativas	11	Las metadatos cumplen con los requisitos funcionales del repositorio.	Cumple / no cumple
	12	Los vocabularios controlados utilizados están alineados con los tipos de usuarios a servir y los requisitos funcionales del repositorio.	Cumple / no cumple
	13	Los registros contienen la información suficiente para cumplir con las expectativas de los usuarios a servir y las necesidades de la institución.	Tasa de campos poblados por registro. Grado de satisfacción de usuarios

⁴ Refiere a que el campo contiene algún dato, pero éste no representa al objeto de información que describe.

Consistencia lógica y coherencia	14	Los datos cumplen con la sintaxis (lista de valores, formatos, restricciones, etc.) especificada en las Directrices.	Porcentaje de datos (de ser el caso) que cumplen con la sintaxis especificada.
	15	Existen contradicciones entre los datos de un mismo registro.	Porcentaje de registros que contienen datos contradictorios.
	16	Existen datos o valores que se repiten en diferentes campos (redundancia).	Porcentaje de campos con datos redundantes.
	17	Campos con datos inconsistentes (sin significado o irrelevantes).	Porcentaje de campos con datos inconsistentes.
Actualidad	18	Las políticas o procedimientos se actualizan cuando nuevos metadatos son agregados.	Cumple / no cumple
	19	Los datos se actualizan cuando los objetos de información cambian.	Cumple / no cumple
	20	Los vocabularios controlados se actualizan cuando quedan obsoletos o no cumplen con describir los objetos de información.	Cumple / no cumple
Visibilidad y accesibilidad	21	Figura en directorios o recolectores nacionales e internacionales.	Cumple / no cumple
	22	Utilizan identificadores persistentes y demás datos para la localización y acceso al objeto de información.	Porcentaje de campos con datos de localización o acceso que enlazan al objeto de información original.
	23	Los datos son fácilmente accesibles y recuperables (búsqueda, evaluación, obtención de estadísticas).	Grado de recuperación y accesibilidad a la información.
Políticas	24	Las Directrices o políticas cumplen con estándares internacionales para el registro, recuperación, integración e interoperabilidad.	Cumple / no cumple
	25	Existe documentación de acceso público sobre la misión, objetivos, directivas y normativas del repositorio.	Cumple / no cumple
	26	Existen políticas de acceso abierto y derechos de autor.	Cumple / no cumple
	27	Existe una política de preservación, uso y reutilización de los datos.	Cumple / no cumple
	28	Existen personas de contacto y ayudas disponibles.	Cumple / no cumple

Seguridad de los datos	29	Existe un procedimiento para copias de seguridad.	Cumple / no cumple
	30	Se cuenta con un registro de datos eliminados.	Cumple / no cumple
	31	Se cuenta con un procedimiento para la validación de datos.	Cumple / no cumple
	32	Cada registro cuenta con un identificador unívoco e información sobre su origen y modificaciones.	Cumple / no cumple

V. Resultados del análisis de los metadatos cosechados en ALICIA

Para efectos del presente documento se han contemplado las siguientes categorías y criterios de calidad: completitud e integridad (criterios 1 y 2), exactitud y precisión (criterios 4 al 7), consistencia lógica y coherencia (criterios 14 y 17), y visibilidad y accesibilidad (criterio 22).

Hay que considerar que algunos de estos criterios analizados pueden afectar a un mismo metadato, por ello resulta complejo determinar qué tipos de errores califican en un solo nivel; por ejemplo, los errores tipográficos u ortográficos atañen a varias de las categorías de calidad descritas, como la integridad, exactitud, coherencia y accesibilidad.

De manera general, podemos afirmar que del estudio realizado los errores más recurrentes en los metadatos que contiene el Repositorio Nacional ALICIA son: falta de normalización, duplicidad, datos incompletos, datos inconsistentes, incumplimiento de los esquemas, vocabularios controlados, valores de datos y restricciones especificados en las Directrices, errores tipográficos u ortográficos o datos que no corresponden al metadato donde ha sido consignado.

Dicho esto, pasaremos a mencionar los principales problemas encontrados en cada elemento Dublin Core examinado, en la sección Anexo, se pueden ver ejemplos más detallados de ellos:

- **dc.creator y dc.contributor** (advisor, editor, illustrator): el principal problema encontrado en estos metadatos es la falta de normalización para nombres personales e institucionales, lo que conlleva otros tipos de errores como duplicidad, inconsistencia, falta de completitud y errores tipográficos u ortográficos. Además, se identifican datos que no corresponden al elemento o que contienen más de un dato. En muchos casos se consignan nombres de personas con calificaciones, como “Doctor”, “Ingeniero”, “Licenciado”, “Magister” o sus abreviaturas.
- **dc.coverage.spatial**: en este elemento encontramos que las Directrices de ALICIA no son lo suficientemente específicas, lo que puede llevar a confusión en el uso de este metadato. Principalmente hallamos problemas con la normalización de los datos, por lo que debería considerarse el empleo de un tesoro o vocabulario controlado de nombres geográficos. Por otro lado, se observan datos que no corresponden a este elemento y errores tipográficos u ortográficos, como el uso indistinto de mayúsculas/minúsculas.
- **dc.date.issued**: las Directrices especifican el uso de la Norma ISO 88601 para el formato de fechas; sin embargo, el problema más recurrente en este elemento es precisamente que no se cumple con esta norma. Ejemplo: fechas registradas utilizando la barra oblicua (/) para separar año, mes, día; uso de guiones o puntos al final del dato del año (cuando no se cuenta con la fecha completa); uso de letras correspondientes a siglos, etc.
- **dc.format**: metadato poco utilizado, los datos deben registrarse, según las Directrices, conforme a la lista proporcionada por IANA; sin embargo, se observa que muchos de

los datos no cumplen con la sintaxis ni los valores requeridos. También se presentan problemas de integridad al registrarse datos que no corresponden al elemento; además de datos inconsistentes y uso indistinto de mayúsculas/minúsculas y espaciados.

- **dc.identifier.citation:** metadato poco utilizado en el que se debe consignar la cita de la publicación en el formato preferido por la institución. Vemos que en las Directrices de ALICIA no quedan claras en cuanto a su contenido, lo que ocasiona que solo se esté registrando como valor el nombre del formato, por ejemplo “APA”; adicionalmente, se consignan de diversas maneras.
- **dc.identifier.journal:** se encuentran principalmente errores en cuanto a la duplicidad y al uso indistinto de mayúsculas/minúsculas y tildes; así como datos que no pertenecen al elemento.
- **dc.identifier.isbn/issn:** en las Directrices se confunde los valores para registrar el ISBN (sin guiones) y el ISSN (con guiones), siendo este uno de los problemas encontrados en el análisis. Por otro lado, se encuentran registros que adicionalmente al número consignan las iniciales palabras “ISBN” o “ISSN”. También se presentan datos inconsistentes (sin semántica).
- **dc.identifier.uri:** metadato para el registro del DOI o Handle de la publicación. Es un elemento clave para identificar unívocamente el objeto de información. Si bien es cierto no se ha realizado un reporte de enlaces rotos, si se han encontrado casos en los que el enlace no funciona. Por otro lado, se encuentran también otro tipo de errores, como datos incompletos, datos que no corresponden al metadato y valores inconsistentes.
- **dc.language.iso:** según las Directrices, para registrar el idioma de la publicación es necesario respetar la norma ISO-639-3, que codifica los países con 3 letras. Los principales problemas encontrados son: datos que no cumplen con la norma (pe. códigos de solo 2 letras), uso indistinto de mayúsculas/minúsculas, datos inconsistentes o sin significado, y datos que no corresponden al campo.
- **dc.publisher:** en este metadato también persisten problemas con la normalización, además errores tipográficos u ortográficos, que conllevan también duplicidad. Así mismo, se encuentra falta de consistencia (datos sin semántica).
- **dc.relation y dc.relation.uri:** la descripción de estos metadatos no se encuentra clara en las Directrices. Encontramos en ellos especialmente datos que no le corresponden al elemento y datos inconsistentes.
- **dc.rights:** se pueden apreciar principalmente errores con la sintaxis que se especifica en las Directrices, que incluye igualmente errores tipográficos. Además, se encuentran datos que no corresponden al elemento.
- **dc.rights.uri:** refiere a la licencia de la publicación para su uso en acceso abierto; sin embargo, en las Directrices no se especifica una sintaxis en particular, siendo este el inconveniente primordial que afecta la calidad de este metadato. Se presenta el uso de texto libre, datos inconsistentes e incompletos o que cuentan con errores tipográficos. Así también, datos que no corresponden al elemento.
- **dc.source:** este metadato se utiliza tanto para registrar el nombre de la institución como del repositorio institucional, lo que origina no solo errores al momento de registrar los datos sino redundancia en la información. Los errores más importantes

encontrados en este metadato son: duplicidad, errores tipográficos y ortográficos (que también originan duplicidad), datos que no corresponden al elemento, datos sin significado (inconsistentes) e incompletos.

- **dc.subject:** es un elemento clave para la búsqueda y recuperación de la información. Debido a que cada institución utiliza diferentes vocabularios (encabezamientos de materia, tesauros o simplemente palabras clave no controladas) la calidad de este metadato se ve fuertemente afectada, siendo uno de los metadatos que presenta mayores problemas. Además, encontramos registros con este metadato vacío, datos inconsistentes, datos incompletos, duplicidad, datos que no corresponden al metadato y dos o más datos en el mismo campo (separados por barras oblicuas, comas, punto y comas o conectores: “y”, “o”).
- **dc.title:** se encuentran datos inconsistentes (sin semántica) lo que afecta la recuperación de la información. Además, errores tipográficos u ortográficos (uso indistinto de mayúsculas/minúsculas, tildes, caracteres no reconocidos, espaciado, uso de comillas, etc.). En menor medida, se presentan datos que podrían ser duplicados.
- **dc.type:** se visualizan problemas con la sintaxis y la semántica, al no respetarse los valores, incluso valores inexistentes (p.ej.: “info:pe-repo/semantics/jurisprudencia”) o con texto libre (p.ej.: “Artículo revisado por pares”), además de errores tipográficos. También se revelan datos que no corresponden al elemento, en gran medida datos sobre la metodología utilizada en la investigación. A pesar que en las políticas de cosecha de LA Referencia se recomienda su uso para indicar también el estado del proceso de publicación, apreciamos que no está siendo utilizado para este propósito.
- **dc.description.uri:** según las Directrices, el campo se viene utilizando para el registro de obra realizada para la obtención de un grado o título profesional, contenido que no le corresponde. Por otro lado, se visualizan errores de tipo tipográfico (uso indistinto de mayúsculas y minúsculas, falta de tildes, palabras mal escritas); así como valores que no respetan la tipología especificada (p.ej.: Tesis para maestría, Tesis / Trabajo de investigación, Informe de tesis).

VI. Conclusiones y recomendaciones

- Mediante este análisis podemos concluir que la calidad de los metadatos cosechados en el Repositorio Nacional ALICIA afecta en la búsqueda, recuperación, localización y visibilidad de la investigación académica y científica del país. Se debe garantizar una estructura estandarizada de datos en los repositorios institucionales que correspondan a los mismos metadatos que cosecha ALICIA.
- Cada institución aportante debe establecer los procedimientos para revisar y validar sus registros antes de dar por concluido el proceso de catalogación de cada ítem, así evitar errores que luego son transmitidos a ALICIA.
- Las Directrices de ALICIA deben ser consideradas de uso obligatorio por las instituciones que la integran, a manera de documento de consulta ante cualquier duda; así se mantendrá la uniformidad y calidad de los datos. Por ejemplo, incluir una sección de “no confundir” para metadatos cuyos valores pueden ser similares o prestarse a equivocaciones; así como relaciones entre metadatos recíprocos.
- Las Directrices de ALICIA cuentan, por un lado, con metadatos que están siendo utilizados para albergar contenido que no les corresponde o son “infrautilizados”; y por otro, algunos metadatos no se describen lo suficiente para evitar confusiones o inconsistencias. Se debe especificar mejor el contenido de los siguientes elementos: dc.coverage.spatial, dc.date*, dc.description.uri, dc.format, dc.type, dc.relation*, dc.source, dc.description.provenance.
- Realizar periódicamente controles de calidad para la detección de posibles datos duplicados en elementos críticos como dc.creator, dc.contributor*, dc.publisher, dc.subject y dc.source. Además un control de enlaces rotos.
- Se debe prestar atención en la normalización de los datos a través de listas de autoridades en los siguientes metadatos: dc.creator, dc.contributor*, dc.subject, dc.publisher, dc.source⁵. Llegar a acuerdos a través de RENARE para el uso de tesauros controlados para las materias.
- Se debe enfatizar el respeto al uso de esquemas, vocabularios controlados, valores de datos o restricciones que se especifican en las Directrices de ALICIA, especialmente en metadatos como: dc.date, dc.format, dc.identifier.isbn, dc.identifier.issn, dc.identifier.uri, dc.language, dc.relation, dc.rights, dc.rights.uri y dc.type.
- Para los datos de fecha (dc.date*) se debe especificar con mayor detalle a qué dato se refiere cada uno de los calificadores.
- Para el metadato dc.format, las Directrices de ALICIA recomiendan el uso de la lista de IANA; sin embargo, tanto en el DCMES como en las políticas de LA Referencia se menciona que no solo refiere al formato del archivo electrónico, sino también al medio físico o dimensiones del recurso; por ello, la información encontrada es muy variada. Se recomienda el uso de calificadores como *extent (para la extensión del recurso, pe. “21 minutos”, “24 MB”, “15 folios”) y *medium (para describir al recurso físico, pe. “DVD”, “óleo sobre madera”).

⁵ El contenido de este elemento se recomienda ser migrado a dc.identifier*

- En el caso del metadato dc.identifier.citation, vemos que no se está cumpliendo con su propósito, sino simplemente se está utilizando para la declaración de un estilo de citación, por ejemplo “APA” y no la cita correspondiente al recurso.
- Utilizar la riqueza que ofrece DCMI para establecer relaciones entre diferentes objetos de información a través de calificadores en el elemento dc.relation, como por ejemplo, isPartof, isReplacedBy o isVersionOf.
- Se debe establecer claramente la diferencia entre los metadatos dc.identifier.uri y dc.relation.uri.
- Para el metadato dc.rights, considerar las recomendaciones de LA Referencia en relación a la fecha de acceso para los documentos con embargo:
“Si el nivel de acceso es: info:eu-repo/semantics/embargoedAccess
Se debe consignar la fecha de fin de embargo
info:eu-repo/date/embargoEnd/<YYYY-MM-DD>
<dc:date>info:eu-repo/date/embargoEnd/2015-12-31</dc:date>”
- El elemento dc.source se viene usando para el nombre de la institución y del repositorio; sin embargo, LA Referencia especifica que: “Utilizar únicamente si el recurso descrito es el resultado de la digitalización de originales no digitales. En caso contrario, utilizar el elemento Relation. Opcionalmente, se pueden añadir metadatos sobre la ubicación actual y la signatura de la publicación digitalizada”. Adicionalmente, recomiendan utilizar los elementos dc.identifier.instname (para el nombre de la institución), dc.identifier.reponame (para el nombre del repositorio) y dc.identifier.repourl (para la URL del repositorio).
- Utilizar el elemento dc.type para registrar, como sugieren las Directrices de LA Referencia, el estado del proceso de la publicación, con la sintaxis y vocabulario controlado sugerido.
- En cuanto a la búsqueda y recuperación de información en ALICIA:
 - o Los campos dc.description.abstract y grado académico (para tesis) deberían ser recuperables.
 - o Identificar los repositorios desde donde se están cosechando datos en los que los caracteres (como los acentuados) no se están reconociendo adecuadamente.
 - o El criterio de búsqueda “Etiqueta” no es entendible.
- Con respecto a la interfaz pública de ALICIA:
 - o Incluir un “botón” (siguiente o anterior) para poder navegar entre registros resultantes de una búsqueda
 - o En la página principal, en las secciones “Fuentes de información”, “Áreas de investigación” e “Instituciones integrantes” se debería enlazar los números de documentos que se mencionan, no en los nombres de las etiquetas.
 - o De ser posible, incluir estadísticas de uso del registro (visualizaciones, clics al documento / repositorio original, descargas).
- Finalmente, medir la calidad de los metadatos con los indicadores propuestos y su evolución (mejora) en el tiempo. Así también, realizar estudios de usabilidad y de usuarios.

VII. Referencias

- Azorín Millaruelo, C., Bernal Martínez, I., Cívico Martín, R., Cózar Santiago, A., Guzmán Pérez, C., Losada Yáñez, M., ... Prats Prat, J. (2017). Guía para la evaluación de repositorios institucionales. Recolecta; FECYT; Ministerio de Economía, Industria y Competitividad de España, CRUE. Recuperado de <https://www.recolecta.fecyt.es/sites/default/files/contenido/documentos/2017GuiaEvaluacionRecolectaFECYT.pdf>
- Baptista, A. A. (2017). Desafios á comunidade iberoamericana de metadatos em repositórios digitais para maximização da interoperabilidade. En *Repositórios digitais : teoria e prática* (Vol. 91, pp. 195–203). Curitiba: Universidade Tecnológica Federal Do Paraná. Recuperado de <http://repositorio.utfpr.edu.br/jspui/handle/1/2495>
- Congreso de la República. Ley 30035 – Ley que regula el Repositorio Nacional Digital de Ciencia, Tecnología e Innovación de Acceso Abierto, 2013.
- Consejo Nacional de Ciencia, Tecnología e Innovación Tecnológica (2017). Directrices para el procesamiento de información en los repositorios institucionales. Recuperado de: http://portal.concytec.gob.pe/images/documentos/alicia/directrices_repositorio.pdf
- Consejo Nacional de Ciencia, Tecnología e Innovación Tecnológica (2016). Política Nacional para el Desarrollo de la Ciencia, la Tecnología y la Innovación Tecnológica – CTI. Recuperado de: <http://portal.concytec.gob.pe/index.php/concytec/estrategias/628-politica-nacional-para-el-desarrollo-de-la-ciencia-la-tecnologia-y-la-innovacion-tecnologica-cti>
- De Giusti, M. R., Oviedo, N. F., Lira, A. J., & Villarreal, G. L. (2013). Control de integridad y calidad en repositorios DSpace. En *III Conferencia de Bibliotecas y Repositorios Digitales de América Latina (BIREDIAL) y VIII Simposio Internacional de Bibliotecas Digitales (SIBD) (Costa Rica, 2013)* (pp. 1–14). San José, Costa Rica. Recuperado de <http://hdl.handle.net/10915/30524>
- Dublin Core Metadata Initiative. *Dublin Core Metadata Element Set, Version 1.1: Reference Description*. Recuperado de: <http://dublincore.org/documents/dces/>
- Dushay, N., & Hillmann, D. I. (2003). Analyzing metadata for effective use and re-use. En *Dublin Core Metadata Conference* (p. 17). Seattle, Washington. Recuperado de <http://dl.acm.org/citation.cfm?id=1383296.1383318>
- Fushimi, M., Genovés, P., Pené, M., & Unzurrunzaga, C. (2011). Indicadores para evaluar repositorios universitarios argentinos, de la teoría a la práctica. En *Segundo Taller de Indicadores de Evaluación de Bibliotecas* (p. 26). La Plata, Argentina: Universidad de la Plata. Facultad de Humanidades y Ciencias de la Educación. https://doi.org/ISSN_1853-7324
- Hillmann, D. I. (2008). Metadata quality: from evaluation to augmentation. *Cataloging and Classification Quarterly*, 46(1), 65–80. <https://doi.org/10.1080/01639370802183008>
- Hillmann, D. I., & Bruce, T. R. (2004). The continuum of metadata quality: defining, expressing, exploiting. *Metadata in Practice*. Recuperado de <https://bit.ly/2GY480y>

- Hillmann, D. I., Dushay, N., & Phipps, J. (2004). Improving metadata quality: augmentation and recombination. En *International Conference on Dublin Core and metadata applications metadata across languages and cultures* (pp. 1–8). Shanghai, China. Recuperado de <http://portal.acm.org/citation.cfm?id=1383385.1383395>
- Información y documentación - Procesos de gestión de documentos - Metadatos para la gestión de documentos. Parte 1: Principios. ISO 23081-1:2006 (2008). *Revista Española de Documentación Científica*, 31(2), 273-301.
- Medrano, J. F., Figuerola, C. G., & Alonso Berrocal, J. L. (2012). Repositorios digitales en España y calidad de los metadatos. *Scire: Representación Y Organización del Conocimiento*, 18(2), 109–121. Recuperado de <https://www.iberid.eu/ojs/index.php/scire/article/view/4231/3811>
- Myrseth, P., Stang, J., & Dalberg, V. (2011). A data quality framework applied to e-government metadata: a prerequisite to establish governance of interoperable e-services. En IEEE (Ed.), *E-Business and E-Government (ICEE), 2011 International Conference on* (pp. 1–4). Recuperado de http://www.semicolon.no/wp-content/uploads/2013/02/Myrseth_Shanghai_2011_paper.pdf
- Ochoa, X., & Duval, E. (2006). Metrics for learning object metadata. En Association for the Advancement of Computing in Education (Ed.), *EdMedia: World Conference on Educational Media and Technology* (pp. 1004–1101).
- Ochoa, X., & Duval, E. (2009). Automatic evaluation of metadata quality in digital repositories. *International Journal on Digital Libraries*, 10(2), 67–91. <https://doi.org/10.1007/s00799-009-0054-4>
- OpenAIRE (2019). *OpenAIRE Guidelines for Literature Repository Managers v4*. Recuperado de: https://guidelines.openaire.eu/en/latest/literature/index_guidelines-lit_v3.html
- Open Society Institute (2002). *Iniciativa de Budapest para el Acceso Abierto*. Recuperado de <https://www.budapestopenaccessinitiative.org/translations/spanish-translation>
- Palavitsinis, N., Manouselis, N., & Sánchez-Alonso, S. (2014). Metadata quality in digital repositories: empirical results from the cross-domain transfer of a quality assurance process. *Journal of the Association for Information Science and Technology*, 65(6), 1202–1216. Recuperado de <https://onlinelibrary.wiley.com/doi/pdf/10.1002/asi.23045>
- Red Federada de Repositorios Institucionales de Publicaciones Científicas (2019). LA Referencia. Recuperado de <http://www.lareferencia.info/es/>
- Reglamento de la Ley Nº 30035, Ley que regula el Repositorio Nacional Digital de Ciencia, Tecnología e Innovación en Acceso Abierto : Decreto Supremo Nº 006-2015-PCM, 2015.
- Rousidis, D., Garoufallou, E., Balatsoukas, P., & Sicilia, M. A. (2014). Metadata for big data: a preliminary investigation of metadata quality issues in research data repositories. *Information Services and Use*, 34(3–4), 279–286. <https://doi.org/10.3233/ISU-140746>

Stivilia, B., Gasse, L., Twidale, M. B., Shreeves, S. L., & Cole, T. W. (2004). Metadata quality for federated collections: IQ concepts, models, case studies. En *Proceedings of the Ninth International Conference on Information Quality (ICIQ-04)* (pp. 111–125). Cambridge. Recuperado

de <https://www.ideals.illinois.edu/bitstream/handle/2142/721/MetaDataQuality4FederatedCollections.pdf?sequence=4&isAllowed=y>

UNESCO (2019). *¿Qué es acceso abierto?* Recuperado de <https://es.unesco.org/open-access/%C2%BFqu%C3%A9-es-acceso-abierto>

VIII. Anexo

Se presentan los cuadros elaborados al analizar los metadatos del Repositorio Nacional ALICIA. Se consideran los principales metadatos, la casuística encontrada y los ejemplos que las grafican:

Metadato:	dc.creator
Descripción:	Consigna el nombre del autor de la publicación.
Ejemplo:	-Cervantes Vizcarra, José Gabriel -Instituto Nacional de Salud
Casuística:	
1. Datos no normalizados	Gerencia de Promoción y Difusión, INDECOPI
	Instituto Nacional de Salud, Perú, Rev Peru Med Exp Salud Publica,
	Administración Técnica del Distrito de Riego Ilave
	Facultad de Medicina Humana de la Universidad de San Martín de Porres
2. Tildes no reconocidas	Per?. Consejo Nacional del Ambiente
3. Errores tipográficos / ortográficos	UNIVERSIDAD CATÓLICA SEDES SAPIENTIAE (mayúsculas / minúsculas)
	ALCANTARA-BOCANEGRA, Fernando (mayúsculas / minúsculas)
	León-Kanashiro, Laura-Raquel (guiones para separar palabras)
	Pumacayo Puño Gladys Giovanna (sin coma entre apellidos y nombres)
	BULEJES ARREDONDO , DANIEL ROSWELD (espacio demás entre la coma)
	TORRES CONCHA STEIN; BRUNO (punto y coma en vez de coma)
	TEJADA ZANABRIA., YELKA ELIZABETH (uso de punto + coma)
	Pusztai, L (falta punto en inicial del nombre)
. Martinez Bustamante, Jessica (punto y espacio antes del nombre)	
4. Dos o más datos en el mismo campo (algunos sin separación)	Facultad de Ciencias Naturales y Matemáticas. Universidad Nacional FedericoVillarreal (Perú). (falta de espacio entre dos palabras)
	Escorra-Mayaute, Luis Miguel; Universidad de Lima (Perú)
	Llano Quispe, Luzmila Ramos Vilcatoma, Mabel Erika
5. Uso de cronológicos	Ceroni Stuva, Aldo; Universidad Nacional Agraria La Molina (Perú).
	Vargas Bianchi, Lizardo, 1976

6. Duplicados	Instituto Nacional de Defensa de la Competencia y de la Protección de la Propiedad Intelectual, INDECOPI. Tribunal de Defensa de la Competencia / Tribunal de Defensa de la Competencia y de la Protección de la Propiedad Intelectual
	PNUD. Programa de Naciones Unidas para el Desarrollo / Programa de las Naciones Unidas para el Desarrollo (PNUD)
7. Datos inconsistentes o incompletos	<p>-, -</p> <p>-, Vicerrectorado de Investigación</p> <p>M., C. M.</p> <p>Taipe (solo aparece el apellido)</p> <p>Bendezu Hinostriza HectorBendezu Hinostriza HectorBendezu Hinostriza HectorBendezu Hinostriza Hector</p> <p>et al.</p> <p>U</p> <p>Otros</p> <p>-, Varios</p> <p>F., Tordoya (sin apellido completo)</p> <p>33, VOX JURIS; Universidad de San Martín de Porres</p> <p>Facultad de Derecho</p> <p>Unidad de Biblioteca y Repositorio Institucional</p> <p>EQI (no se conoce el significado)</p>
8. Datos que no corresponden al elemento (o datos de más)	<p>El documento presenta información estadística sobre la labor de las diferentes comisiones, direcciones, servicios y salas que componen el Indecopi. (descripción o resumen)</p> <p>Castillo, Daladier; Universidad Nacional Jorge Basadre Grohmann, Av. Miraflores S/N (Ciudad Universitaria), Apartado Postal 316 Tacna-Perú.</p> <p>Lora G., Alfredo; Egresado de la Universidad Nacional Agraria La Molina (Perú).</p> <p>Castagnino Pastor, Claudio Stefano S.; economista; Oracle Certified Professional, Java Programmer y CompTIANetwork+Certified.</p>

Metadato:	dc.contributor: advisor, editor, illustrator
Descripción:	Consigna el nombre de los asesores de tesis, responsables de la edición/supervisión del contenido del recurso, responsables de creación o generación de diseños o ilustraciones, respectivamente.
Ejemplo:	Cifre Wibrow, Patricia
Casística:	
1. Datos inconsistentes, incompletos o que no identifica a la persona o institución	DIR
	Hidalgo
	-, Vicerrectorado de investigación
	N.E
	ánchez Díaz, Sebatían S
	Jorge
	-, -
	., -
	[s.a.]
	"No especifica"
	Chauca Saavedra, Oscar David Chauca Saavedra Oscar David Chauca Saavedra Oscar David Chauca Saavedra Oscar David
A, A	
metal	
2. Caracteres acentuados no reconocidos	Fern?andez Henr?quez, Luz Olinda
	Rodr?guez Ch?vez, Iv?n
3. Datos no normalizados	Williams León de Castro, Marta Leonor del Rosario
	VICERRECTORADO DE INVESTIGACIÓN DE LA UNIVERSIDAD NACIONAL DE SAN MARCO - LIMA - PERÚ
	del Carpio Flores, Sofia
	DERECHO PUCP
	Francisco Quiroz Chueco
4. Errores tipográficos / ortográficos	VICERRECTORADO DE INVESTIGACIÓN DE LA UNIVERSIDAD NACIONAL DE SAN MARCO - LIMA - PERÚ (mayúsculas / minúsculas)
	PALACIOS SÁNCHEZ, JOSÉ MANUEL (mayúsculas / minúsculas)
	MEza Minanya, Edgar William (mayúsculas / minúsculas)
	De-la-Puente-Arbaiza, Carlos-Alberto (guiones entre las palabras)
	Vilchez Reyes, Maria Adriana (sin tildes)
5. Uso de cronológicos	Cruz Alcedo, Gastón Educado, 1963-
	Castillo Córdova, Luis, 1971-

6. Uso de calificadores	Dr. Campos Martinez, Hernando Martin
	Dr. Campos Martinez, José Jorge
	Mg. Zairi Sacsi, Doris Margarita
	Otros: Dra., Ms, C.P.C., MG, Obs., Dr. C.D., Lic., Q.F., Ing.; MG CPC... ;MG ING.
7. Datos que no corresponden al elemento	REVISTA DE INVESTIGACIÓN ALTOANDINAS
	antonio.a67@gmail.com
	TESIS
	jdelcarpio@uni.edu.pe
	info:eu-repo/semantics/openAccess
	https://creativecommons.org/licenses/by-nc/3.0/
	CONSTRUÇÃO DA SEXUALIDADE NA IDADE GÊNERO DE REPRODUÇÃO PERSPECTIVA MÉDICA. Bases culturais sócio-histórico
	Estudio financiado con fondos del CSI y CONCITEC
	Estudio financiado con fondos del CONCyTEC
	N°096-2014-FONDECYT-DE
	N° K13051.
	612.1 - Sangre y circulación En: Journal of Geophysical research, v. 111, n. A6, (June 2006), p. A06306, 1-3.
8. Posibles duplicados	Domingo Flores, Antonio Domingo Flores, Antonio Mariano
	Pajuelo Bustamante de Morales, Rossana Pajuelo Bustamante de Morales, Roxana
	Paca Pantigoso, F. Romero Paca Pantigoso, F. Romeo
9. Dos o más datos en el mismo campo	FINCYT e Instituto Tecnológico de la Producción del Ministerio de la Producción
	FIDECOM (Fondo de Investigación y Desarrollo para la Competitividad) del Ministerio de la Producción del Perú y de la empresa Andean Roots SRL.
	Neciosup Severino, David; Huamán Larios, Julissa Bertha
10. Elemento vacío	

Metadato:	dc.coverage.spatial
Descripción:	País o ciudad donde se realizó el documento
Ejemplo:	Perú
Casística:	
1. Datos no normalizados	Lima. Perú
	Lima, Perú
	Cordillera Blanca
	Cuenca Piura
	Cuenca Quilca - Vitor – Chili
	Costa peruana; Desierto costero
	SECTOR METALMECÁNICO EN LIMA
2. Datos que no corresponden al elemento	2014
	2016
	Siglo XX
	Descriptiva-Explicativa
	se tomaron muestras de biomasa aérea (hojas, altura, estípite) y suelo para palma aceitera
3. Datos inconsistentes o incompletos	Public
	actual
	Regional
	2012-2016
	Global
	Trabajadores
	Buddlejaceae, Buddleja, Peru, endemism, endemic plants.Buddlejaceae, Buddleja, Peru, endemism, endemic plants.
4. Datos tipográficos / ortográficos	caribe
	chimbote
	Cuenca Tacna
	LAMBAYEQUE
	andes

Metadato:	dc.date
Descripción:	Fecha de publicación. Cumpe con la norma ISO-88601.
Ejemplo:	2012-10-28 2016
Casuística:	
1. Dos o más datos en el mismo elemento	2013-12-01 / 2014-10-18
2. No cumple con las directrices	21/03/2014
	s. XVI-XVII
	[193-]
	2013- 2011.

Metadato:	dc.format
Descripción:	Se consigna el formato de la publicación de acuerdo a la lista de IANA.
Ejemplo:	application/pdf application/msword
Casuística:	
1. No cumple con las directrices	PDF
	6 páginas
	application/vnd.openxmlformats-officedocument.wordprocessingml.document
	0,67 MB
	5 láminas
	1:24:14
2. Datos inconsistentes o incompletos	application/flv
	application/epub
	audio/x-wav
	49 Application/pd
3. Errores tipográficos	aplication/pdf Application/pdf
4. Datos que no corresponden al elemento	Enfermería en Centro Quirúrgico
	En: Annales Geophysicae, v. 24, n. 5 (January 2006), p. 1305-1310
	Educación básica primaria
	Restringido

Metadato:	dc.identifier.citation
Descripción:	La institución adoptará un estándar para la citación de los documentos, según la política de la institución (APA, ISO, Vancouver; Chicago, otros).
Casuística:	
1. No cumple con las directrices	APA
	Formato APA
	Normas de Vancouver

Metadato:	dc.identifier.journal
Descripción:	Consigna el nombre completo de la revista (obligatorio cuando el tipo de documento es un artículo)
Ejemplo:	Revista Peruana de Biología
Casuística:	
1. Errores tipográficos / ortográficos	Un vicio absurdo / Un Vicio Absurdo (mayúsculas / minúsculas)
	TECNIA (mayúsculas / minúsculas)
	Boletin de Investigacion Instituto Tecnologico Pesquero del Peru (sin tildes)
2. Duplicados	Un vicio absurdo / Un Vicio Absurdo
3. Datos que no corresponden al elemento	Inf. Inst. Mar Perú nº 135, 1998
	[Fotografía del Museo de Arqueología Josefina Ramos de Cox]. (Lima, 2018). Colección de textiles. Museo de Arqueología Josefina Ramos de Cox- IRA-PUCP. Lima, Lima.
	Evaluación de los niveles sonoros en la ciudad de Bagua. Departamento Amazonas, 2015.pdf

Metadato:	identifier.isbn / issn
Descripción:	Número internacional asignado a un monográfico y publicación seriada, respectivamente
Ejemplo:	20754469 (ISBN - sin guiones) 0378-7702
Casuística:	
1. No cumple con las directrices	ISSN 1684-1662
	ISSN (impresa): 1726-2216
2. Sin guiones (ISBN)	19978812
	0378 7702
3. Datos inconsistentes o incompletos	10
	7
	MAICA 6 2017 3

Metadato:	identifier.uri
Descripción:	Se utiliza para el registro de identificadores internacionales como handle o DOI. También, puede ser utilizado para el identificador personalizado de la institución
Ejemplo:	http://tesis.pucp.edu.pe/repositorio/handle/123456789/6008 http://dx.doi.org/10.15381/rpb.v22i3.11431
Casuística:	
1. Datos inconsistentes o incompletos	https://doi.org/
2. Datos que no corresponden al elemento	http://orcid.org/0000-0001-7168-8613
	http://creativecommons.org/licenses/by-nc-nd/4.0/
	http://www.osinergmin.gob.pe/seccion/centro_documental/gart/publicaciones/gart-card/GartCard2016-01/Master%20Disco%20%20-%20Publicaciones/archivos/contenido/pdf/InfEconFin/Financiera2015T2.pdf
	http://revistas.unife.edu.pe/index.php/phainomenon/article/view/105
3. Handle que no funcionan	http://hdl.handle.net/11724/5216

Otra casuística general en los campos dc.identifier	
1. Datos inconsistentes o incompletos	-
	Limaq
	index-oti2018 (NB)
	s.l
	1a.
	EST
	sss
	Llontop, D. (2017). Dos teorías informacionales de la conciencia. Persona, (020), 141-152. doi: http://dx.doi.org/10.26439/persona2017.n20.1744
2. Datos que no corresponden al elemento	Lima
	Perú
	2014
	[casos clínicos]. Lima: Universidad de San Martín de Porres, Facultad de Odontología;
3. Elemento vacío	

Metadato:	dc.language
Descripción:	Registra el lenguaje de la publicación, usa los códigos de idiomas reconocidos en el estándar ISO-639-3 (3 letras)
Ejemplo:	spa
Casuística:	
1. No cumple con las directrices	Español
	es
	en
2. Datos inconsistentes o incompletos	other
	de
	pt
	ing
	en_US
	T_MED.HUMA_2126
3. Errores tipográficos / ortográficos (mayúsculas / minúsculas)	Spa
	SPA
	Eng
4. Datos que no corresponden al elemento	Universidad Nacional Mayor de San Marcos. Programa Cybertesis PERÚ
	Indecopi

Metadato:	dc.relation
Descripción:	Enlaza una publicación con datos que se encuentran accesibles en un repositorio de datos institucional
Ejemplo:	info:pe-repo/semantics/dataset
Casuística:	
1. Datos que no corresponden al elemento	SUNEDU
	Acrobat Reader
	2017
	Tesis
	Ministerio de Transportes y Comunicaciones
	http://creativecommons.org/licenses/by/4.0/
	Gaceta Constitucional: jurisprudencia de observancia obligatoria para abogados y jueces
	2015-09-18

2. Datos inconsistentes o incompletos	Dato
	1
	-
	Número;1
	Vacio
	TBMC/2017;
	<iframe width="854" height="480" src="https://www.youtube.com/embed/a9KOveUYBJE" frameborder="0" gesture="media" allowfullscreen></iframe>

Metadato:	dc.publisher
Descripción:	Consignar la entidad responsable de la publicación del contenido del recurso
Ejemplo:	Universidad Nacional de Ingeniería
Casuística:	
1. Datos que no corresponden al elemento	Perú
	Lima
	Trujillo
	Economía
	THÈMIS-Revista de Derecho
	Magíster en Ingeniería de Sistemas-Mención: Gerencia en Tecnologías de Información y Comunicación
	Calvo, E.
2. Datos no normalizados	Universidad de San Martín de Porres USMP
	Instituto de Defensa Legal. Pontificia Universidad Católica del Perú
	Facultad de Ciencias Sociales y Jurídicas de la Universidad de Tarapacá
	Escuela de Posgrado. Universidad Nacional de Educación Enrique Guzmán y Valle
3. Datos inconsistentes o incompletos	Lima:
	s.l
	Uso de Bases de datos de Publicaciones Científicas
	Cuzco : Imprenta Republicana
	s.n
	u
	s. e.
L.F.	

	?b?ne
	dc.publisher.department
	Lima: , junio 2011. 61 p. Fotos.
4. Errores tipográficos / ortográficos	Universida Católica de Santa María
	Associaç,ã o Brasileira de Psiquiatria
	: Proyecto Piloto Rioplus, PDRS/GTZ
	Instituto Geofísico del Perú }
	Universidad Nacional Jose Faustino Sanchez Carrion (sin tildes)
	Pontificia Universidad Católica de Valparaiso (sin tildes)
	UNIVERSIDAD DE HUÁNUCO (mayúsculas / minúsculas)
5. Duplicados	Biomed Central Ltd BioMed Central Ltd.
	Universidad Nacional Mayor de San Marcos. Programa Cybertesis.PERÚ Universidad Nacional Mayor de San Marcos. Programa Cybertesis PERÚ
	Universidad Nacional Amazónica de Madre de Dios- UNAMAD Universidad Nacional Amazónica de Madre de Dios -UNAMAD
	Siglo XXI Editores Siglo XXI

Metadato:	dc.rights
Descripción:	Indica el tipo de acceso a un documento: acceso abierto, restringido, embargo o restringido en su totalidad
Ejemplo:	info:eu-repo/semantics/openAccess
Casuística:	
1. No cumple con las directrices	Restringido
	Tesis con Acceso Restringido
2. Datos que no corresponden al elemento	Luis Castillo Córdova
	Universidad de Piura
3. Errores tipográficos / ortográficos	info:eu- repo/semantics/openAccess (espacio de más luego del guión)
	info:eu-repo/semantics/RestrictedAccess (mayúscula en la R)
	info:eu-repo/semantics/restrictedAcces (falta una "s" al final)
	info:eu-repo/semantics/openAccess Licencia:

Metadato:	dc.rights.uri
Descripción:	Se indica el tipo de permisos que se otorga para el uso de la publicación de acceso abierto
Ejemplo:	https://creativecommons.org/licenses/by/4.0/
Casuística:	
1. No cumple con las directrices	Atribución 2.5 Perú
	Artículo en acceso abierto
	Creative Commons Atribución-NoComercial-SinDerivadas 2.5 Perú
	Copyright (c) 2017 Revista Enfermeria Herediana (casos de repositorio de revistas)
	Abierto
	Derechos de autor 2017 Pedro RUÍZ-CUBILLAS, Charles B. DAVEY
	© Universidad San Ignacio de Loyola
2. Datos que no corresponden al elemento	Universidad Continental
	http://www.scielo.org.pe/scielo.php?script=sci_arttext&pid=S2077-99172015000400010
	http://tesis.unsm.edu.pe/jspui/handle/11458/217
	http://repositorio.ana.gob.pe/handle/ANA
3. Datos inconsistentes o incompletos	nfo:eu-repo/semantics/openAccess
	creativecommons.org
	eu-repo/semantics/closedAccess
	An error occurred getting the license - uri.
	odc-by
	T_descriptor
	odc-odbl
	i
	licencia_001
<img alt="Licencia Creative Commons" style="border-width:0" ...	
4. Errores tipográficos	" https://creativecommons.org/licenses/by-nc-nd/2.5/pe/ " (uso de comillas)
	: http://creativecommons.org/licenses/by-nc-nd/2.5/pe/ (dos puntos al inicio de más)
	<info:eu-repo/semantics/openAccess (caracter al inicio de más)
	tps://creativecommons.org/licenses/by-nc-nd/4.0/deed.es (falta la "h" al inicio)
	https://i.creativecommons.org/l/by-nc/4.0/88x31.png

Metadato:	dc.source
Descripción:	Registra el nombre completo de la institución
Ejemplo:	Instituto del Mar del Perú
Casística:	
1. Datos inconsistentes o incompletos	INDECOPI-Institucional Repositorio INDECOPI
	Facultad de Ingeniería
	10.21895/incres.2013.v4n2
	\\STG-POLIMEDIA\Polimedia\Storage Polimedia\ re
	U
	71-86
2. Datos que no corresponden al elemento	1609-9419
	1025-5583
	https://pirhua.udep.edu.pe
	Revista Peruana de Medicina Experimental y Salud Pública; 2013 Vol 30 (2) <i>(nombres de la revista que provienen de OJS)</i>
	https://creativecommons.org/licenses/by-nc-nd/4.0/ application/pdf
3. Duplicados	Repositorio Institucional USIL Repositorio Institucional de USIL
	Universidad Privada Antenor Orrego - UPAOO Universidad Privada Antenor Orrego - UPAOAO Universidad Privada Antenor Orrego - Piura Universidad Privada Antenor Orrego - UPAO
	Universidad Nacional de San Martín. Universidad Nacional de San Martín-UNSM-T Universidad Nacional de San Martín-Tarapoto.
	Universidad Privada Antenor Orregospa Universidad Privada Antenor Orrego - UPAOO Universidad Privada Antenor Orrego - UPAOAO
	Universidad Peruana de Ciencias <i>aplicadas</i> (UPC) Universidad Nacional de la Amazonia <i>Peuana</i> Universidad Nacional Amazónica de <i>madre</i> de Dios Repositorio <i>Institucional</i> - ULADECH Repositorio <i>institucional</i> - UNITRU REPOSITORIO ACADÉMICO USMP
4. Errores tipográficos / ortográficos	Universidad Privada Antenor Orregospa Universidad Privada Antenor Orrego - UPAOO Universidad Privada Antenor Orrego - UPAOAO
	Universidad Peruana de Ciencias <i>aplicadas</i> (UPC)
	Universidad Nacional de la Amazonia <i>Peuana</i>
	Universidad Nacional Amazónica de <i>madre</i> de Dios
	Repositorio <i>Institucional</i> - ULADECH Repositorio <i>institucional</i> - UNITRU
	REPOSITORIO ACADÉMICO USMP

Metadato:	dc.subject
Descripción:	Consigna las palabras clave o descriptores
Ejemplo:	Psicología Ingeniería del transporte Comunicación – Aspectos sociales
Casuística:	
1. Elemento vacío	
2. Dos o más datos en el mismo campo	Ciencias empresariales y económicas / Finanzas
	Humanidades / Literatura
	Odontología, Cirugía Oral y Medicina Oral
	zoología; biodiversidad
	autoeficacia, estrés prácticas, profesionales, estudiantes, universidad
3. Errores tipográficos / ortográficos	derecho (primera letra mayúscula)
	Mercado Integrado Latinoamericano-Estadísticas (mayúsculas / minúsculas)
	Salud Pública (mayúsculas / minúsculas)
	MEDICINA (mayúsculas / minúsculas)
	Peru (sin tilde)
	linguística (sin diéresis ni tilde)
	Huancayo. (punto al final)
	Tingo marí a (espacio de más)
	– Adolescente
	, trabajo colaborativo
Sánchez, Luis Alberto 1900-1994 (falta coma antes del cronológico)	
EsSaludJaén (palabras juntas)	
4. Uso de descriptores en inglés	Law
	Peruvian writers
5. Datos no normalizados	Universidad
	Otras Ciencias Agrícolas
	Otras ingenierías y Tecnologías
	Lima (dpto)
	Hotel
	Hospital
	5S

	Kant
	Estudiantes de escuelas secundarias - Perú - Lima (Lima) - Psicología
	Francisco Pizarro
	Vargas Llosa, Jorge Mario Pedro
	Peruano (Lima, Perú)
	Enfermeros/normas
6. Datos inconsistentes o incompletos	Subject 1
	Día treinta
	Tesina
	con
	Partir
	Tercer
	Pedro
	--
	influyen
	plan
	to
	sobre
	a
	3240
	Señor
	Msnm
	...
	no
	No.
	■ ARTESANÍA TEXTIL: Se conoce como artesanía textil a todos aquellos accesorios que son hechos a mano por un artesano, quien durante su labor ...
	r e g r e s i ó n
	Research Subject Categories::SOCIAL SCIENCES::Social sciences::Psychology
	LATEX DE PAPAYA AL 2%,4%,6%,8%
YBa2Cu3O7	
7. Duplicados	Análisis estructural
	Análisis estructural (Ingeniería)
	Análisis financiero
	Analisis financiero
	Análisis Financiero
	Lima metropolitana

	Lima (Perú : Area Metropolitana)
	TIC TICs Tecnología de la información y comunicación Tecnologías de la Información y Comunicación Tecnologías de información y comunicaciones
	Helados Helado
	Maca Maca (Planta)
	Mamas - Cáncer Cáncer de Mamas
	5S Metodología 5s
	Rendimiento académico - Perú - Lima (Area Metropolitana) Rendimiento académico - Perú - Lima (Área Metropolitana)
	Matemáticas-Estudio y enseñanza primaria Matemáticas--Estudio y enseñanza (Primaria)
	Anillos (Álgebra) Anillos (Algebra)
	Texto autoinstructivo Textos autoinstructivos
8. Datos que no corresponden al elemento	2014
	Facultad de Derecho
	El juego
	Una Mirada De La Vivencia Del Rol Materno
	Comisión de Protección al Conusmidor N° 3
	Asesor: Jorge León Bejarano
9. Tildes no reconocidas	Deshidrataci?n
	c?ncer de cuello uterino
	Poblaci?n

Metadato:	dc.title
Descripción:	Registrar el nombre original, el orden y la ortografía del título del recurso
Ejemplo:	Una metodología para el desarrollo de habilidades de localización y valoración de la información en los estudiantes de la Universidad Peruana de Ciencias Aplicadas (UPC)
Casuística:	
1. Elemento vacío	
2. Tildes no reconocidas	"Morfometr?a de especies de Aphididae (Hemiptera) procedentes de Cieneguilla (Lima-Per?)."
	Dise?o e implementaci?n de un robot humanoide asistencial controlado por computadora para aplicaciones en pacientes parap?jicos
3. Datos inconsistentes o incompletos	D ifu s i ó n de cloruro d e sodio y á cido ac ét ico En e l músculo de a nchov e t a p e ru an a (Engraulis Ringens) dur an te e l proceso de mari na do
	Embalse Bramadero ? Proyecto De Abastecimiento De Agua A La Mina ?La Zanja? Y Labores Agropecuarias
	-
4. Errores tipográficos / ortográficos	Insercion de un equipamiento educativo artistico en el centro historico de Arequipa Escuela de Musica en el EX-COLEGIO San Francisco y Serafico (sin tildes)
	Disen~ o de un Puente Tipo Compuesto con Vigas de Acero y Losa de Concreto Ubicado sobre la Quebrada Japolccaya, Distrito de Maca, Caylloma, Arequipa
	La novela, ?obra de mala fe?
	?Efecto del cloruro de mercurio sobre la regeneraci?n celular de la planaria de agua dulce Girardia festae (Borelli, 1898) (Tricladida: Dugesiiidae)?
	"plan estratégico de marketing para consolidar el posicionamiento de la marca prolimsa en los consumidores del mercado de productos de limpieza en la provincia de trujillo"
	Sílabo de Comportamiento del Consumidor (solo nombres propios con mayúscula)
	Enfermedades inmunoprevenibles: Tos ferina (subtítulo sin mayúscula)
	Identifying Musculoskeletal Discomfort in a Manufacturing Plant in Los Olivos, Lima–Peru (solo nombres propios con mayúscula)
5. Posibles duplicados	conductas asertivas y trabajo en equipo en niños de 5 años de la I.E.P. Cuna de Bendición Horacio Zeballos, Ate, 2017 (Primera palabra en mayúscula)
	Estudio Prospectivo del Mercado del Esparrago en el Norte Chico del Perú 2007 - 2017
	EFICACIA DE UN SISTEMA DE ASPIRACION CERREDO COMPARADO CON EL SISTEMA DE ASPIRACIÓN ABIERTO EN LOS CAMBIOS HEMODINAMICOS DE PACIENTES CON VENTILACIÓN MECÁNICA
	El talento humano y el rendimiento laboral del personal de la administración de la UNJFSC, 2014 / El talento humano y el rendimiento laboral del personal de la administración de la universidad nacional José Faustino Sánchez Carrión

Metadato:	dc.type
Descripción:	Registrar el tipo de publicación de acuerdo a valores establecidos según las directrices DRIVER
Ejemplo:	info:eu-repo/semantics/article info:eu-repo/semantics/masterThesis
Casuística:	
1. Elemento vacío	
2. Datos inconsistentes o incompletos	info:eu-repo/semantics/publishedVersion
	info:eu-repo/semantics/other
	info:eu-repo/semantics/Sílabus
	info:eu-repo/semantics/reporte
	info:pe-repo/semantics/jurisprudencia
	info:eu-repo/semantics/presentation
	info:eu-repo/semantics/Thesis
	NO_PUBLICACION
	info:eu-repo/semantics/boletín
	www
	deductivo-inductivo
	Inductivo- Deductivo y Analítico
	infor:eu-repo/semantics/conferencepaper
	T_M_Manejo Forestal
	método participativo
	método científico empírico - analítico
	info:eu-repo/semantics
	nfo:eu-repo/semantics/contributionToPeriodical
eu-repo/semantics/report	
3. No cumple con las directrices	Artículo
	Artículo revisado por pares
	Nota de prensa
	Tesis de Licenciatura
	Documentos Internos
	Tesis de Maestría
	Tesis
	Libro / Monografía
	T_en proceso
	Trabajo Académico de Segunda Especialidad

4. Datos que no corresponden al elemento	Se utilizó el medio de cultivo basal de Murashige & Skoog (MS) suplementado con vitaminas y hormonas. Se aplicó un diseño completamente al azar...
	image/jpeg

Metadato:	dc.description.uri
Descripción:	Tipo de obra conducente a grado académico
Ejemplo:	Tesis Tesis de segunda especialidad Trabajo de investigación Trabajo de suficiencia profesional Trabajo académico
Casuística:	
1. Errores ortográficos / tipográficos	Trabajo academico (sin tilde)
	Tesis de Segunda Especialidad (mayúsculas / minúsculas)
	Trabado de investigacion (error tipográfico y falta tilde)
	Trabajo de Suficiencia profesional (mayúsculas / minúsculas)
2. No cumple con las directrices	Tesis para maestría
	Tesis para título profesional
	Tesis para Magister
	Tesis / Trabajo de investigación
	Informe de tesis