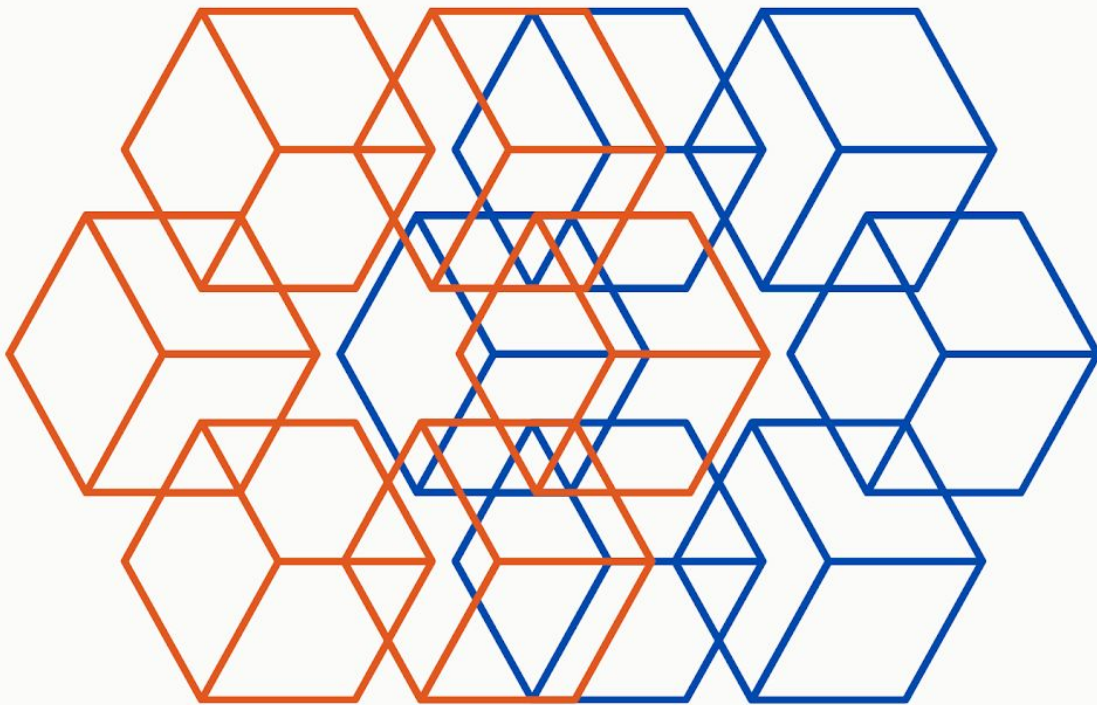




Padrões de interoperabilidade para repositórios de dados de pesquisa

Instituto Brasileiro de Informação em Ciência e Tecnologia | Ibict



4º Plano de Ação Nacional em Governo Aberto Brasil

Compromisso 3. Estabelecer mecanismos de governança de dados científicos para o avanço da Ciência Aberta no Brasil

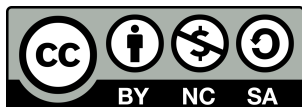
Marco 8. Proposição de padrões de interoperabilidade para repositórios de dados de pesquisa

PADRÕES DE INTEROPERABILIDADE PARA REPOSITÓRIOS DE DADOS DE PESQUISA

Brasília

2020





Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT)
Comissão Nacional de Energia Nuclear (CNEN)
Rede Nacional de Ensino e Pesquisa (RNP)

Elaboração:

Lucas Nóbrega Paganine (Ibict)

Washington Luís R. de Carvalho Segundo (Ibict)

João Luiz R. Moreira (University of Twente, Enschede - Holanda)

Luiz F. Sayão (CNEN)

Vanderlino Coelho Barreto Neto (CNEN)

Leandro Neumann Ciuffo (RNP)

Carolina Howard Felicíssimo (RNP)

Gustavo Neves Dias (RNP)

Arte da capa:

Fhillipe de Freitas Campos

Revisão:

Luciana dos Santos Nahuz

Dados Internacionais de Catalogação na Publicação (CIP)

P124

Padrões de interoperabilidade para repositórios de dados de pesquisa / Lucas Nóbrega Paganine...[et al] . -- Brasília, DF: IBICT , RNP; Rio de Janeiro: CNEN; Enschede: University of Twente, 2020.

31p. : il. ; 28 cm.

1. Repositório de dados. 2. Interoperabilidade. 3. Dados de pesquisa. 4. Gestão de dados de pesquisa. 5. Curadoria digital. 6. Ciência aberta. I. Paganine, Lucas Nóbrega. II. Comissão Nacional de Energia Nuclear. III. Rede Nacional de Ensino e Pesquisa. IV. University of Twente. VI. Título.

CDU 004.62

CDD 004

Ficha catalográfica elaborada pela Bibliotecária Luciana Nahuz CRB1 nº 1665

Setor de Autarquias Sul (SAUS) - Quadra 05 Lote 06
Bloco H - 5º Andar
Cep: 70070-912 - Brasília, DF
Telefones: 55 (61) 3217-6360/ 55 (61) 3217-6350
<http://www.ibict.br>

Rua Lauro Muller, 455 - 4º Andar - Botafogo
Cep: 22290-160 - Rio de Janeiro, RJ
Telefones: 55 (21) 2275-0321
Fax: 55 (21) 2275-3590
<https://ibict.br/pesquisa-e-pos-graduacao/pos-graduacao-em-ciencia-da-informacao>
<http://www.ppgi.ufrj.br/>

Esta publicação está licenciada com a Licença CC BY
Creative Commons Atribuição 4.0 Internacional.

Este documento é resultado do Marco 8 ‘Definição de padrões de interoperabilidade para repositórios de dados de pesquisa’ do Compromisso 3 ‘Estabelecer mecanismos de governança de dados científicos para o avanço da Ciência Aberta no Brasil’ do 4º Plano de Ação Nacional em Governo Aberto.

Instituições parceiras do Marco 8

Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict)

Comissão Nacional de Energia Nuclear (CNEN)

Rede Nacional de Ensino e Pesquisa (RNP)

University of Twente

CITAR

PAGANINE, Lucas. *et al.* **Padrões de interoperabilidade para repositórios de dados de pesquisa.** Brasília, DF: IBICT, 2020. 44 p.

Apresentação

A Parceria para Governo Aberto (do inglês, Open Government Partnership - OGP) é uma iniciativa internacional composta, atualmente, por mais de 79 nações. O Brasil é co-fundador dessa parceria que foi criada, em setembro de 2011, com o objetivo de difundir e incentivar globalmente práticas governamentais relacionadas à transparência dos governos, ao acesso à informação pública e à participação social.

As ações relativas à OGP são operacionalizadas, em cada país, por meio de um Plano de Ação Nacional, onde são especificados os compromissos em Governo Aberto assumidos pelo país e as estratégias e atividades para concretizá-los. Os planos de ação possuem duração de até dois anos e, ao longo desse período, os Governos precisam publicar, anualmente, um relatório de autoavaliação sobre a execução dos compromissos assumidos.

O Brasil publicou, em 29 de outubro de 2018, o seu 4º Plano de Ação Nacional composto por 11 compromissos, os quais foram co-criados com o envolvimento de 105 pessoas, representantes de 88 instituições, sendo 39 organizações da sociedade civil, 39 órgãos da Administração Pública Federal e 10 órgãos das Administrações Públicas Estaduais e Municipais.

O Compromisso 3 do 4º Plano de Ação Nacional do Brasil tem o objetivo de “Estabelecer mecanismos de governança de dados científicos para o avanço da Ciência Aberta no Brasil”. Esse compromisso, conhecido como Compromisso pela Ciência Aberta, é coordenado pela Empresa Brasileira de Pesquisa Agropecuária (Embrapa) e conta com a parceria dos seguintes órgãos governamentais e da sociedade civil: Ministério da Ciência, Tecnologia e Inovações (MCTI), Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), Fundação Oswaldo Cruz (Fiocruz), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Rede Nacional de Ensino e Pesquisa (RNP), Open Knowledge Brasil (OKBR), Comissão Nacional de Energia Nuclear (CNEN), Arquivo Nacional, Instituto de Pesquisa do Jardim Botânico do Rio de Janeiro (JBRJ) e Universidade de Twente.

A execução desse compromisso conta com ações desenvolvidas em nove marcos, sendo que o Marco 8 teve o objetivo de “definir padrões de interoperabilidade para repositórios de dados de pesquisa”, o qual esteve sob a responsabilidade do Ibict em parceria com o CNEN e a RNP, também vale destacar a colaboração com a Universidade de Twente e cujo resultado é este documento.

As recomendações aqui propostas aplicam-se a todas as instituições de pesquisa que desenvolvem pesquisas financiadas com recursos públicos e tem o objetivo de orientar e promover a abertura e interoperabilidade dos dados científicos em apoio ao movimento da Ciência Aberta no Brasil. Os padrões de interoperabilidade aqui definidos visam guiar as instituições de pesquisa a construir ou aprimorar seus repositórios de dados científicos.

SUMÁRIO

1. INTRODUÇÃO.....	1
2. OBJETIVOS.....	3
3. METODOLOGIA.....	3
4. DIRETRIZES GERAIS PARA REPOSITÓRIOS DE DADOS DE PESQUISA.....	4
5. DIRETRIZES PARA REPOSITÓRIOS DE DADOS DE PESQUISA POR ÁREA DO CONHECIMENTO.....	14
5.1 Biologia	14
5.2 Ciências sociais	15
5.3 Agricultura.....	16
6. CONCLUSÃO	16
REFERÊNCIAS	17

1 Introdução

A partir Movimento Acesso Aberto surgiram novas necessidades e perspectivas que demandam a abertura de todo o processo científico para a sociedade de forma geral, surgindo então a Ciência Aberta (um termo guarda-chuva que abrange as práticas de abertura seguindo os princípios da transparência e colaboração). Um dos temas tratados na Ciência Aberta é o compartilhamento de dados de pesquisa.

Os repositórios de dados científicos são apontados como ferramenta para organização, tratamento, preservação e disseminação de dados científicos. Porém, com a vasta quantidade de modelos de descrição de conjuntos de dados, que são gerados nas mais variadas áreas do conhecimento, surge a necessidade de um estudo sobre os padrões existentes, com vistas a se obter um conjunto de descrição central, assim como o tratamento de especificidades via pequenas extensões deste conjunto, para que se viabilize a interoperabilidade dos repositórios de dados científicos de forma geral em diversos níveis.

O projeto Digital Repository Infrastructure Vision for European Research (DRIVER) em 2007 publicou as Guidelines for content providers: Exposing textual resources with OAI-PMH, que contém recomendações que permitem interoperabilidade entre repositórios. E desde então essas recomendações foram evoluindo e estão atualmente em sua versão 4.0 sendo chamadas de OpenAIRE Guidelines for Literature Repository Managers v4. A partir de 2013 a Comissão Européia (CE) com a publicação do Guidelines on Open Access to Scientific Publications and Research Data está de acordo com os requerimentos do Horizon 2020 abrangendo desde repositórios de literatura à repositórios de dados e sistemas CRIS.

A diretriz é dividida em três partes: Uma introdução onde ela é apresentada; Uso do OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting), onde são apresentadas algumas explicações sobre o uso do protocolo OAI-PMH e; por fim uma visão de um perfil de aplicação. A aplicação é baseada em quatro padrões de metadados: Dublin Core; Dublin Core qualificado; Datacite e; Oaire (um padrão de metadados próprio desenvolvido pela OpenAIRE), além de serem utilizados diversos vocabulários controlados.

Ressalta-se que mesmo com o extensivo papel do protocolo OAI-PMH na interoperabilidade sintática o paradigma atual tecnológico está se desenvolvendo para cobrar mais. Assim outras iniciativas se destacam como os W3C: PROV-O e DCAT 2.0, que incorpora classes para descrever os serviços de dados visando a implementação FAIR.

Nas diretrizes OpenAIRE V4. foram definidos quatro níveis de obrigatoriedade para metadados: Mandatório (M); Mandatório se aplicável (MA); Recomendado (R) e; Opcional (O). Seguem as explicações de cada um deles.

- Mandatório (M): Qualquer campo com esse nível de obrigatoriedade deve ser preenchido independentemente da situação, caso isso não ocorra estará em não conformidade com as diretrizes. (ex: Título)
- Mandatório se aplicável (MA): Quando o campo apresenta esse nível de obrigatoriedade ele deve ser preenchido caso a informação em questão esteja disponível. (ex: Contribuidor)

- Recomendado (R): O campo que apresentar esse nível de obrigatoriedade não necessita ser preenchido porém é benéfico que o faça. (ex: Identificador Alternativo)
- Opcional (O): Este nível de obrigatoriedade significa que seu preenchimento é de caráter completamente discricionário, caso se julgue que seu preenchimento enriquecerá o registro pode-se preenchê-lo ou não. (ex: Tamanho)

É apresentado no perfil de aplicação uma tabela com os campos, elementos e vocabulários utilizados. São 32 campos no total, com seis mandatórios, oito mandatórios se aplicáveis, quinze recomendados e três opcionais.

O FAIR Data Point (FDP) é uma aplicação web independente para repositórios de dados que fornece metadados e outros objetos digitais FAIR. O FDP usa uma API REST para criar, armazenar e veicular metadados. O FDP é um software que permite que os proprietários dos dados exponham os conjuntos de dados de maneira FAIR além de permitir que os usuários descubram metadados sobre os conjuntos de dados oferecidos e, de acordo com as condições da licença, acessar os dados. Uma coleção de dados em FDP visa solucionar problemas de interoperabilidade promovendo a encontrabilidade, acessibilidade, interoperabilidade e reutilização.

O software FDP é um aplicativo da web independente. Porém sua funcionalidade e comportamento podem ser incorporados em quaisquer outros aplicativos fornecendo dados FAIR e acessibilidade aos metadados e dados. Por exemplo, um repositório de dados existente pode optar por implementar os metadados do FDP, comportando-se dessa maneira também como um FDP.

O padrão de metadados FDP é baseado no Esquema re3data e DCAT_x, dividido em 4 camadas com 2 opções de nível de obrigatoriedade autoexplicativas, sendo elas nível obrigatório (com 10 campos) ou nível opcional (com 12 campos). As 3 camadas de metadados se referem cada uma a um tipo de objeto digital complexo que pode ser descrito por metadados em um repositório de dados. São elas:

- Camada de metadados do repositório: contém informações relativas ao repositório como um todo
- Camada de metadados do catálogo: contém informações relativas à coleção de conjuntos de dados
- Camada de metadados do conjunto de dados: contém informações relativas ao Conjunto de dados, ou dataset
- Camada de metadados de distribuição: contém informações relativas aos arquivos individuais contidos nos conjuntos de dados

Vale destacar ainda, outras iniciativas da CE como: a *European Open Science Cloud* (EOSC) de 2015 que visa implementar uma infraestrutura de práticas de ciência aberta; a promoção dos princípios FAIR desde 2014 com a publicação de diversos guias e o

relatório *Turning FAIR into reality: Final Report and Action Plan on FAIR Data* em 2018; as diretrizes PSI (*Public Sector Information*) de 2003, que em 2019 se transformaram nas Diretrizes para dados abertos e reuso de informações do setor público (*Directive on open data and the re-use of public sector information*); o programa e publicação *Open Innovation, Open Science and Open to the World: a vision for Europe* de 2016; por último reconhece-se a atuação da CE no Research Data Alliance (RDA) que em 2018 lançou a quarta fase de seu programa na plataforma, garantindo a participação de *stakeholders* a partir de projetos e financiamentos.

2 Objetivos

Desenvolver e aplicar um conjunto mínimo de descrição para dados de pesquisa, realizando-se extensões para domínios específicos dos conhecimentos, com base nos padrões e diretrizes internacionais existentes.

3 Metodologia

Realizou-se um estudo qualitativo com pesquisa exploratória, utilizando-se respectivamente as definições de Creswell (2007) e Gil (2011, p. 28). Foram realizadas observações dos padrões adotados por repositórios temáticos e institucionais (multitemáticos), tomando-se como base os registros de repositórios de dados de pesquisa encontrados no re3data.org (Registry of Research Data Repositories), em consultas realizadas entre outubro de 2019 e março de 2020.

O estudo realizado foi desenvolvido em 2 etapas, a primeira para o estabelecimento de um padrão de metadados geral multitemático, e a segunda para o desenvolvimento de extensões por área do conhecimento. Entre as diretrizes existentes para repositórios de dados multitemáticos foram elegidas as [OpenAIRE Guidelines for Data Archives](#), e no que se refere ao paralelo e conexão com os repositórios de publicações científicas, foram também tomadas como base as [OpenAIRE Guidelines for Literature Repositories](#). As diretrizes *OpenAIRE* são uma evolução das diretrizes *Digital Repository Infrastructure Vision for European Research* (DRIVER) para repositórios de publicações científicas e são amplamente utilizadas no contexto internacional. No entanto, estudos sobre interoperabilidade no nível semântico apontam a necessidade de extensões das diretrizes *OpenAIRE* para a criação de repositórios que atendam aos princípios FAIR (WILKINSON; et al, 2016). Uma extensão inicial foi realizada para a adaptação das diretrizes *OpenAIRE* ao padrão exigido em (SANTOS; et. al, 2016), na criação do chamado *Fair Data Point* (FDP).

Para o desenvolvimento das extensões por área do conhecimento (Anexo III), foram tomadas como base a tabela de áreas do conhecimento CNPq (CNPQ, s.d.), a tabela de áreas exibida no manual Frascati (OECD, 2015) e a divisão de áreas utilizada pelo *Data Curation Centre* (DCC) da *Research Data Alliance* (RDA) (DFG, 2020).

Por fim foi realizada análise sobre os padrões utilizados nos repositórios observados, identificando-se que extensões seriam necessárias às diretrizes OpenAIRE com a extensão pelo padrão de descrição adotado no FDP, para 3 áreas selecionadas: Biologia, Agricultura e Ciências Sociais.

4 Diretrizes gerais para repositórios de dados de pesquisa

A partir da comparação (Anexo II) entre os campos OAIRE (Anexo I) e os FDP chegou-se ao seguinte núcleo de metadados:

Campos obrigatórios

- **Título (Title)**

Metadado: datacite:title e dct:title

Definição: Nome ou título pelo qual o conjunto de dados será conhecido

Exemplo: National Institute for Environmental Studies and Center for Climate System Research Japan: A survey

Exemplo de aplicação:

```
<datacite:title xml:lang="en-US">
  National Institute for Environmental Studies and Center
  for Climate System Research Japan
</datacite:title>
<datacite:title xml:lang="en-US" titleType="Subtitle">A
survey</datacite:title>
```

e

```
dcterms:title "National Institute for Environmental Studies
and Center for Climate System Research Japan: A survey";
```

Vocabulário controlado: [DC terms](#) e [title type](#)

Justificativa: Obrigatório no FDP e OAIRE

- **Criador (Creator)**

Metadado: datacite:creator e dct:publisher

Definição: Autor principal do recurso (pode ser correlato ao usuário e automatizado), vale destacar a importância de se utilizar aqui os atributos datacite:affiliation com a instituição de origem do pesquisador e datacite:nameIdentifier com um identificador

persistente único de autor, preferencialmente o IDLattes mas possivelmente o ORCID

Exemplo: Evans, R.J.

Exemplo de aplicação:

```
<datacite:creators>
  <datacite:creator>
    <datacite:creatorName>Evans, R.J.</datacite:creatorName>
    <datacite:affiliation>Institute of Science and
Technology</datacite:affiliation>
    <datacite:nameIdentifier nameIdentifierScheme="IDLattes"
      schemeURI="http://lattes.cnpq.br/">
      7718494581206998
    </datacite:nameIdentifier>
  </datacite:creator>
</datacite:creators>
```

e

```
dcterms:publisher "Evans, R.J.";
```

Vocabulário controlado: [name type](#)

Justificativa: Obrigatório no FDP e OAIRE

- **Data (Date)**

Metadado: datacite:date e dct:issued

Definição: Data de depósito do conjunto de dados (pode ser automatizado pelo sistema)

Exemplo: 2000-12-25

Exemplo de aplicação:

```
<datacite:date dateType="Available">2000-12-25</datacite:date>
```

e

```
dcterms:issued "2000-12-25";
```

Vocabulário controlado: [date type](#)

Justificativa: Obrigatório no OAIRE

- **Tipo (Type)**

Metadado: oaire:resourceType e rdf:type

Definição: O gênero do recurso, tipo de dado científico

Exemplo: Text

Exemplo de aplicação:

```
<oaire:resourceType resourceTypeGeneral="dataset"
uhttp://purl.org/coar/resource_type/c_ddb1">text</oaire:resourceType>
```

Vocabulário controlado: [re3data](#) (pág. 19)

Justificativa: Obrigatório no FDP e OAIRE

- **Identificador (Identifier)**

Metadado: datacite:identifier

Definição: Identificador único do conjunto de dados (automatizado pelo sistema)

Exemplo: <http://hdl.handle.net/1234/5628> ou <https://doi.org/10.1109/5.771073>

Exemplo de aplicação:

```
<datacite:identifier
identifierType="Handle">http://hdl.handle.net/1234/5628</datacite:identifier>
```

e

```
dcterms:identifier "http://hdl.handle.net/1234/5628" .
```

ou

```
<datacite:identifier
identifierType="DOI">https://doi.org/10.1109/5.771073</datacite:identifier>
```

e

```
dcterms:identifier "http://hdl.handle.net/1234/5628" .
```

Vocabulário controlado: [identifier type](#)

Justificativa: Obrigatório no OAIRE

- **Direitos de acesso (Access Rights)**

Metadado: datacite:rights e dct:accessRights

Definição: Nível de abertura do conjunto de dados (dependendo do escopo do repositório)

pode ser automatizado pelo sistema)

Exemplo: Open access

Exemplo de aplicação:

```
<datacite:rights
rightsURI="http://purl.org/coar/access_right/c_abf2">open
access</datacite:rights>
```

e

```
dcterms:accessRights <http://136.243.4.200:8087/fdp/accessRights>;
```

Vocabulário controlado: [COAR Access Right Vocabulary](#)

Justificativa: Obrigatório no OAIRE

- **Assunto (Subject)**

Metadado: datacite:subject e dcat:theme

Definição: Palavras ou termos que descrevem o conteúdo do conjunto de dados

Exemplo: Earth sciences and geology

Exemplo de aplicação:

```
<datacite:subjects>
  <datacite:subject>Earth sciences and geology</datacite:subject>
  <datacite:subject subjectScheme="DDC"
schemeURI="http://dewey.info/" valueURI="">
  551 Geology, hydrology, meteorology
</datacite:subject>
</datacite:subjects>
```

e

```
dcat:theme <http://dbpedia.org/resource/Text_mining>,
<http://semanticscience.org/resource/statistical-association> .
```

Vocabulário controlado: -

Justificativa: Obrigatório no FDP

- **Versão (Version)**

Metadado: datacite:version e dct:hasVersion

Definição: Versão numérica do conjunto de dados (pode ser automatizado pelo sistema)

Exemplo: 1.0

Exemplo de aplicação:

```
<oaire:version>1.0</oaire:version>
```

e

```
dcterms:hasVersion "1.0";
```

Vocabulário controlado: [COAR Version Vocabulary](#)

Justificativa: Obrigatório no FDP

- **Parte de (Part of)**

Metadado: dct:isPartOf

Definição: Link ou identificador da comunidade, dataverse, repositório ou qualquer outro grupo maior ao qual o conjunto de dados pertença (idealmente automatizado pelo sistema, utilização apenas possível ao se adotar um nível de descrição de metadados do repositório)

Exemplo: <http://136.243.4.200:8087/fdp/catalog/textmining>

Exemplo de aplicação:

```
dcterms:isPartOf
```

```
<http://136.243.4.200:8087/fdp/catalog/textmining>;
```

Vocabulário controlado: -

Justificativa: Obrigatório no FDP

- **Identificador dos metadados (Metadata Identifier)**

Metadado: fdp:metadataIdentifier e datacite:relatedIdentifier

Definição: Link ou identificador dos metadados associados ao conjunto de dados (idealmente automatizado pelo sistema, utilização apenas possível ao se adotar um nível de descrição da camada dos conjuntos de metadados)

Exemplo: http://purl.org/biosemantics-lumc/fdp/dataset/gene_disease_association

Exemplo de aplicação:

```
<datacite:relatedIdentifiers>
```

```
<datacite:relatedIdentifier
```

```
relatedIdentifierType="MetadataIdentifierURL"
```

```
relationType="IsMetadataFor">http://purl.org/biosemantics-lumc/fdp
/dataset/gene_disease_association</datacite:relatedIdentifier>
</datacite:relatedIdentifiers>
```

e

```
fdp:metadataIdentifier
<http://purl.org/biosemantics-lumc/fdp/dataset/gene_disease_associ
ation>;
```

Vocabulário controlado: [FDP Ontology](#)

Justificativa: Obrigatório no FDP

- **Data dos metadados (Metadata Date)**

Metadado: fdp:metadataIssued e datacite:date dateType="MetadataIssued"

Definição: Data na qual os metadados foram preenchidos (pode ser automatizado pelo sistema)

Exemplo: 2018-03-20

Exemplo de aplicação:

```
<datacite:date dateType="Accepted">2018-03-20</datacite:date>
```

e

```
fdp:metadataIssued "2018-03-20T10:30:18.662Z"^^xsd:dateTime;
```

Vocabulário controlado: -

Justificativa: Obrigatório no FDP

- **Data da modificação dos metadados (Metadata Modification Date)**

Metadado: fdp:metadataModified e datacite:date dateType="MetadataModified"

Definição: Data na qual os metadados foram alterados (pode ser automatizado pelo sistema)

Exemplo: 2018-03-20

Exemplo de aplicação:

```
fdp:metadataModified "2018-08-20T13:09:55"^^xsd:dateTime;
```

Vocabulário controlado: -

Justificativa: Obrigatório no FDP

- **Distribuição (Distribution)**

Metadado: dcat:distribution e datacite:relatedIdentifier

Definição: Link ou identificador dos arquivos contidos no conjunto de dados (pode ser automatizado pelo sistema)

Exemplo:

http://136.243.4.200:8087/fdp/distribution/gene_disease_association_nquads_gzip

Exemplo de aplicação:

```
<datacite:relatedIdentifiers>
  <datacite:relatedIdentifier
    relatedIdentifierType="DistributionIdentifierURL"
    relationType="IsPartOf">http://purl.org/biosemantics-lumc/fdp/data-set/gene\_disease\_association</datacite:relatedIdentifier>
</datacite:relatedIdentifiers>
```

e

```
dcat:distribution
<http://136.243.4.200:8087/fdp/distribution/gene\_disease\_association\_nquads\_gzip>;
```

Vocabulário controlado: [DCAT](#)

Justificativa: Obrigatório no FDP

Campos recomendados

- **Área do conhecimento CNPq**

Metadado: datacite:subject subjectScheme=CNPq

Definição: Designação de área do conhecimento de acordo com a tabela CNPq

Exemplo: Cosmologia

Exemplo de aplicação:


```
<datacite:subjects>
  <datacite:subject>Cosmologia</datacite:subject>
  <datacite:subject subjectScheme="CNPq"
schemeURI="http://lattes.cnpq.br/web/dgp/arvore-do-conhecimento>
1.00.00.00-3 Ciências Exatas e da Terra 1.04.00.00-1 Astronomia
1.04.04.04-0 Cosmologia
</datacite:subject>
</datacite:subjects>
```

Vocabulário controlado: [Tabela CNPq](#)

- **Área do conhecimento Internacional**

Metadado: datacite:subject subjectScheme=DFG

Definição: Designação de área do conhecimento de acordo com a tabela DFG (re3data/datacite)

Exemplo: Culturas antigas

Exemplo de aplicação:

```
<datacite:subjects>
  <datacite:subject>Earth sciences and geology</datacite:subject>
  <datacite:subject subjectScheme="DDC"
schemeURI="https://www.dfg.de/en/dfg_profile/statutory_bodies/revi
ew_boards/subject_areas/index.jsp>
  101 Ancient Cultures
</datacite:subject>
</datacite:subjects>
```

Vocabulário controlado: [DFG Classification](#)

- **Conjunto de dados ou publicação relacionado**

Metadado: datacite:relatedIdentifier

Definição: Identificador persistente único de um conjunto de dados que tenha alguma relação com o conjunto em questão

Exemplo: <http://hdl.handle.net/1234/5628>

Exemplo de aplicação:

```
<datacite:relatedIdentifiers>
```

```

<datacite:relatedIdentifier
relatedIdentifierType="Handle">http://hdl.handle.net/1234/5628</da
tacite:relatedIdentifier>
</datacite:relatedIdentifiers>

```

Vocabulário controlado: -

- **Financiador**

Metadado: oaire:fundingReferences

Definição: Nome da instituição financiadora

Exemplo: Comissão Europeia

Exemplo de aplicação:

```

<oaire:fundingReferences>
  <oaire:fundingReference>
    <oaire:funderName>European Commission</datacite:funderName>
    <oaire:funderIdentifier
funderIdentifierType="ROR">https://ror.org/00k4n6c32</oaire:funde
rIdentifier>
    <oaire:fundingStream>Horizon          2020          Framework
Programme</oaire:fundingStream>
    <oaire:awardNumber
awardURI="http://cordis.europa.eu/project/rcn/194062_en.html">643
410</oaire:awardNumber>
    <oaire:awardTitle>Open Access Infrastructure for Research in
Europe 2020</oaire:awardTitle>
  </oaire:fundingReference>
</oaire:fundingReferences>

```

ou

```

<oaire:fundingReferences>
  <oaire:fundingReference>
    <oaire:funderName>European Commission</datacite:funderName>
    <oaire:funderIdentifier funderIdentifierType="Crossref Funder
ID">http://doi.org/10.13039/100010661</oaire:funderIdentifier>
    <oaire:fundingStream>Horizon          2020          Framework
Programme</oaire:fundingStream>

```

```
<oaire:awardNumber
awardURI="http://cordis.europa.eu/project/rcn/194062_en.html">643
410</oaire:awardNumber>
  <oaire:awardTitle>Open Access Infrastructure for Research in
Europe 2020</oaire:awardTitle>
</oaire:fundingReference>
</oaire:fundingReferences>
```

Vocabulário controlado: [ROR](#)

- **Licença**

Metadado: oaire:licenseCondition

Definição: Licença de uso do conjunto de dados

Exemplo: CC-BY-NC

Exemplo de aplicação:

```
<oaire:licenseCondition
startDate="2019-02-01"
uri="http://creativecommons.org/licenses/by-nc/4.0/">Creative
Commons Attribution-NonCommercial</oaire:licenseCondition>
```

Vocabulário controlado: -

- **Palavra-chave**

Metadado: datacite:subject subjectScheme=keyword

Definição: Palavras ou termos que descrevem o conteúdo do conjunto de dados

Exemplo: Título

Exemplo de aplicação:

```
<datacite:subjects>
  <datacite:subject>Título</datacite:subject>
  <datacite:subject subjectScheme="keyword">
  </datacite:subject>
</datacite:subjects>
```

Vocabulário controlado: -

- **Chamado**

Metadado: datacite:description descriptionType=chamado

Definição: Número do chamado que gerou o conjunto de dados

Exemplo: 08/2019

Exemplo de aplicação:

```
<datacite:description>  
  <datacite:description>08/2019</datacite:description>  
  <datacite:description descriptionType="chamado">  
  </datacite:description>  
</datacite:description>
```

Vocabulário controlado: -

- Projeto

Metadado: datacite:description descriptionType=projeto

Definição: Nome do projeto que gerou o conjunto de dados

Exemplo: Descrição morfoagronômica e diversidade de acessos

Exemplo de aplicação:

```
<datacite:description>  
  <datacite:description>Descrição morfoagronômica e diversidade de  
acessos</datacite:description>  
  <datacite:description descriptionType="projeto">  
  </datacite:description>  
</datacite:description>
```

Vocabulário controlado: -

5 Diretrizes para repositórios de dados de pesquisa por área do conhecimento

Biologia

Exatamente por se tratar de um área de posicionamento complexo a Biologia foi escolhida como primeira área para se realizar a elaboração de uma adaptação dos padrões para atender as necessidades de um repositório do gênero.

Para construção da proposta foram selecionados e analisados 3 grandes padrões de metadados da área, são eles o MIBBI, o Darwin Core e o ABCD. O padrão MIBBI possui 40 módulos, porém apenas 23 principais e destes 23 apenas 17 plenamente desenvolvidos. O padrão Darwin Core possui 12 pacotes de descrição e carga semântica. E por fim o ABCD possui 38 categorias de metadados expansíveis. Em uma busca nos repositórios cadastrados no re3data percebemos ao realizar uma busca por repositórios de dados de

pesquisa da área da biologia ao se excluir padrões genéricos e de outros campos do conhecimentos correlatos resulta-se apenas 2 ocorrências do MIBBI e uma do Darwin Core (dos 42 resultados).

A partir da comparação foram levantados metadados de aparente importância na área, depois eliminadas redundâncias e por fim campos generalizados, resultando no quadro a seguir com 8 elementos na linguagem Darwin Core que possibilita a representação de todos os campos selecionados:

Quadro 1- Campos Biologia

Campo	Metadado
Organismo	dwc:higherClassification
Contexto Geológico	dwc:locality
Coleta	dwc:Event
Amostra	dwc:materialSampleID
Responsável (contato)	dwc:scientificNameAuthorship
Ocorrência (data, hora, local)	dwc:Occurrence
Taxonomia	dwc:Taxon
MeasurementOrFact	dwc:measurementID

Fonte: Elaboração própria.

Ciências sociais

Para o campo das ciências sociais por se tratar de um campo mais abrangente foram abordadas padrões também abrangentes, foram selecionados então os padrões MODS, METS, MARC, CERIF e Dublin Core (DC). O METS apresenta 7 classes, o MODS 20, o DC 15, o MARC 9 e o CERIF 22. As classes e seus elementos foram então comparadas após selecionados os campos foi possível realizar uma tradução para o padrão Datacite o que possibilitou a comunicação com o padrão generalista anterior.

Quadro 2 - Campos Ciências Sociais

Campo	Metadado
Origem/Coleta	datacite:description (descriptionType=methods)
Versão	datacite:version
Língua	datacite:language e dc:language
Palavras-chave	datacite:subject (subjectScheme=keyword)
Descrição/resumo	datacite:description (descriptionType=abstract)

Identificador interno de autor	datacite:identifier (identifierType=idlattes)
Instituição/publicador	datacite:affiliation
Financiamento	datacite:funderName

Fonte: Elaboração própria.

Agricultura

Foram selecionados os padrões AgMES e AGRIS. O padrão Agmes, com 21 campos baseados no Dublin Core, visa abranger questões de padrões semânticos no domínio da agricultura com relação à descrição, descoberta de recursos, interoperabilidade e troca de dados para diferentes tipos de recursos de informação enquanto o AGRIS, com 16 campos, é um perfil de aplicação para o sistema internacional de informação sobre diretrizes de ciências e tecnologias agrícolas sobre boas práticas para a informação. Na plataforma re3data não foi identificado, em repositórios de dados de agricultura, o uso de nenhum destes dois padrões, sendo mais utilizado padrões genéricos como o Datacite ou o Dublin core. Com a comparação entre os padrões os campos escolhidos foram:

Quadro 3 - Campos agricultura

Campo	Metadado
Corporação autora	Creator. CreatorCorporate
Classificação	Subject.Classification
Thesaurus	Subject. subjectThesaurus
Localização	availabilityLocation

Fonte: Elaboração própria.

6 Conclusão

Entre as dificuldades encontradas uma das mais proeminentes é não adoção de padrões de metadados já conhecidos e bem estabelecidos ou de ao menos padrões compatíveis entre si.

A metodologia apresentada parece ser suficiente para áreas do conhecimento que já possuem trabalhos direcionados à descrição de seus conteúdos e será um trabalho em constante desenvolvimento. Vale a pena destacar ainda que o método foi desenvolvido no

contexto do Instituto Brasileiro de Informação em Ciência e Tecnologia como desenvolvimento do marco 8 (proposição de padrões de interoperabilidade para repositórios de dados de pesquisa) do Compromisso 3 do 4º Plano de Ação Nacional em Governo Aberto, que visa ‘Estabelecer mecanismos de governança de dados científicos para o avanço da Ciência Aberta no Brasil’. O Compromisso é oficialmente coordenado pela Empresa Brasileira de Pesquisa Agropecuária (Embrapa), com a colaboração de uma rede crescente de instituições de governo e da sociedade civil, entre elas: Ministério da Ciência, Tecnologia e Inovações (MCTI), Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), Coordenação de Aperfeiçoamento de Pessoal do Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação Oswaldo Cruz (Fiocruz), Universidade de Brasília (UnB), Open Knowledge Foundation (OKBR), SciELO e Rede Nacional de Ensino e Pesquisa (RNP).

O processo está sendo aplicado também no desenvolvimento do Lattes Data. O Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e o Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) assinaram um acordo de cooperação (ACT) para a criação de um repositório de dados científicos, o Lattes Data. A assinatura foi realizada no Ministério da Ciência, Tecnologia, Inovações e Comunicações, no dia 12 de dezembro de 2019. Futuramente além de avançar em outras áreas do conhecimento e continuar explorando soluções para a área da saúde, será realizada a análise de um Corpus. Esse Corpus será um conjunto de termos utilizados para descrever os dados de cada área do conhecimento, retirados dos repositórios de dados temáticos (que utilizam o protocolo de comunicação OAI-PMH) recuperados a partir da plataforma re3data. O Corpus será formado a partir da coleta automática com o software coletador LaReferencia dos campos título, resumo, assunto e palavras-chave e depois organizados e visualizados pelo software VosViewer.

Por fim vale ressaltar que este não é um trabalho que se esgota no relatório, muito pelo contrário, ele será desdobrado em uma incessante busca por extensões e adaptações das diretrizes nas mais diversas áreas do conhecimento.

Referências

CONSELHO NACIONAL DE DESENVOLVIMENTO CIENTÍFICO E TECNOLÓGICO. **Tabela de Áreas do Conhecimento**. Brasília: CNPQ, 2020. Disponível em:

<http://lattes.cnpq.br/documents/11871/24930/TabeladeAreasdoConhecimento.pdf/d192ff6b-3e0a-4074-a74d-c280521bd5f7>. Acesso em: 30 de abr. 2020.

CRESWELL, John W. **Projetos de pesquisa: métodos qualitativo, quantitativo e misto**. 2. ed. Porto Alegre: Artmed, 2007. 248 p

DATA CURATION CENTRE. **Disciplinary Metadata**. [S.l.], 2020. Disponível em: <http://www.dcc.ac.uk/resources/metadata-standards>. Acesso em: 30 abr. 2020.

DEUTSCHE FORSCHUNGSGEMEINSCHAFT. **DFG Classification of Subject Area, Review Board, Research Area and Scientific Discipline**. Bonn, 2020. Disponível em:

https://www.dfg.de/en/dfg_profile/statutory_bodies/review_boards/subject_areas/index.jsp. Acesso em: 30 abr. 2020.

EUROPEAN OPEN SCIENCE INFRASTRUCTURE. **DRAFT: OpenAIRE Guidelines for Literature Repository Managers v4**. [S.l.], 2018. Disponível em: <https://openaire-guidelines-for-literature-repository-managers.readthedocs.io/en/v4.0.0/>. Acesso em: 30 abr. 2020.

FAIR DATA POINT. **FAIR Data Point Specification**. [S.l.], 2016. Disponível em: <https://github.com/FAIRDataTeam/FAIRDataPoint/wiki/FAIR-Data-Point-Specification>. Acesso em: 30 de abr. 2020.

GIL, Antonio Carlos. **Métodos e técnicas de pesquisa social**. 6. ed. 4. reimpr. São Paulo: Atlas, 2011. 216 p.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. **Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development, The Measurement of Scientific, Technological and Innovation Activities**. Paris: OECD Publishing, 2015. Disponível em: <https://www.oecd.org/publications/frascati-manual-2015-9789264239012-en.htm>. Acesso em: 30 abr. 2020.

RESEARCH DATA ALLIANCE. **Metadata Directory**. [S.l.], 2020. Disponível em: <http://rd-alliance.github.io/metadata-directory/>. Acesso em: 30 de abr. 2020.

RE3DATA. **Registry of research data repositories**. [S.l.], 2020. Disponível em: <http://service.re3data.org/about>. Acesso em: 30 abr. 2020.

SANTOS, Luiz O. B. da Silva. FAIR Data Points Supporting Big Data Interoperability. In: MERTINS, K. *et al* (org.). **Enterprise Interoperability in the Digitized and Networked Factory of the Future**. Londres: ISTE Press, 2016. p. 270-279.

WILKINSON. Mark D. *et. al.* The FAIR Guiding Principles for scientific data management and stewardship. **Sci Data**, [S.l.], v.3, 2016. DOI: <https://doi.org/10.1038/sdata.2016.18>.

ANEXO I - Especificações OpenAIRE v4

O quadro a seguir apresenta os campos Mandatários (M):

Quadro 4- Campos Mandatários

Campo	Metadado	Definição	Exemplo	Exemplo de aplicação	Vocabulário controlado
Title (Título)	datacite:title	Nome ou título pelo qual o recurso é conhecido.	National Institute for Environmental Studies and Center for Climate System	<datacite:title xml:lang="en-US"> National Institute for Environmental Studies and Center for Climate System Research Japan </datacite:title>	title type (https://openaire-guidelines-for-literature-repository-managers.readthedocs.io/en/v4.0.0/vocab_titletype.html#vocab-titletype-titletype)
Creator (Criador)	datacite:creator	Autor do recurso, pode ser uma instituição ou nome pessoal (No caso de nome pessoal no formato: Sobrenome, Nome).	Evans, R.J.	<datacite:creator> <datacite:creatorName>Evans, R.J.</datacite:creatorName> </datacite:creator>	name type (https://openaire-guidelines-for-literature-repository-managers.readthedocs.io/en/v4.0.0/vocab_nametype.html#vocab-nametype-nametype)
Publication Date (Data de Publicação)	datacite:date	Data com alguma associação com o ciclo de vida do recurso, deve ser utilizado o formato W3C (https://www.iso.org/iso-8601-date-and-time-format.html).	2000-12-25	<datacite:date dateType="Issued">2000-12-25</datacite:date>	date type (https://openaire-guidelines-for-literature-repository-managers.readthedocs.io/en/v4.0.0/vocab_datetype.html#vocab-datetype-datetype)
Resource Type (Tipo de Recurso)	oaire:resourceType	O gênero do recurso, tipo de produto científico.	Journal Article (Artigo de periódico)	<oaire:resourceType resourceTypeGeneral="literature" uri="http://purl.org/coar/resource_type/c_6501">journal article</oaire:resourceType>	COAR Resource Type Vocabulary (http://vocabularies.coar-repositories.org/documentation/resource_types/)

Resource Identifier (Identificador do Recurso)	datacite:identifier	Identificador único do recurso.	http://hdl.handle.net/1234/5628	<datacite:identifier identifierType="Handle">http://hdl.handle.net/1234/5628</datacite:identifier>	identifier type (https://openaire-guidelines-for-literature-repository-managers.readthedocs.io/en/v4.0.0/vocab_identifiertype.html#vocab-identifiertype-identifiertype)
Access Rights (Direitos de Acesso)	datacite:rights	Informações sobre como o recurso pode ser acessado, utilizado e disseminado.	Open Access (Acesso Aberto)	<datacite:rights rightsURI="http://purl.org/coar/access_right/c_abf2">open access</datacite:rights>	COAR Access Right Vocabulary (http://vocabularies.coar-repositories.org/documentation/access_rights/)

Fonte: Elaboração própria.

A seguir no próximo quadro aborda-se o nível de obrigatoriedade Mandatório se aplicável (MA):

Quadro 5- Campos Mandatórios se aplicáveis

Campo	Metadado	Definição	Exemplo	Exemplo de aplicação	Vocabulário controlado
Contributor (Colaborador)	datacite:contributor	Pessoa ou instituição que tenha tido algum tipo de colaboração na elaboração do recurso, se utilizado faz também ser obrigatório o uso do atributo contributorType como demonstrado no exemplo (contributorType possui vocabulário controlado em https://openaire-guidelines-for-literature-repository-managers.readthedocs.io/en/v4.0.0/field_contributor.html#dc-contributor)	Evans, R. J.	<datacite:contributor> <datacite:contributorName>Evans, R. J.</datacite:contributorName> <datacite:contributor>	name type (link já fornecido) e contributor type (https://openaire-guidelines-for-literature-repository-managers.readthedocs.io/en/v4.0.0/vocab_contributortype.html#vocab-contributortype-contributortype)
Funding Reference (Referência de financiamento)	oaire:fundingReference	Informações sobre apoio financeiro recebido. Subpropriedade funderName obrigatória quando utilizar este campo como demonstrado no exemplo.	European Commission (Comissão Euripéia)	<oaire:fundingReference> <oaire:funderName>European Commission</datacite:funderName> </oaire:fundingReference>	funderIdentifier type (https://openaire-guidelines-for-literature-repository-managers.readthedocs.io/en/v4.0.0/vocab_funderidentifier.html#vocab-funderidentifier)
Embargo Period Date (Período de embargo)	datacite:date	Datas relevantes para descrever o período de embargo. O campo direitos de acesso devem estar como “embargoed access” para utilizar esse campo e devem ser	2011-12-01 2012-12-01	<datacite:dates> <datacite:date dateType="Accepted">2011-12-01</datacite:date> <datacite:date dateType="Available">2012-12-01</datacite:date> </datacite:dates>	date type (link já fornecido)

		utilizados as datas de início e fim do período de embargo além dos atributos dateType como indicados no exemplo.			
Language (Língua)	dc:language	Idioma do conteúdo do recurso.	eng (Inglês)	<dc:language>eng</dc:language>	IETF BCP 47 (https://tools.ietf.org/rfc/bcp/bcp47.txt) e ISO 639-3 (https://iso639-3.sil.org/)
Publisher (Publicador)	dc:publisher	O publicador do recurso, usar apenas nome completo.	John Wiley & Sons, Inc. (US)	<dc:publisher>John Wiley & Sons, Inc. (US)</dc:publisher>	-
Description (Descrição)	dc:description	Um texto descrevendo o conteúdo do recurso (não se trata de uma descrição física).	A number of problems in quantum state and system identification are addressed (Um número de problemas no estado quântico e sistemas de identificação são abordados).	<dc:description xml:lang="en-US"> A number of problems in quantum state and system identification are addressed. </dc:description>	-
Subject (Assunto)	datacite:subject	Assunto geral, palavra-chave, ou qualquer outro código ou termo que representa o recurso, se utilizar algum código utilizar o atributo subjectScheme identificando-o caso seja possível.	Earth sciences and geology (Ciências da terra e geologia).	<datacite:subject>Earth sciences and geology</datacite:subject>	-
File Location (Localização do arquivo)	oaire:file	referência para desambiguação do arquivo, utilizar um HTTP URI.	http://purl.org/coar/access_right/c_abf2 " mimeType="application/pdf	<oaire:file accessRightsURI="http://purl.org/coar/access_right/c_abf2"	-

				<pre> mimeType="application/pdf" objectType="fulltext">http://link-to-the-fulltext.org</oaire:file> </pre>	
--	--	--	--	---	--

Fonte: Elaboração própria.

O próximo quadro se trata dos campos Recomendados (R) :

Quadro 6- Campos Recomendados

Campo	Metadado	Definição	Exemplo	Exemplo de aplicação	Vocabulário controlado
Alternate Identifier (Identificador alternativo)	datacite:alternateIdentifier	Outro identificador que não seja o principal. Caso esse campo seja utilizado é obrigatório o uso do atributo alternateIdentifierType.	http://someUrl	<pre> <datacite:alternateIdentifiers> <datacite:alternateIdentifier alternateIdentifierType="URL">http://someUrl</datacite:alternateIdentifier> </datacite:alternateIdentifiers> </pre>	alternateIdentifier type (https://openaire-guidelines-for-literature-repository-managers.readthedocs.io/en/v4.0.0/vocab_alternateidentifiertype.html#vocab-alternateidentifiertype-identifiertype)
Related Identifier (Identificador relacionado)	datacite:relatedIdentifier	Identificador de um outro recurso associado ao recurso principal descrito. Obrigatório o uso do atributo relatedIdentifierType (vocabulário controlado disponível) e o uso do atributo relationType (também utiliza vocabulário disponibilizado).	http://someUrl	<pre> <datacite:relatedIdentifiers> <datacite:relatedIdentifier relatedIdentifierType="URL" relationType="HasPart">http://someUrl</datacite:relatedIdentifier> </datacite:relatedIdentifiers> </pre>	relatedIdentifier type (https://openaire-guidelines-for-literature-repository-managers.readthedocs.io/en/v4.0.0/vocab_relatedidentifiertype.html#vocab-relatedidentifiertype-identifiertype), relation type (https://openaire-guidelines-for-literature-repository-managers.readthedocs.io/en/v4.0.0/vocab_relationtype.html#vocab-relationtype-rel

					ationtype) e resourcetype general (https://openaire-guidelines-for-literature-repository-managers.readthedocs.io/en/v4.0.0/vocab_resourcetypegeneral.html#vocab-resourcetypegeneral)
Format (Formato)	dc:format	Manifestação física ou digital do recurso.	application/xml	<dc:format>application/xml</dc:format>	-
Source (Fonte)	dc:source	Referência à outro recurso do qual o original descrito é derivado.	Ecology Letters (1461023X) vol.4 (2001)	<dc:source>Ecology Letters (1461023X) vol.4 (2001)</dc:source>	-
License Condition (Condição da licença)	oaire:licenseCondition	Informações sobre direitos e licença como uso e acesso ao recurso.	Creative Commons Attribution-Non Commercial	<oaire:licenseCondition startDate="2019-02-01" uri="http://creativecommons.org/licenses/by-nc/4.0/">Creative Commons Attribution-Non Commercial</oaire:licenseCondition>	-
Coverage (Cobertura)	dc:coverage	O escopo ou abrangência do conteúdo recurso. Podendo ser geográfico, temporal, assunto o qualquer outro.	2000-2010	<dc:coverage>2000-2010</dc:coverage>	-
Resource Version (Versão do recurso)	oaire:version	Dependendo do tipo de recurso pode indicar, o	1.0.3	<oaire:version>1.0.3</oaire:version>	COAR Version Vocabulary (http://vocabulari

		número da versão ou status da publicação.			es.coar-repositori es.org/document ation/version_t ypes/)
Citation Title (Título para citação)	oaire:citationTit le	Título de onde o recurso está inserido (revista, livro, etc.) conforme constará na citação bibliográfica.	Journal Title	<oaire:citationTit le>some Journal Title</oaire:citati onTitle>	-
Citation Volume (Volume para citação)	oaire:citationVol ume	Volume conforme constará na citação bibliográfica.	10	<oaire:citationV olume>10</oaire :citationVolume >	-
Citation Issue (Exemplar de citação)	oaire:citationIssu e	Exemplar conforme constará na citação bibliográfica.	1	<oaire:citationIss ue>1</oaire:citat ionIssue>	-
Citation Start Page (Página inicial para citação)	oaire:citationStar tPage	Página inicial conforme constará na citação bibliográfica.	100	<oaire:citationSt artPage>100</oa ire:citationStartP age>	-
Citation End Page (Página final para citação)	oaire:citationEnd Page	Página final conforme constará na citação bibliográfica.	105	<oaire:citationEn dPage>105</oair e:citationEndPag e>	-
Citation Edition (Edição para citação)	oaire:citationEdit ion	Edição conforme constará na citação bibliográfica.	2	<oaire:citationEd ition>2</oaire:cit ationEdition>	-
Citation Conference Place (Local da Conferência para citação)	oaire:citationCon ferencePlace	Local da conferência conforme constará na citação bibliográfica.	Berlin	<oaire:citationCo nferencePlace>B erlin</oaire:citati onConferencePla ce>	-
Citation Conference Date (Data da Conferência)	oaire:citationCon ferenceDate	Data da conferência conforme constará na	2013-10-22	<oaire:citationCo nferenceDate>20 13-10-22</oaire: citationConferen	-

para citação)		citação bibliográfica, deve ser conforme as regras W3C já apresentadas.		ceDate>	
----------------------	--	---	--	---------	--

Fonte: Elaboração própria.

Por fim o quadro com os campos opcionais (O):

Quadro 7- Campos opcionais

Campo	Metadado	Definição	Exemplo	Exemplo de aplicação	Vocabulário controlado
Size (Tamanho)	datacite:size	Tamanho do recurso, pode ser digital ou físico.	6 MB	<datacite:size>6 MB</datacite:size>	-
Geo Location (Geo localização)	datacite:geoLocation	Localização geográfica ou espacial à qual o recurso se refere ou onde foi desenvolvido. Utilizar latitude e longitude como apresentado no exemplo. Caso deseje utilizar áreas ao invés de ponto de localização utilizar as subpropriedades disponíveis.	31.233 -67.302	<datacite:geoLocation><datacite:geoLocationPlace>Atlantic Ocean</datacite:geoLocationPlace><datacite:geoLocationPoint><datacite:pointLongitude>31.233</datacite:pointLongitude><datacite:pointLatitude>-67.302</datacite:pointLatitude></datacite:geoLocationPoint>	-
Audience (Público)	dcterms:audience	Classe o entidade à qual o recurso pode ser útil ou foi pensado para.	Researchers (Pesquisadores)	<dcterms:audience>Researchers</dcterms:audience>	-

Fonte: Elaboração própria.

ANEXO II- Comparação OAIRE e FDP

Foi realizada uma comparação entre OpenAIRE e FDP a fim de encontrar discrepâncias e sinônimos entre os campos a tabela apresentada a seguir foi elaborada objetivando visualizar as equivalências semânticas entre os campos:

Quadro 8- Comparação OAIRE FDP

	Equivalente em OAIRE	Nome
Obrigatórios FDP (10)	oaire:resourceType (diferentes vocabulários)	rdf:type
	datacite:title (diferentes vocabulários)	dct:title
	datacite:creator	dct:publisher
	oaire:version (no OAIRE é opcional)	dct:hasVersion
	datacite:relatedIdentifier (pode ser usado para outros níveis de relações) (é opcional no OAIRE aqui é obrigatório por causa do catálogo)	dct:isPartOf
	datacite:identifier (diferentes vocabulários)	fdp:metadataIdentifier
	datacite:date (especificar o dateType como metadado)	fdp:metadataIssued
	datacite:date (especificar o dateType como última modificação dos metadados)	fdp:metadataModified
	Não encontrado	dcat:distribution
	datacite:subject	dcat:theme
Obrigatórios OAIRE (6)	acima	datacite:title
	acima	datacite:creator
	dct:issued (opcional no FDP)	datacite:date
	acima	oaire:resourceType
	acima	datacite:identifier
	dct:accessRights (opcional no FDP)	datacite:rights

Fonte: Elaboração própria.

A partir do quadro 5 apresentada podemos perceber q a maioria dos campos dos padrões selecionados apresentam equivalências entre si mesmo que com níveis de obrigatoriedade diferentes. O quadro 8 acima relaciona apenas os campos obrigatórios a fim de estabelecer o núcleo de metadados visado,

Alguns dos problemas encontrados a partir da comparação, além das diferenças de obrigatoriedade já citados, foram: a necessidade de especificação de alguns campos no OAIRE a partir das funções Type (ex: dateType), a adoção de diversos vocabulários controlados diferentes, e a incompatibilidade do campo dcat:distribution do FDP, que se refere ao registro dos metadados da distribuição (arquivo).

ANEXO III- Estudo áreas do conhecimento

Para definição dos metadados temáticos, a primeira barreira encontrada foi o caráter multidisciplinar de um repositório do tipo, ou seja qual divisão de áreas do conhecimento será abordadas para seleção dos padrões vigentes, para tanto foram analisadas 3 tabelas de áreas do conhecimento, sendo elas: a tabela CNPq; a encontrada no manual Frascati e por fim a divisão utilizada pelo DCC/RDA.

Quadro 9- Comparação áreas do conhecimento

CNPq	Frascati	RDA/DCC
Ciências exatas e da terra	Ciências Naturais*	Matemáticas e Físicas
Engenharias	Engenharia e Tecnologia	
Ciências Biológicas*	-	-
Ciências da Saúde	Ciências Médicas e de saúde	Ciências da vida*
Ciências agrárias	Agricultura e veterinária	
Ciências sociais aplicadas	Ciências sociais	Ciências sociais e comportamentais
Ciências humanas	Humanidades e artes	Artes e humanidades
Linguística, letras e artes		

Fonte: Elaboração própria.

As diferenças para a tabela CNPq estão na primeira e última categorias, a tabela CNPq separa a primeira categoria entre ciências naturais e ciências biológicas. E a última categoria, também separa entre Humanidades, e Linguística e artes.

ANEXO IV- Padrões de metadados por área do conhecimento

A partir das diferentes divisões das áreas do conhecimento se iniciou o processo de busca por padrões de metadados, partindo-se então para uma lista, publicada pela DCC, contendo os diversos padrões de metadados já divididos por suas áreas do conhecimento que foram analisados, deduplicados e brevemente explicados, a seguir:

Artes e Humanidades

- [Encoded Archival Description \(EAD\)](#)

Um padrão para codificar auxiliares de busca de recursos de busca arquivísticos usando XML em repositórios de arquivos e manuscritos, implementando as recomendações do Conselho Internacional de Arquivos ISAD (G): Padrão Internacional de Descrição Geral de Arquivos.

- [TEI - Text Encoding Initiative](#)

As diretrizes da Iniciativa de Codificação de Texto fazem recomendações sobre maneiras adequadas de representar as características de recursos textuais. Eles especificam um conjunto de tags XML para marcar os metadados textuais, a estrutura do texto, o relacionamento entre imagens e transcrições e outros recursos de interesse. As diretrizes descrevem o esquema formal de codificação TEI e documentam o vocabulário de marcação TEI.

- [MARC \(Machine-Readable Cataloging\)](#)

O MARC é um formato de serialização padrão para representar metadados bibliográficos, originalmente projetado como uma maneira de trocar registros bibliográficos entre catálogos de bibliotecas. Várias versões diferentes foram definidas, principalmente com escopo nacional ou regional, das quais o MARC 21 é provavelmente o mais amplamente utilizado.

Ciências da vida

- [ABCD \(Access to Biological Collection Data\)](#)

O Esquema de Acesso a Dados de Coleções Biológicas (ABCD) é um padrão abrangente em evolução para o acesso e o intercâmbio de dados sobre espécimes e observações (também conhecidos como dados primários de biodiversidade). O esquema ABCD é altamente estruturado, suportando dados de uma ampla variedade de bancos de dados e compatível com vários padrões de dados existentes. Estruturas paralelas existem para que um (ou ambos) dados atomizados e texto livre possam ser acomodados.

Patrocinado pelo TDWG - Padrões de Informações sobre Biodiversidade - o Grupo de Trabalho sobre Bancos de Dados Taxonômicos.

- [Darwin Core](#)

Um corpo de padrões, incluindo um glossário de termos, destinados a facilitar o compartilhamento de informações sobre diversidade biológica, fornecendo definições de referência, exemplos e comentários . Patrocinado pelos Padrões de Informação sobre Biodiversidade (TWDG).

- [EML \(Ecological Metadata Language\)](#)

A Linguagem de Metadados Ecológicos (EML) é uma especificação de metadados especialmente desenvolvida para a disciplina de ecologia. Baseia-se em trabalhos anteriores realizados pela Sociedade Ecológica da América e em esforços associados. Patrocinado pela ecoinformatics.org.

- [Genome Metadata](#)

O metadados genoma no PATRIC consistem em 61 campos de metadados diferentes, chamados atributos, organizados nas sete categorias principais a seguir: Informações do organismo, Informações do isolado, Informações do hospedeiro, Informações da sequência, Informações do fenótipo, Informações do projeto e Outros.

- [ISA-Tab](#)

O formato delimitado por tabulação Investigação / Estudo / Ensaio (ISA) (TAB) é uma estrutura de uso geral com a qual coletar e comunicar metadados complexos (isto é, características da amostra, tecnologias usadas, tipo de medidas feitas). Criado por desenvolvedores principais da Universidade de Oxford, o ISA-TAB foi lançado em novembro de 2008.

- [MIBBI \(Minimum Information for Biological and Biomedical Investigations\)](#)

Um portal comum para um grupo de quase 40 listas de verificação de informações mínimas para várias disciplinas biológicas. O conceito foi realizado inicialmente através dos esforços conjuntos da Iniciativa Proteomics Standards, do Genomic Standards Consortium e dos Grupos de Trabalho MGED RSBI.

- [Observ-OM](#)

O Observ-OM baseia-se em quatro conceitos básicos para representar qualquer tipo de observação: Destinos, Recursos, Protocolos (e suas Aplicações) e Valores.

- [OME-XML \(Open Microscopy Environment XML\)](#)

OME-XML é um formato de arquivo para dados de imagem biológica, com ênfase em metadados que suportam microscopia de luz. Ele pode ser usado como um formato de arquivo de dados por si só ou como uma maneira de codificar metadados em um arquivo TIFF ou BigTIFF (para esse fim, existe a especificação OME-TIFF). O padrão é mantido pelo Open Microscopy Environment Consortium.

- [PDBx/mmCIF \(Protein Data Bank Exchange Dictionary and the Macromolecular Crystallographic Information Framework\)](#)

O arquivo Protein Data Bank (PDB) é o único repositório mundial de informações sobre estruturas 3D de proteínas, ácidos nucleicos e conjuntos complexos, gerenciados pelo PDB Mundial (wwPDB). O PDB Exchange Dictionary (PDBx) é usado pelo wwPDB para definir o conteúdo dos dados para deposição, anotação e arquivamento de entradas do PDB. O PDBx incorpora a representação de metadados padrão da comunidade, a Estrutura de Informações Cristalográficas Macromoleculares (mmCIF), desenvolvida originalmente sob os auspícios da União Internacional de Cristalografia (IUCr). O PDBx foi estendido pelo wwPDB para incluir descrições de outros métodos experimentais que produzem modelos de estrutura macromolecular 3D, como espectroscopia de ressonância magnética nuclear, microscopia eletrônica de varredura 3D e tomografia.

- [Protocol Data Element Definitions](#)

Um conjunto de elementos exigidos pelas Instituições Nacionais de Saúde (EUA) para o envio de informações de teste e dados ao banco de dados de resultados e registro do ClinicalTrials.gov.

- [UKEOF](#)

Um padrão de metadados para descrever atividades, programas, redes e instalações de monitoramento ambiental publicados pelo UK Environmental Observation Framework (UKEOF).

Ciências Matemáticas e Físicas

- [AgMES \(Agricultural Metadata Element Set\)](#)

Padrão semântico desenvolvido pela Organização das Nações Unidas para Agricultura e Alimentação (FAO), o AgMES permite descrição, descoberta de recursos, interoperabilidade e troca de dados de diferentes tipos de recursos de informação em todas as áreas relevantes para a produção de alimentos, nutrição e desenvolvimento rural. Patrocinado pela ONU AIMS.

- [AVM \(Astronomy Visualization Metadata\)](#)

O esquema AVM suporta a pesquisa cruzada de coleções de imagens astronômicas prontas para impressão e prontas para tela, renderizadas a partir de observações telescópicas. O esquema é compatível com a especificação do Adobe XMP, para que os metadados possam ser incorporados em formatos de imagem comuns, como JPEG, TIFF e PNG. Essas imagens podem combinar dados adquiridos em diferentes bandas de onda e de diferentes observatórios. Especificamente, o subconjunto mais geral desse esquema também é apropriado para descrever obras de arte e ilustrações de assuntos astronômicos.

- [CF \(Climate and Forecast\) Metadata Conventions](#)

O padrão CF foi originalmente estruturado como um padrão para dados gravados no formato netCDF, tendo particularmente em mente os dados de previsão climática gerados pelo modelo. No entanto, é igualmente aplicável a conjuntos de dados observacionais e pode ser usado para descrever outros formatos. É um padrão que visa distinguir quantidades (como descrição física, unidades e processamento anterior) e localizar os dados no espaço-tempo. Patrocinado pela Convenção de Metadados do Clima e Previsão do NetCDF.

- [CIF \(Crystallographic Information Framework\)](#)

Uma estrutura de arquivo padrão bem estabelecida para o arquivamento e distribuição de informações cristalográficas, o CIF é usado regularmente para relatar determinações de estruturas cristalinas para o Acta Crystallographica e outros periódicos. Patrocinado pela União Internacional de Cristalografia.

- [CIM \(Common Information Model\)](#)

O Common Information Model (CIM) descreve dados climáticos, os modelos e softwares dos quais eles derivam, as grades geográficas usadas para calculá-los e projetá-los e os processos experimentais (geralmente simulações) que os produziram. A CIM foi originalmente desenvolvida pelo Projeto Metafor, financiado pela UE. Agora é mantido e desenvolvido pela Earth Science Documentation (ES-DOC).

- [CSMD \(Core Scientific Metadata Model\)](#)

Um modelo orientado a dados de estudo, principalmente no suporte ao software de infraestrutura de gerenciamento de dados do ICAT. O CSMD foi projetado para suportar dados coletados no fluxo de trabalho científico de uma instalação em grande escala. Patrocinado pelo Science and Technologies Facilities Council.

- [DIF \(Directory Interchange Format\)](#)

Uma iniciativa inicial de metadados da comunidade de ciências da Terra, destinada à descrição de conjuntos de dados científicos. Inclui elementos focados em instrumentos que capturam dados, características temporais e espaciais dos dados e

projetos aos quais o conjunto de dados está associado. É definido como um esquema XML do W3C. Patrocinado pelo Diretório principal de mudanças globais.

- [FGDC/CSDGM \(Federal Geographic Data Committee Content Standard for Digital Geospatial Metadata\)](#)

Um padrão amplamente utilizado, que define o conteúdo de informações para um conjunto de dados geoespaciais digitais exigidos pelo governo federal dos EUA. O CSDGM foi patrocinado pelo Comitê Federal de Dados Geográficos dos EUA.

- [FITS \(Flexible Image Transport System\)](#)

O FITS é um formato de arquivo de dados de imagem para codificar dados astronômicos. As convenções do WCS (World Coordinate System) mapeiam elementos em matrizes de dados para coordenadas físicas padrão no céu. O FITS possui provisões para metadados de imagem codificados em um cabeçalho ASCII no início dos arquivos.

- [International Virtual Observatory Alliance Technical Specifications](#)

As especificações técnicas definidas pela IVOA (International Virtual Observatory Alliance) permitem a interoperabilidade e a integração de arquivos astronômicos em todo o mundo em um observatório virtual internacional. Eles incluem vários modelos de dados que atuam como esquemas de metadados para tipos de dados específicos: por exemplo, dados de fotometria, dados de simulação, coordenadas espaço-temporais, dados de linhas espectrais, dados espectrais, dados observacionais e o espaço físico de parâmetros de conjuntos de dados astronômicos.

- [ISO 19115](#)

Um esquema adotado internacionalmente para descrever informações e serviços geográficos. Ele fornece informações sobre a identificação, extensão, qualidade, esquema espacial e temporal, referência espacial e distribuição de dados geográficos digitais.

Patrocinado pela International Standards Organization, foi dividida em partes: a ISO 19115-1: 2014 contém os fundamentos da norma; A ISO 19115-2: 2009 contém extensões para imagens e dados em grade; e a ISO / TS 19115-3: 2016 fornece uma implementação de esquema XML para os conceitos fundamentais compatíveis com a ISO / TS 19138: 2007 (XML de metadados geográficos ou GMD).

- [NeXus](#)

O NeXus é um padrão internacional para armazenamento e troca de dados experimentais de nêutrons, raios-x e múons. A estrutura dos arquivos NeXus é extremamente flexível, permitindo o armazenamento de conjuntos de dados simples, como um único conjunto de dados e seus eixos, e dados altamente complexos e seus metadados associados, como medições em um instrumento multicomponente ou simulações numéricas. O NeXus é construído sobre o formato HDF5 do contêiner e adiciona regras

específicas do domínio para organizar dados nos arquivos HDF5, além de um dicionário de nomes de campos específicos do domínio bem definidos.

- [SDAC \(Standard for Documentation of Astronomical Catalogues\)](#)

O Padrão para Documentação de Catálogos Astronômicos é um conjunto de convenções para arquivar dados astronômicos. Além de convenções de caminho, nome de arquivo e formato de dados, também especifica como construir um arquivo de descrição de texto sem formatação para documentar os arquivos de dados. Foi desenvolvido como uma alternativa ao FITS, que seria mais adequada para arquivos, permitir inspeção humana e manipulação através de ferramentas de linha de comando padrão do Unix. O SDAC foi desenvolvido pelo CDS (Centre de Données astronomiques de Strasbourg).

- [SPASE Data Model](#)

Um modelo de informação para descrever os elementos do ambiente de dados de heliofísica e um conjunto de tipos de recursos que podem ser usados para descrever dados, juntamente com seu contexto científico, fonte, proveniência, conteúdo e localização. Ele foi projetado para suportar um sistema de dados federado em que os dados podem residir em locais diferentes e podem ser separados dos metadados que os descrevem. O formulário de expressão preferido é XML. O esforço de busca e extração de arquivos de física espacial (SPASE) é implementado pelo consórcio SPASE, composto por representantes da comunidade internacional de dados de Heliofísica.

Ciências Sociais e Comportamentais

- [DDI \(Data Documentation Initiative\)](#)

Um padrão internacional amplamente usado para descrever dados das ciências sociais, comportamentais e econômicas. Atualmente, duas versões do padrão são mantidas em paralelo:

O DDI Codebook (ou DDI versão 2) é o mais simples dos dois e destina-se a documentar dados simples de pesquisa para troca ou arquivamento. A versão 2.5 foi lançada em janeiro de 2014.

O ciclo de vida do DDI (ou DDI versão 3) é mais rico e pode ser usado para documentar conjuntos de dados em cada estágio do seu ciclo de vida, desde a conceituação até a publicação e reutilização. É modular e extensível.

Ambas as versões são baseadas em XML e definidas usando esquemas XML. Eles foram desenvolvidos e são mantidos pela Aliança DDI.

- [MIDAS-Heritage](#)

Um padrão de patrimônio cultural britânico para registrar informações sobre prédios, sítios arqueológicos, naufrágios, parques e jardins, campos de batalha, áreas de interesse e artefatos. Patrocinado pelo Fórum sobre Padrões de Informação no Patrimônio.

- [QuDEX \(Qualitative Data Exchange Format\)](#)

O padrão / esquema QuDEX é um formato neutro em software para dados qualitativos que preserva anotações e relacionamentos entre dados e outros objetos relacionados. Ele pode ser visto como o modelo ideal de troca de dados da linha de base para arquivamento e intercâmbio de dados e metadados.

- [SDMX \(Statistical Data and Metadata Exchange\)](#)

Um conjunto de padrões e diretrizes técnicas e estatísticas comuns a serem usados para o intercâmbio e o compartilhamento eficientes de dados e metadados estatísticos. As instituições patrocinadoras incluem BIS, BCE, EUROSTAT, FMI, OCDE, ONU e Banco Mundial.

Dados de pesquisa gerais

- [CERIF \(Common European Research Information Format\)](#)

O Common European Information Information Format é o padrão que a UE recomenda a seus estados membros para registrar informações sobre as atividades de pesquisa. Ele inclui suporte específico para a gravação de metadados para conjuntos de dados.

- [Data Package](#)

O Data Package é um formato genérico de wrapper para troca de dados. Embora ele suporte metadados arbitrários, o formato define campos obrigatórios, recomendados e opcionais para o pacote como um todo e para os recursos contidos nele. Uma especificação separada mas vinculada fornece uma maneira de descrever as colunas de uma tabela de dados; descrições deste formulário podem ser incluídas diretamente nos metadados do pacote de dados.

- [DataCite Metadata Schema](#)

Um conjunto de metadados obrigatórios que devem ser registrados no DataCite Metadata Store ao criar um identificador persistente de DOI para um conjunto de dados. As propriedades sem domínio específico foram escolhidas por sua capacidade de ajudar na identificação precisa e consistente de dados para fins de citação e recuperação. Patrocinado pelo consórcio DataCite.

- [DCAT \(Data Catalog Vocabulary\)](#)

O DCAT é utilizado para descrever conjuntos de dados em catálogos de dados, os editores aumentam a capacidade de descoberta e permitem que os aplicativos acessem facilmente metadados de vários catálogos. Além disso, permite a publicação descentralizada de catálogos e facilita a pesquisa federada de conjuntos de dados nos sites. Os metadados DCAT agregados podem servir como um arquivo de manifesto para facilitar a preservação digital.

- [Dublin Core](#)

Um padrão básico, independente de domínio, que pode ser facilmente entendido e implementado e, como tal, é um dos padrões de metadados mais conhecidos e mais amplamente utilizados. Patrocinado pela Dublin Core Metadata Initiative, o Dublin Core foi publicado como padrão ISO 15836 em fevereiro de 2009.

- [MODS \(Metadata Object Description Schema\)](#)

O Esquema de Descrição de Objeto de Metadados (MODS) é um padrão de metadados bibliográficos implementado em XML. Ele reimplementa um subconjunto dos elementos do MARC (Catalogação legível por máquina) usando tags baseadas em idioma em vez de numéricas e as agrupa de maneira um pouco diferente. Ele se destina tanto a uma versão simplificada do MARC 21 quanto a uma alternativa mais rica ao Dublin Core para aplicações como organização / coleta de metadados e documentação de pacotes de informações digitais. Foi desenvolvido em 2002 pelo setor de Desenvolvimento de Rede da Biblioteca do Congresso e pelo Escritório de Padrões do MARC, juntamente com um grupo de especialistas interessados.

- [OAI-ORE \(Open Archives Initiative Object Reuse and Exchange\)](#)

O objetivo desse padrão é expor o conteúdo rico em agregações de recursos da Web a aplicativos que oferecem suporte à criação, depósito, troca, visualização, reutilização e preservação. O padrão apóia a natureza mutável da bolsa de estudos e da comunicação acadêmica, e a necessidade de infra-estrutura cibernética para apoiar essa bolsa, objetivando desenvolver um padrão que generalize todas as informações baseadas na Web, incluindo as crescentes redes sociais populares da "Web 2.0".

- [Observations and Measurements \(Observações e Medidas\)](#)

Essa codificação é uma dependência essencial para o padrão de interface do Serviço de Observação de Sensor OGC (SOS). Mais especificamente, esse padrão define esquemas XML para observações e recursos envolvidos na amostragem ao fazer observações. Eles fornecem modelos de documentos para o intercâmbio de informações que descrevem os atos de observação e seus resultados, dentro e entre diferentes comunidades científicas e técnicas.

- [PREMIS](#)

O Dicionário de Dados PREMIS (Preservation Metadata: Implementation Strategies) define um conjunto de metadados para os repositórios de objetos digitais registrarem e usar para preservar esses objetos a longo prazo. Ele tem suas raízes no Modelo de Referência do Sistema de Informações de Arquivamento Aberto, mas foi fortemente influenciado pela experiência prática de tais repositórios. Embora o Dicionário de Dados possa ser usado com outros padrões para influenciar a criação de perfis de aplicativos locais, um Esquema XML é fornecido para permitir que os metadados sejam serializados independentemente.

O PREMIS foi desenvolvido inicialmente pelo Grupo de Trabalho de Preservação de Metadados: Estratégias de Implementação, organizado pela OCLC e RLG, e atualmente é mantido pela Atividade de Manutenção do PREMIS, liderada pela Biblioteca do Congresso.

- [PROV](#)

A proveniência é a informação sobre entidades, atividades e pessoas envolvidas na produção de um dado ou coisa, que pode ser usada para formar avaliações sobre sua qualidade ou confiabilidade. A família de documentos PROV define um modelo, serializações correspondentes e outras definições de suporte para permitir o intercâmbio interoperável de informações de proveniência em ambientes heterogêneos, como a Web.

- [RDF Data Cube Vocabulary](#)

O padrão fornece um meio para publicar dados multidimensionais, como estatísticas, na Web, de forma que possam ser vinculados a conjuntos de dados e conceitos relacionados usando o padrão W3C RDF (Resource Description Framework).

Porém a lista não apresentou uma especificidade considerada suficiente dentro das áreas do conhecimento então foi utilizado o projeto mantido pela comunidade que deu origem a essa lista. Ele é mantido pela RDA e pode ser encontrado em <http://rd-alliance.github.io/metadata-directory/>.

A partir da lista de padrões da DCC e de sua distribuição pela RDA foram selecionados os padrões de cada área do conhecimento para análise. Após essa seleção inicial uma verificação na plataforma re3data foi realizada visando conferir quais padrões estão realmente sendo utilizados.