

UNIVERSIDADE FEDERAL DE MINAS GERAIS
ESCOLA DE CIÊNCIA DA INFORMAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM GESTÃO E ORGANIZAÇÃO DO
CONHECIMENTO

Filipi Miranda Soares

**PRINCÍPIOS PARA A CRIAÇÃO DE UMA EXTENSÃO DE METADADOS SOBRE
INTERAÇÕES ECOLÓGICAS NA AGROBIODIVERSIDADE PARA O PADRÃO
DARWIN CORE**

Belo Horizonte
2019

FILIPY MIRANDA SOARES

**PRINCÍPIOS PARA A CRIAÇÃO DE UMA EXTENSÃO DE METADADOS SOBRE
INTERAÇÕES ECOLÓGICAS NA AGROBIODIVERSIDADE PARA O PADRÃO
DARWIN CORE**

Dissertação apresentada ao Programa de Pós-Graduação em Gestão e Organização do Conhecimento da Escola de Ciência da Informação da Universidade Federal de Minas Gerais, como requisito para a obtenção do título de mestre.

Área de Concentração: Ciência da Informação

Linha de Pesquisa: Arquitetura e Organização do Conhecimento

Orientadora: Dra. Benildes Coura Moreira dos Santos Maculan

Coorientadora: Dra. Debora Pignatari Drucker

Belo Horizonte
2019

S676p Soares, Filipi Miranda

Princípios para a criação de uma extensão de metadados sobre interações ecológicas na agrobiodiversidade para o padrão Darwin Core [recurso eletrônico] / Filipi Miranda Soares. – 2019.

1 recurso eletrônico (241 f. : il., color): pdf.

Orientador: Benildes Coura Moreira dos Santos Maculan.

Dissertação (mestrado) – Universidade Federal de Minas Gerais, Escola de Ciência da Informação.

Referências: f. 189-206.

Apêndices: f. 207-231.

Anexos: f. 232-240.

1. Representação da informação. 2. Metadados. 3. Darwin Core. 4. Agrobiodiversidade. 5. Interações ecológicas. I. Maculan, Benildes Coura Moreira dos Santos. II. Universidade Federal de Minas Gerais, Escola de Ciência da Informação. III. Título.

CDD: 025.316



FOLHA DE APROVAÇÃO

**PRINCÍPIOS PARA A CRIAÇÃO DE UMA EXTENSÃO DE METADADOS SOBRE
INTERAÇÕES ECOLÓGICAS NA AGROBIODIVERSIDADE PARA O PADRÃO
DARWIN CORE**

FILIPI MIRANDA SOARES

Dissertação submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em GESTÃO E ORGANIZAÇÃO DO CONHECIMENTO, como requisito para obtenção do grau de Mestre em GESTÃO E ORGANIZAÇÃO DO CONHECIMENTO, área de concentração CIÊNCIA DA INFORMAÇÃO, linha de pesquisa Arquitetura e Organização do Conhecimento.

Aprovada em 18 de dezembro de 2019, pela banca constituída pelos membros:

Prof(a). Benildes Coura Moreira dos Santos Maculan (Orientadora)
ECI/UFMG

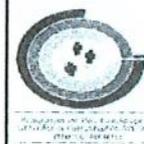
Dr(a). Debora Pignatari Drucker
Embrapa [por videoconferência] (Coorientadora)

Prof(a). Antonio Mauro Saraiva
USP [por videoconferência]

Dr(a). Ivo Fierozzi Júnior
Embrapa [por videoconferência]

Prof(a). Célia da Consolação Dias
ECI/UFMG

Belo Horizonte, 18 de dezembro de 2019.



ATA DA DEFESA DA DISSERTAÇÃO DO ALUNO FILIPI MIRANDA SOARES

Realizou-se, no dia 18 de dezembro de 2019, às 14:00 horas, Sala 1000 - ECI/UFMG, da Universidade Federal de Minas Gerais, a defesa de dissertação, intitulada *PRINCÍPIOS PARA A CRIAÇÃO DE UMA EXTENSÃO DE METADADOS SOBRE INTERAÇÕES ECOLÓGICAS NA AGROBIODIVERSIDADE PARA O PADRÃO DARWIN CORE*, apresentada por FILIPI MIRANDA SOARES, número de registro 2018666848, graduado no curso de BIBLIOTECONOMIA, como requisito parcial para a obtenção do grau de Mestre em GESTÃO E ORGANIZAÇÃO DO CONHECIMENTO, à seguinte Comissão Examinadora: Prof(a). Benildes Coura Moreira dos Santos Maculan - ECI/UFMG (Orientadora), Dr(a). Debora Pignatari Drucker - Embrapa [por videoconferência] (Coorientadora), Prof(a). Antonio Mauro Saraiva - USP [por videoconferência], Dr(a). Ivo Pierozzi Júnior - Embrapa [por videoconferência], Prof(a). Célia da Consolação Dias - ECI/UFMG.

A Comissão considerou a dissertação:

Aprovada

Reprovada

Finalizados os trabalhos, lavrei a presente ata que, lida e aprovada, vai assinada por mim e pelos membros da Comissão.

Belo Horizonte, 18 de dezembro de 2019.

Prof(a). Benildes Coura Moreira dos Santos Maculan

Dr(a). Debora Pignatari Drucker

Prof(a). Antonio Mauro Saraiva

Dr(a). Ivo Pierozzi Júnior

Prof(a). Célia da Consolação Dias

Dedico este trabalho aos meus pais, meus maiores incentivadores e apoiadores.

Também a Adriana Bogliolo (*in memoriam*), por instigar em mim o desejo da investigação científica.

AGRADECIMENTOS

A Deus e à Espiritualidade amiga que me acompanham e me trazem tanto esclarecimento, especialmente neste processo de desenvolvimento científico.

À minha orientadora Benildes Maculan, para quem eu poderia escrever uma dissertação inteira só de agradecimentos. Esteve comigo desde 2015, me conduzindo com maestria pelos caminhos da pesquisa e me apoiando de diversas formas para que eu pudesse me desenvolver completamente como pesquisador. Das noites discutindo e escrevendo os resultados dos muitos projetos que trabalhamos em conjunto, às conversas descontraídas e confidências trocadas, nos tornamos amigos e, por tudo isso, eu tenho um sentimento imenso de gratidão. Obrigado por vibrar com as minhas conquistas e por comprar as minhas ideias e trabalhar nelas comigo. Carrego dentro de mim muito amor e respeito por você!

À Debora por ter aceitado me coorientar nesta pesquisa e por me instigar a adentrar no campo da biodiversidade, que tanto gosto. Obrigado por me auxiliar no entendimento desse campo tão complexo que é a agrobiodiversidade e por me estimular a conquistar novos espaços. Quero muito que essa parceria continue no futuro, em um doutorado.

Ao meu pai, Eustáquio, e minha mãe, Lineia, por todo o apoio, amor e por compreenderem a minha ausência durante esse período em que estive totalmente dedicado à pesquisa. Sou muito grato aos valores que me ensinaram, principalmente perseverança e respeito ao próximo, que tanto me auxiliam no desenvolvimento da minha carreira. Amo vocês, que são os melhores pais do mundo!

Às minhas irmãs, Viviane e Thaís, fonte de inspiração e motivação. Gratidão por estarem sempre comigo e por me confortarem com seu amor, mesmo à distância!

À minha avó Elza, quem amo muito e que tanto me apoiou durante essa pesquisa. Obrigado por rezar por mim e por manter o carinho que sempre teve comigo, desde quando eu era criança e costuma correr pela sua casa.

Ao Edgar Quintanilha por ser parceiro e me apoiar nessa jornada, sempre com palavras de motivação e apoio. Obrigado por me levar para passear no meio do mato quando as coisas apertavam e por me acompanhar nas trilhas.

Agradeço a Raissa e Marília por serem as melhores amigas, colegas de classe e companheiras de cachaça dessa vida. Obrigado pelos cafés na universidade, pelas parcerias acadêmicas, pelas discussões conceituais que tanto me ajudaram a avançar com a pesquisa, e pelas muitas “*girls' night out*” que tanto me serviram de apoio emocional e divertimento.

No meu pacote de melhores pessoas encontradas nesse caminho do mestrado, também coloco minha querida amiga Jeane, um presente de Deus e sábia conselheira. Obrigado por me trazer luz nos momentos difíceis da pesquisa e por todo o apoio emocional e espiritual.

Aos amigos e colegas de turma Cristiano e Joana, que tornaram as aulas das disciplinas incrivelmente mais divertidas. Adoro vocês!

Aos meus queridos amigos Adalberto, André, Eliz, Pâmela, Sâmara, Washington, obrigado por serem companhias tão agradáveis e por me entenderem até quando eu mesmo tenho dificuldade de me entender.

Agradecimento especial a Átila Lúcio, Benildes Maculan, Bruno Berg, Debora Drucker, Denise Marques, Edgar Quintanilha, Vó Elza Soares, D. Fátima Rocha, Fernanda Resende, Júnia Flor, Leonardo de Castro, Lúcia de Távora Queiroz, Mariana Silva, Mauricio Almeida, Ronaldo Araújo e Sérgio Sevilhano por terem contribuído com a minha campanha de financiamento para participação no Biodiversity Next, tão importante para a minha pesquisa.

Agradeço muito pelas contribuições recebidas do Prof. Guilherme Ataíde, Profa. Célia Dias e Dr. Ivo Pierozzi durante a minha qualificação e em especial à Profa. Célia e ao Dr. Ivo, que aceitaram contribuir novamente com a minha pesquisa na defesa.

Ao Prof. Antonio Saraiva pela companhia agradável e produtiva durante o Biodiversity Next. Também por ter aceitado o meu convite para compor a minha banca de defesa.

A huge thanks to Jennifer Hammock, who has kindly answered all my questions concerning biodiversity metadata (I loved to see you in Leiden!).

Thank you, Karen Coyle, for being such a guiding light on my way gently answering all my questions about the simple version of the DCAM.

Ao Philip Russo por me auxiliar, com tanta paciência, a encontrar as respostas que eu tanto precisava no domínio da biodiversidade.

Não poderia deixar de agradecer o apoio recebido dos meus queridos colegas de trabalho, que se tornaram amigos, da Faculdade Minas Gerais (em ordem alfabética): Adriana, Heleninha, Nara, Thomas e Vivi.

À Profa. Gercina, por ter enxergado em mim o potencial para a pesquisa em 2014 e ter me indicado para ser bolsista de iniciação científica.

À Profa. Beatriz Cendón pelas aulas de metodologia científica muito esclarecedoras e estimulantes.

Às agências de fomento que contribuíram com a disponibilização de bolsas e ajudas de custo para a realização deste mestrado, incluindo as participações em eventos, em especial à Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e à Pró-Reitoria de Pós-Graduação (PRPG) da UFMG.

A arte de pensar sem riscos. Não fossem os caminhos da emoção a que leva o pensamento, pensar já teria sido catalogado como um dos modos de se divertir. Não se convidam amigos para o jogo por causa da cerimônia que se tem em pensar. O melhor modo é convidar apenas para uma visita, e, como quem não quer nada, pensa-se junto, no disfarçado das palavras.

Isso, enquanto jogo leve. Pois para pensar fundo – que é o grau máximo do hobby – é preciso estar sozinho. Porque entregar-se a pensar é uma grande emoção, e só se tem coragem de pensar na frente de *outrem* quando a confiança é grande a ponto de não haver constrangimento em usar, se necessário, a palavra *outrem*. Além do mais exige-se muito de quem nos assiste pensar: que tenha um coração grande, amor, carinho, e a experiência de também se ter dado ao pensar. Exige-se tanto de quem ouve as palavras e os silêncios – como se exigiria para sentir. Não, não é verdade. Para sentir exige-se mais. Bom, mas quanto a pensar como divertimento, a ausência de riscos o põe ao alcance de todos. Algum risco tem, é claro. Brinca-se e pode-se sair de coração pesado. Mas de um modo geral, uma vez tomados os cuidados intuitivos, não tem perigo. Como hobby, apresenta a vantagem de ser por excelência transportável. Embora no seio do ar ainda seja melhor, segundo eu. Em certas horas da tarde, por exemplo, em que a casa cheia de luz mais parece esvaziada pela luz, enquanto a cidade inteira estremece trabalhando e só nós trabalhamos em casa, mas ninguém sabe – nessas horas em que a dignidade se refaria se tivéssemos uma oficina de consertos ou uma sala de costuras – nessas horas: pensa-se. Assim: começa-se do ponto exato em que se estiver, mesmo que não seja de tarde; só de noite é que não aconselho.

Uma vez por exemplo – no tempo em que mandávamos roupa para lavar fora – eu estava fazendo o rol. Talvez por hábito de dar título ou por súbita vontade de ter caderno limpo como em escola, escrevi: rol de... e foi nesse instante que a vontade de não ser séria chegou. Este é o primeiro sinal do *animus brincandi*, em matéria de pensar – como – hobby. E escrevi esperta: rol de sentimentos. O que eu queria dizer com isto, tive que deixar para ver depois – outro sinal de se estar no caminho certo é o de não ficar aflita por não entender; a atitude deve ser: não se perde por esperar, não se perde por não entender.

[...] *Mas devo avisar.* Às vezes começa-se a brincar de pensar, e eis que inesperadamente o brinquedo é que começa a brincar conosco. Não é bom. É apenas frutífero.

Clarice Lispector. Brincar de pensar, 19 de agosto de 1967.

RESUMO

A informação é elemento intrínseco às relações humanas e, quando veiculada ou armazenada em meio digital, necessita ser bem descrita para que possa ser eficientemente recuperada, acessada e interpretada pela sociedade. Dentre os temas de interesse da sociedade na contemporaneidade, a agrobiodiversidade é um conceito amplo que envolve organismos e ecossistemas relacionados à produção agrícola e lavouras. Para sistemas computacionais, a representação da informação produzida sobre a agrobiodiversidade pode ser feita com metadados. Entretanto, os padrões de metadados existentes não contemplam plenamente a representação de alguns conceitos da agrobiodiversidade, a exemplo do que ocorre no padrão Darwin Core (DwC). O objetivo desta pesquisa foi criar princípios para a criação de uma extensão de metadados para o padrão DwC, tendo como recorte as interações ecológicas no contexto da agrobiodiversidade. Para atingir o objetivo da pesquisa, a metodologia, caracterizada como exploratória, qualitativa, aplicada e descritiva, foi dividida em duas etapas: 1) exploração dos insumos metodológicos e terminológicos; 2) definição terminológica e modelagem de metadados. A execução da primeira etapa da metodologia foi organizada em quatro subetapas: a) análise sistemática da literatura sobre interações ecológicas; b) análise do núcleo principal de termos do DwC; c) análise das extensões do padrão de metadados DwC; d) análise terminológica correlata das classes do modelo <onto.biodiversidade>, DwC e conceitos das interações ecológicas. A primeira subetapa teve como resultado um modelo conceitual sobre as interações ecológicas; a segunda subetapa teve como resultado a tradução das definições dos termos do núcleo principal do DwC e respectivas análises; a terceira subetapa apresentou uma síntese do conteúdo das extensões de metadados desenvolvidas para o DwC em outros projetos; e, por fim, a quarta subetapa consistiu em analisar de maneira correlacionada as classes do modelo <onto.biodiversidade>, que foi desenvolvido pela Embrapa com o intuito de organizar informações sobre a agrobiodiversidade, as classes de termos do DwC e de suas extensões e o modelo conceitual das interações ecológicas. A segunda etapa da metodologia teve como resultado três elementos de metadados, criados para representar a interação de parasitismo, representados como um registro de metadados em Extensible Markup Language (XML). Considera-se que a maior contribuição desta pesquisa foi apresentar um conjunto de princípios metodológicos para a criação de uma extensão de metadados para a representação das interações ecológicas no contexto da agrobiodiversidade, o que pode fomentar melhorias nas práticas agrícolas, que são importantes para toda a sociedade, assim como para o campo de estudos da Ciência da Informação.

Palavras-chave: Representação da informação. Metadados. Darwin Core. Agrobiodiversidade. Interações ecológicas.

ABSTRACT

Information is an intrinsic element to human relations. When transmitted or stored in digital media, it needs to be well described so it can be efficiently retrieved, accessed and interpreted by society. Among the themes of interest to society today, agrobiodiversity is a broad concept that involves organisms and ecosystems related to agricultural production and crops. For computational systems, the representation of information produced about agrobiodiversity can be done with metadata. However, the existing metadata standards do not fully include the representation of some concepts of agrobiodiversity, as in the Darwin Core (DwC) standard. The aim of this research was to present principles for the creation of an extension of metadata for the DwC standard, having as scope the ecological interactions in the context of agrobiodiversity. To achieve the research aim, the methodology, characterized as exploratory, qualitative, applied and descriptive, was divided into two stages: 1) exploration of methodological and terminological inputs; 2) terminological definition and metadata modeling. The execution of the first stage of the methodology was organized into four substeps: a) systematic analysis of the literature on ecological interactions; b) analysis of the main core of DwC terms; c) analysis of the extensions to the DwC metadata standard; d) correlated terminological analysis of the classes of the model <onto.biodiversity>, DwC and concepts of ecological interactions. The first substep resulted in a conceptual model on ecological interactions; the second substep resulted in the translation of the definitions of the terms of the main core of the DwC and respective analyzes; the third substep presented a summary of the content of the metadata extensions developed for DwC in other projects; and, finally, the fourth substep consisted of analyzing in a correlated manner the classes of the <onto.biodiversity> model, which was developed by Embrapa in order to organize information on agrobiodiversity, the DwC term classes and their extensions and the conceptual model of ecological interactions. The second stage of the methodology resulted in three metadata elements, created to represent the interaction of parasitism, represented as a metadata record in Extensible Markup Language (XML). It is considered that the greatest contribution of this research was to present a set of methodological principles for the creation of an extension of metadata for the representation of ecological interactions in the context of agrobiodiversity, which can foster improvements in agricultural practices, which are important for the entire community. society, as well as for the field of information science studies.

Keywords: information representation. Metadata. Darwin Core. Agrobiodiversity. Ecological interactions.

LISTA DE ESQUEMAS

Esquema 1 – Registro de metadados XML	68
Esquema 2 – Regra 5: usar códigos para metadados do DC e do DwC	93
Esquema 3 – Registro XML de metadados DwC com a propriedade <i>parasiteOf</i>	172
Esquema 4 – Alternativa de descrição da classe <i>Parasitism</i>	175

LISTA DE FIGURAS

Figura 1 – Riqueza de espécies do Brasil	29
Figura 2 – Teia alimentar.....	46
Figura 3 – Bússola de interação	48
Figura 4 – Exemplo de predatismo	49
Figura 5 – Exemplo de parasitismo.....	50
Figura 6 – Exemplo de competição por interferência direta	51
Figura 7 – Exemplo de competição por recurso	52
Figura 8 – Exemplo de competição intraespecífica	53
Figura 9 – Exemplo de neutralismo	54
Figura 10 – Exemplos de interações mutualísticas.....	55
Figura 11 – Exemplo de comensalismo	57
Figura 12 – Exemplo de amensalismo	58
Figura 13 – Exemplo de sociedade.....	58
Figura 14 – Exemplo de colônia	59
Figura 15 – Exemplo de organismo que pratica canibalismo	60
Figura 16 – As três partes de uma declaração.....	63
Figura 17 – Tabela de uma base de dados relacional	66
Figura 18 – A estrutura semântica do XML	67
Figura 19 – Registro de metadados da obra Mona Lisa	75
Figura 20 – Metadados de uma foto da Mona Lisa.....	76
Figura 21 – The DCMI resource model	78
Figura 22 – Componentes do nome DOI.....	79
Figura 23 – DCMI Description Set Model	80
Figura 24 – The DCMI vocabulary model.....	83
Figura 25 – Dublin Core Abstract Model simplificado	85
Figura 26 – Relação de triplas no RDF	86
Figura 27 – Regra 1: termo do Darwin Core como nome de campo.....	92
Figura 28 – Diagrama de relação de entidade do modelo de DSW.....	95
Figura 29 – Modelo Crow's foot notation.....	96
Figura 30 – Aplicabilidade das extensões	101
Figura 31 – Exemplo de aplicação de um código internacional de nomenclatura	113
Figura 32 – Estrutura da metodologia	115
Figura 33 – Classificação das interações ecológicas.....	119
Figura 34 – Classes gerais do DwC.....	155
Figura 35 – Termos subordinados às classes do DwC.....	156
Figura 36 – Ficha terminológica do termo <i>kingdom</i>	157

Figura 37 – Apresentação das extensões do DwC	157
Figura 38 – Termos da extensão	158
Figura 39 – Classes gerais do modelo <onto.biodiversidade>	159
Figura 40 – Classes modelo <onto.biodiversidade> e equivalentes no DwC.....	160
Figura 41 – Classes do modelo de representação para a Embrapa.....	162
Figura 42 – Registro de uma extensão autorizada	166
Figura 43 – Ficha terminológica de metadados das extensões.....	167

LISTA DE QUADROS

Quadro 1 – Os tipos de hábitos alimentares	46
Quadro 2 – Grade de interações ecológicas	47
Quadro 3 – Conjunto de atributos de um elemento do Dublin Core	73
Quadro 4 – Metadado do tipo <i>class</i>	89
Quadro 5 – Metadado do tipo <i>property</i>	90
Quadro 6 – Regras de uso do Padrão Darwin Core	91
Quadro 7 – Exemplo de aplicação da extensão de metadados do Dublin Core	98
Quadro 8 – Classe <i>Record-level</i>	122
Quadro 9 – Classe <i>Occurrence</i>	124
Quadro 10 – Classe <i>Organism</i>	127
Quadro 11 – Classe <i>MaterialSample</i>	128
Quadro 12 – Classe <i>Event</i>	130
Quadro 13 – Classe <i>Location</i>	131
Quadro 14 – Classe <i>GeologicalContext</i>	136
Quadro 15 – Classe <i>Identification</i>	137
Quadro 16 – Classe <i>Taxon</i>	138
Quadro 17 – Classe <i>MeasurementOrFact</i>	141
Quadro 18 – Classe <i>ResourceRelationship</i>	142
Quadro 19 – Classe <i>UseWithIRI</i>	143
Quadro 20 – Evidência de ocorrência.....	147
Quadro 21 – Estrutura do Apêndice A e do Apêndice B	149
Quadro 22 – Síntese do conteúdo das extensões estáveis do DwC	150
Quadro 23 – Síntese do conteúdo das extensões em desenvolvimento do DwC	152
Quadro 24 – Definição terminológica funcional	163
Quadro 25 – Classe <i>Parasitism</i>	169
Quadro 26 – Elemento <i>parasiteOf</i>	170
Quadro 27 – Elemento <i>hostOf</i>	176
Quadro 28 – Síntese dos princípios metodológicos para criação de uma extensão de metadados para o DwC	178

LISTA DE ABREVIATURAS E SIGLAS

5W1H	<i>Who</i> (quem), <i>what</i> (o que), <i>why</i> (porque), <i>where</i> (onde), <i>when</i> (quando), <i>how</i> (como)
AAAS	American Association for the Advancement of science
AACR2	Código de Catalogação Anglo-Americano
ABCD	Access to Biological Collections Data schema
AGRIS	International Information System for the Agricultural Sciences and Technology
AGROVOC	Agricultural Information Management Standards
ASFA	Aquatic Science & Fisheries Abstracts
ASP	Academic Search Premier
BI	Informática em Biodiversidade
BU-UFSC	Biblioteca Universitária da Universidade Federal de Santa Catarina
BZN	Bulletin of Zoological Nomenclature
CBD	Convention on Biological Diversity
CDD	Classificação Decimal de Dewey
CDU	Classificação Decimal Universal
CETAF	Consortium of European Taxonomic Facilities
CGen	Conselho de Gestão do Patrimônio Genético
CI	Ciência da Informação
CSV	Comma-separated values
DC	Dublin Core
DCAM	Dublin Core Abstract Model
DCMI Box	DCMI Box Encoding Scheme
DCMI	Dublin Core Metadata Initiative
DCMIType	Vocabulário de tipo DCMI
dcterms	Dublin Core Terms

DiSSCo	Distributed System of Scientific Collections
DOI	Digital Object Identifier
DSW	Darwin Core Semantic Web
DwC	Darwin Core
DwCA	Darwin Core Archive
DwCAB	Darwin Core Agrobiodiversity Extension
DwCMG	Darwin Core Maintenance Group
dwr	Simple Darwin Core
Embrapa	Empresa Brasileira de Pesquisa Agropecuária
EML	Ecological Metadata Language
ENVO	Environment Ontology
EPSG	European Petroleum Survey Group
ER	entidades e relacionamentos
FAO	Food and Agriculture Organization of the United Nations
GBIF	Global Biodiversity Information Facility
GGBN	Global Genome Biodiversity Network
GISIN	Global Invasive Species Information Network
GRBio	Global Registry of Biorepositories
IAPT	International Association for Plant Taxonomy
Ibama	Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis
Ibict	Instituto Brasileiro de Informação em Ciência e Tecnologia
ICBN	International Code of Botanical Nomenclature
ICMBio	Instituto Chico Mendes de Conservação da Biodiversidade
ICN	International Code of Nomenclature for Algae, Fungi and Plants
ICNCP	International Code of Nomenclature for Cultivated Plants
ICNP	International Code of Nomenclature of Prokaryotes
ICTV code	International Committee on Taxonomy of Viruses

ICZ	Congresso Internacional de Zoologia
ICZN ₁	International Commission on Zoological Nomenclature
ICZN ₂	International Code of Zoological Nomenclature
iDigBio	National resource for digitized information about vouchered natural history collections
IEC	International Electrotechnical Commission
IMT	Internet Assigned Numbers Authority
INPE	Instituto Nacional de Pesquisas Espaciais
IPBES	Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services
IRI	Internationalized Resource Identifier
ISHS	International Society for Horticultural Science
ISO	International Organization for Standardization
ISSN	International Standard Serial Number = Número Internacional de Série Padrão
IUBS	União Internacional de Ciências Biológicas
IUCN	International Union for Conservation of Nature
JSON	JavaScript Object Notation
LCC	Library of Congress Classification
LCSH	Library of Congress Subject Headings
LifeWatch ERIC	e-Science and Technology European Infrastructure for Biodiversity and Ecosystem Research
LSI	Identificadores das Ciências da Vida
MARC21	Machine Readable Cataloging
MBG	Marine BioGeography
MCPD	Multi Crop Passport Data standard
MESH	Medical Subject Headings
NBN	National Biodiversity Network
NISO	National Information Standards Organization

NLM	National Library of Medicine Classification
NXF	NBN eXchange Format
OBIS	Ocean Biogeographic Information System
OBO	Open Biological and Biomedical Ontologies
OECD	Organisation for Economic Co-Operation and Development
ONU	Organização das Nações Unidas
PCO	Population and Community Ontology
PDF	Portable Document Format
Period	DCMI Period Encoding Scheme
PGRFA	Recursos Genéticos de Plantas para Alimentação e Agricultura
Point	DCMI Point Encoding Scheme
PPGZoo	Programa de Pós-Graduação em Zoologia
PNB	Política Nacional de Biodiversidade
RDF	Resource Description Framework
RFC	Request for Comments
SGML	Standard Generalized Markup Language
SI	International System of Units
SiBBr	Sistema de Informação sobre a Biodiversidade Brasileira
SOC	Sistemas de Organização do Conhecimento
SRI	Sistema de Recuperação da Informação
SRS	Sistema de Referência Espacial
TDWG	Biodiversity Information Standards
TGN	Getty Thesaurus of Geographic Names
UFMG	Universidade Federal de Minas Gerais
UNCED	United Nations Conference on Environment and Development
UNESCO	Organização das Nações Unidas para Educação, Ciência e Cultura
URI	Uniform Resource Indicator

VLTS	Visual Line Transect Survey
W3C	World Wide Web Consortium
W3C-DTF	W3C Date and Time Formats Specification
WKT	Well-Known Text
WWF-BRASIL	World Wide Fund for Nature Brasil
XML	eXtensible Markup Language

SUMÁRIO

1	INTRODUÇÃO	20
1.1	Problema e justificativa	22
1.2	Objetivos	24
1.2.1	<i>Objetivo geral</i>	25
1.2.2	<i>Objetivos específicos</i>	25
1.3	Estrutura da pesquisa	25
2	FUNDAMENTOS TEÓRICO-METODOLÓGICOS	27
2.1	Biodiversidade	27
2.1.1	<i>Iniciativas para organização de informação sobre biodiversidade</i>	32
2.2	Agrobiodiversidade	38
2.2.1	<i>Interações ecológicas</i>	43
2.2.1.1	<i>Predatismo e parasitismo</i>	49
2.2.1.2	<i>Competição</i>	50
2.2.1.3	<i>Neutralismo</i>	53
2.2.1.4	<i>Mutualismo e protocooperação</i>	54
2.2.1.5	<i>Comensalismo</i>	56
2.2.1.6	<i>Amensalismo</i>	57
2.2.1.7	<i>Sociedade</i>	58
2.2.1.8	<i>Colônia</i>	59
2.2.1.9	<i>Canibalismo</i>	59
2.2.2	<i>Considerações sobre as interações ecológicas</i>	60
2.3	Representação da informação em metadados	61
2.3.1	<i>Bases de dados relacionais</i>	65
2.3.2	<i>eXtensible Markup Language (XML)</i>	66
2.3.3	<i>Linked data e RDF</i>	69
2.4	Princípios teóricos para a criação de metadados	70
2.5	Padrões de metadados da biodiversidade.....	87
2.6	Padrão de metadados Darwin Core	88
2.6.1	<i>Darwin Core Extensions</i>	97
2.7	Esquemas de codificação	103
2.7.1	<i>International Code of Zoological Nomenclature</i>	105
2.7.2	<i>International Code of Nomenclature for algae, fungi, and plants</i>	108
2.7.3	<i>International Code of Nomenclature for Cultivated Plants</i>	111
2.8	Considerações sobre os esquemas de codificação	113
3	METODOLOGIA.....	114
3.1	Primeira etapa: exploração dos insumos metodológicos e terminológicos .	114

3.2	Segunda etapa: definição terminológica e modelagem dos metadados	116
4	RESULTADOS E ANÁLISES	117
4.1	Primeira etapa: exploração dos insumos metodológicos e terminológicos .	117
4.1.1	<i>Definição do escopo de representação de dados da agrobiodiversidade</i>	117
4.1.2	<i>Análise sistemática da literatura sobre interações ecológicas.....</i>	118
4.1.3	<i>Análise do núcleo principal de termos do Darwin Core.....</i>	120
4.1.4	<i>Análise das extensões do padrão de metadados Darwin Core.....</i>	148
4.1.5	<i>Análise terminológica correlata das classes do modelo <onto.biodiversidade>, DwC e conceitos de interações ecológicas</i>	154
4.2	Segunda etapa: definição terminológica e modelagem dos metadados	163
4.3	Sistematização dos princípios para construção de uma extensão ao DwC..	177
5	CONSIDERAÇÕES FINAIS.....	182
	REFERÊNCIAS	189
	APÊNDICE A – LISTA DE EXTENSÕES ESTÁVEIS DO DARWIN CORE.....	207
	APÊNDICE B – EXTENSÕES EM DESENVOLVIMENTO.....	219
	APÊNDICE C – PESQUISA BIBLIOGRÁFICA	227
	APÊNDICE D – TRADUÇÃO DOS QUADROS 20, 21 E 22	231
	ANEXO A – MODELO <ONTO.BIODIVERSIDADE> ESTRUTURA GERAL	232
	ANEXO B – MODELO <ONTO.BIODIVERSIDADE> ESPÉCIMEN E ESPÉCIE .	233
	ANEXO C – MODELO <ONTO.BIODIVERSIDADE> CLASSE COLEÇÃO	234
	ANEXO D – MODELO <ONTO.BIODIVERSIDADE> ECOLOGIA.....	235
	ANEXO E – MODELO <ONTO.BIODIVERSIDADE> CLASSE TAXA	236
	ANEXO F – MODELO <ONTO.BIODIVERSIDADE> CLASSE COLETA.....	237
	ANEXO G – MODELO <ONTO.BIODIVERSIDADE> CLASSE ESPÉCIMEN	238
	ANEXO H – MODELO <ONTO.BIODIVERSIDADE> COLEÇÃO	239
	ANEXO I – MODELO <ONTO.BIODIVERSIDADE> CLASSE LOCALIZAÇÃO ..	240

1 INTRODUÇÃO

A Ciência da Informação (CI) tem como um de seus intentos garantir que as pessoas tenham acesso à informação, a partir do respaldo de teorias e metodologias que dão suporte aos estudos que seus pesquisadores desenvolvem. Em meio digital, torna-se cada vez mais imprescindível a representação do recurso informacional, visto que esta vai estabelecer uma ligação entre os dados e o recurso, e buscará ser uma codificação que atribui semântica aos dados representados. Esses signos de representação são empregados em processos de acesso, recuperação, compartilhamento e reutilização de conhecimento estruturado, estreitamente relacionados com a interação e a comunicação sociais.

A informação é, assim, representada, para que possa ser recuperada, e para isso há diversos modelos e ferramentas, dentre eles os metadados. Metadados são etiquetas que carregam trechos curtos, porém descritivos, de informação sobre determinado objeto, denominado dado. Sem os metadados para informar o que é o dado, torna-se difícil entender a semântica que o dado representa. Considere-se, por exemplo, os campos de um formulário de endereço: rua, bairro, número, CEP, estado, país; são todos metadados. A eficiência dos Sistemas de Recuperação da Informação (SRIs) depende de metadados bem construídos, que permitam a construção de registros sobre o recurso informacional que se deseja indexar no SRI.

Metadados são criados para diversos fins e áreas do conhecimento, como para a biodiversidade e agricultura. A agricultura envolve um conjunto de fatores ambientais e biológicos que estão intrinsecamente relacionados. São fatores biológicos os seres vivos e as funções desempenhadas por eles nos sistemas agroecológicos. As interações entre os seres vivos nas lavouras podem contribuir para o aumento da produção de alimentos ou prejudicá-la. A polinização, por exemplo, é o tipo de interação essencial para a reprodução dos vegetais, e acontece por meio de uma relação harmônica entre polinizadores e as plantas; as lagartas que atacam as plantações, por outro lado, são exemplos de relações desarmônicas em lavouras e que podem prejudicar a produção de alimentos.

Pesquisas são constantemente realizadas em instituições como a Empresa Brasileira de Pesquisa Agropecuária (Embrapa), visando aprimorar técnicas de manejo que

contribuam para o aumento sustentável da produção agrícola, para atender a uma demanda crescente por alimentos.

Esta pesquisa é desenvolvida com suporte e colaboração da Universidade Federal de Minas Gerais (UFMG) e da Embrapa, tendo como norte o desenvolvimento de uma extensão de metadados para o padrão Darwin Core (DwC), de maneira que atenda à representação da informação sobre a agrobiodiversidade brasileira.

O primeiro padrão de metadados voltado à descrição de recursos na web foi criado em 1995, o Dublin Core (DC). O núcleo básico do DC possui 15 elementos que permitem descrever de forma simples (a definição dos elementos é, supostamente, de fácil entendimento para a maioria das pessoas) e eficiente (supostamente adequado para a representação de diversos tipos de recursos digitais e analógicos).

Os 15 elementos do DC são muito genéricos e não possibilitam descrições exaustivas do conteúdo dos recursos informacionais. Por esse motivo, é permitido criar extensões ao padrão de metadados e, também, criar novos padrões de metadados. No domínio da Biodiversidade foram criados alguns padrões de metadados para a organização de coleções biológicas, dados de pesquisas ecológicas e informações sobre espécies que são mundialmente utilizados como: Access to Biological Collections Data schema (ABCD), DwC, Ecological Metadata Language (EML) e Plinian Core.

O DwC é o esquema de metadados para descrição de ocorrências de espécies. Permite representar, principalmente, informações sobre “taxa, sua ocorrência na natureza conforme documentado por observações, espécimes, amostras e informações relacionadas” (DARWIN CORE TASK GROUP, 2009c, *online*, tradução do proponente). Possui um núcleo básico denominado Darwin Core terms contendo 176 elementos, “com o objetivo de facilitar o compartilhamento de informações sobre a diversidade biológica, fornecendo definições de referência, exemplos e comentários” (SISTEMA DE INFORMAÇÃO SOBRE A BIODIVERSIDADE BRASILEIRA, 2018, *online*).

O DwC tem como um de seus princípios básicos a flexibilidade, que permite a extensão do padrão de metadados para atender a contextos específicos. Extensões de metadados, que são conjuntos de metadados criados em complemento a um núcleo básico de metadados, têm sido desenvolvidas para complementar a

representação de dados de biodiversidade de campos específicos que não podem ser representados pelos metadados do DwC (as extensões do DwC aprovadas e em desenvolvimento estão listadas nos Apêndices A e B, respectivamente). A agrobiodiversidade é um campo específico da biodiversidade que produz informações que não podem ser completamente representadas pelos elementos do DwC.

Ainda que a maioria dos dados produzidos pela agrobiodiversidade possam ser descritos por metadados presentes no DwC, alguns dados ecológicos, como aqueles provenientes da determinação de relações ecológicas entre espécies, não podem ser representados pelo DwC.

Dessa forma, destaca-se como elemento norteador desta pesquisa a inexistência de metadados para a representação da informação sobre as interações ecológicas no contexto da agrobiodiversidade.

1.1 Problema e justificativa

As interações ecológicas são relações entre os seres vivos em um ecossistema para obtenção de alimento, abrigo, reprodução ou outro motivo, que pode ser benéfica para ambos os indivíduos envolvidos ou apenas para um (CASSINI, 2005). O estudo das interações ecológicas é importante no contexto da agrobiodiversidade, uma vez que certas espécies podem ser utilizadas para controlar a população de pragas em lavouras, por exemplo.

Os dados sobre interações ecológicas são apontados no *Final Report of the Task Group on GBIF Data Fitness for Use in Agrobiodiversity* como recurso importante para a agricultura e pecuária por oferecerem oportunidades em potencial para a “evolução da diversidade na agricultura e adaptação das culturas, e identificar compensações e riscos em intervenções para conservação, gestão de terras ou ações de restauração” (ARNAUD *et al.*, 2016, p. 23, tradução do proponente). Os autores acrescentam que esses dados podem ser, também, valiosos indicadores sobre os impactos da presença de gado em ambientes, pragas, doenças e demarcar a presença de insetos úteis, como polinizadores, ou sobre predadores de determinadas pragas que possam ser empregados como controle biológico.

A biodiversidade é um recurso que, se manejado corretamente, pode contribuir para

melhorar o desempenho das práticas agropecuárias. A Empresa Brasileira de Pesquisa Agropecuária (Embrapa) desenvolve tecnologias e sistemas para fomentar a melhoria das práticas agropecuárias. Em 2012, o pesquisador da Embrapa Informática Agropecuária, Pierozzi Junior (2012), desenvolveu um modelo conceitual para a representação da informação sobre a biodiversidade, abordando, dentre outras categorias: espécie, espécime, coleção, georreferenciamento e taxonomia, denominado <onto.biodiversidade>. O modelo representa um esquema de conceitos que representam espécimes coletados e permite relacioná-lo a conceitos comportamentais das espécies, como hábitos alimentares e funções ecológicas, e está distribuído entre os Anexos A e I.

Uma análise preliminar mostrou que o modelo <onto.biodiversidade> possui um esquema geral de representação da informação sobre a biodiversidade, possuindo diversas categorias comuns com os conceitos que compõem o vocabulário do DwC. Entretanto, observou-se que o DwC não possui metadados acerca das interações ecológicas no contexto da agrobiodiversidade. Nesse caso, Zermoglio (2018, *online*, tradução do proponente) sugere “propor e testar uma nova extensão para resolver um problema em particular (isso é feito o tempo todo)”. É permitido desenvolver extensões ao padrão DwC e isso já foi feito em outros campos de estudo específicos, como pode ser conferido nos Apêndices A e B. Um levantamento na literatura demonstrou que não existe qualquer proposta de extensão ao DwC para o caso específico da agrobiodiversidade brasileira.

Uma extensão de metadados para a agrobiodiversidade pode contribuir para a gestão de dados das interações ecológicas entre espécies e facilitar o acesso ao conhecimento sobre espécies polinizadoras, parasitos, predadores e outras funções ecológicas com potencial para melhorar as práticas agrícolas em lavouras afetadas por pragas e doenças, por exemplo. Desta maneira, esses dados são relevantes para a Embrapa, pois configuram-se como recurso de conhecimento com potencial para aprimorar práticas agrícolas. Por conseguinte, apresenta-se como pressuposto desta pesquisa: é possível definir o escopo de representação dos dados da agrobiodiversidade, notadamente das interações ecológicas, a partir do modelo conceitual apresentado no <onto.biodiversidade>, e, tendo ele como escopo terminológico, propor princípios para a construção de metadados sobre a agrobiodiversidade, tendo em vista a criação uma extensão para o padrão de

metadados DwC.

No desenvolvimento desta pesquisa, pretende-se responder à seguinte pergunta **Quais princípios são necessários adotar para a criação de uma extensão de metadados que siga as especificações sintáticas e semânticas do DwC, de maneira que possa representar os conceitos da agrobiodiversidade brasileira, especificamente das interações ecológicas?**

A escolha pelo termo “princípios” em lugar do termo “diretrizes”, ou outro termo de conceito semelhante, se relaciona com a etimologia da palavra, que vem do latim *principium* e *primus*, que significa origem ou início de algo, causa próxima ou primária, uma proposição fundamental ou algum conhecimento que ainda pode ser questionado por estar em seu estágio inicial. Tem também por base a definição segundo o Michaelis (PRINCÍPIO, 2020, *online*), que determina que o termo “princípios”, no plural, é, “em uma área de conhecimento, conjunto de proposições fundamentais e diretivas que servem de base e das quais todo desenvolvimento posterior deve ser subordinado”, servindo de orientações para a compreensão de algo, a partir de “suas propriedades essenciais ou características” (PRINCÍPIO, 2020, *online*). Considera-se que o conceito do termo “diretrizes” está mais relacionado com o conteúdo de uma norma, que estabelece, em geral, um conhecimento de construção conjunta. No caso desta pesquisa, o conceito de princípios indica um sentido mais restrito, pois apresenta uma proposição inicial, de acordo com um dado raciocínio, não estando ainda devidamente consolidado para ser considerado diretrizes estabelecidas em uma norma.

Desta maneira, nesta pesquisa foram utilizados os termos *requisitos* e *recomendações* para descrever as diretrizes propostas em normas para a construção de metadados, e o termo *princípios* para descrever os recursos metodológicos apresentados como resultado desta pesquisa.

1.2 Objetivos

A partir da problematização aventada, pretende-se atingir o objetivo geral da pesquisa, que é sustentado pelos objetivos específicos.

1.2.1 *Objetivo geral*

Propor princípios para a criação de uma extensão de metadados para o padrão DwC, com vistas ao atendimento das necessidades de representação da informação sobre interações ecológicas no contexto da agrobiodiversidade brasileira.

1.2.2 *Objetivos específicos*

Pretende-se com esta pesquisa, de forma mais específica:

- a) conhecer o padrão DwC e suas extensões para entender como os dados sobre a biodiversidade são representados, tendo em vista estabelecer uma correlação entre esses dados e a estrutura semântica do DwC;
- b) conhecer os princípios do DCMI e das recomendações da ISO/IEC 11179-4 (2004) para subsidiar a determinação dos princípios para criação de metadados para o DwC;
- c) examinar a literatura sobre interações ecológicas para compreender os conceitos e sua relação com o campo da agrobiodiversidade, com vistas a verificar a possibilidade de uso do padrão DwC para a descrição de dados dessa temática;
- d) contribuir para as práticas agrícolas sustentáveis ao propor princípios que podem ser empregados na representação de informação sobre a agrobiodiversidade, na forma de metadados.

1.3 *Estrutura da pesquisa*

Para elucidar a organização desta pesquisa, aplica-se a ferramenta 5W2H¹:

- a) **QUEM (WHO)**: padrão de metadados DwC aplicado à representação da informação para atender às demandas de clientes que trabalham com aplicações sobre a agrobiodiversidade brasileira, como produtores rurais;
- b) **O QUE (WHAT)**: estudar o modelo <onto.biodiversidade>, que relaciona

¹ *Who* (quem), *what* (o que), *why* (porque), *where* (onde), *when* (quando), *how* (como), *how much* (quanto).

conceitos da biodiversidade a agrobiodiversidade, e os aspectos sintáticos e semânticos do padrão de metadados DwC para propor princípios para a criação de uma extensão de metadados para a agrobiodiversidade, especificamente as interações ecológicas (ver objetivos, Seção 1.2);

- c) **PORQUE (WHY)**: inexistência de um modelo de metadados para representação da informação da agrobiodiversidade, especificamente as interações ecológicas; carência de metadados no padrão DwC que representem conceitos específicos da agrobiodiversidade (ver problema e justificativa, Seção 1.1);
- d) **ONDE (WHERE)**: recorte na agrobiodiversidade, notadamente nas interações ecológicas;
- e) **QUANDO (WHEN)**: a pesquisa teve início em março de 2018 e término em dezembro de 2019;
- f) **COMO (HOW)**: aplicação dos princípios para construção de metadados do DCMI Abstract Model (DCAM), recomendações da ISO/IEC 11179-4 (2004) e características semânticas e sintáticas do DwC (ver Metodologia, Seção 3);
- g) **QUANTO (HOW MUCH)**: o esforço de imersão no campo da biodiversidade, por meio de formação acadêmica na área de biologia, literatura especializada, visitas a laboratório e participação em conferência internacional na área.

2 FUNDAMENTOS TEÓRICO-METODOLÓGICOS

Nesta seção estão descritos e discutidos, primeiramente, os conceitos que envolvem a temática da biodiversidade e de seu componente, a agrobiodiversidade, tendo em vista o alinhamento do conhecimento sobre as questões que perpassam o tratamento desse tipo de informação. Em seguida são trazidos conceitos de representação da informação, notadamente na forma de metadados, relevantes para o entendimento dos procedimentos que foram executados na pesquisa, ou seja, o desenvolvimento da metodologia. No âmbito dos metadados, são descritos os princípios teóricos para a sua criação, as recomendações de melhores práticas para a criação de metadados, as características semânticas, sintáticas e estruturais do padrão de metadados DwC e de suas extensões já publicadas, assim como outros esquemas de codificação aplicados no campo da biodiversidade e que podem ser relevantes na pesquisa sobre a agrobiodiversidade.

2.1 Biodiversidade

A United Nations Conference on Environment and Development (UNCED) ou The Earth Summit in Rio de Janeiro, como é popularmente conhecida em inglês, e Rio-92, em português, aconteceu em 1992 no Rio de Janeiro e teve como resultado tratados importantes: Agenda 21, Rio Declaration on Environment and Development, Statement of Forest Principles, the United Nations Framework Convention on Climate Change e a United Nations Convention on Biological Diversity (EARTH SUMMIT+5, 1997). Dentre esses documentos, a Convention on Biological Diversity (CBD) apresentou a definição oficial sobre o conceito de diversidade biológica, que

significa a variabilidade de organismos vivos de todas as origens, compreendendo, dentre outros, os ecossistemas terrestres, marinhos e outros ecossistemas aquáticos e os complexos ecológicos de que fazem parte, compreendendo ainda a diversidade dentro de espécies, entre espécies e de ecossistemas. (CBD, 1992, p. 3, tradução do proponente²).

O termo “biodiversidade” foi cunhado com base no conceito de diversidade biológica, e, antes de ser definido oficialmente pela CBD, foi considerado sinônimo de riqueza

² "Biological diversity means the variability among living organisms from all sources including, inter alia, terrestrial, marine and other aquatic ecosystems and the ecological complexes of which they are part; this includes diversity within species, between species and of ecosystems" (CBD, 1992, p. 3).

de espécies. Desde 1982, esse conceito já havia ampliado o seu significado para diversidade genética, e, em seguida, após 1986, expandiu seu sentido também para a diversidade ecológica (ou dos ecossistemas) (HESSE, 2008).

A diversidade genética é “a variabilidade presente no conjunto de indivíduos da mesma espécie”, tornando-o único entre os seus pares (HESSE, 2008, p. 23). O autor esclarece que a diversidade genética compreende a capacidade dos seres vivos em se adaptar às novas condições ambientais, que são dinâmicas, para sobreviverem. A diversidade de espécies “é responsável pela manutenção de uma série de serviços que a natureza nos presta: polinização, ciclagem de nutrientes, conservação de solos e controle de pragas e doenças”, entre outras ações (HESSE, 2008, p. 23-24). Segundo o autor, a diversidade genética e de espécies é fundamental para a diversidade de ecossistemas, uma vez que engloba “comunidades de animais, plantas e microrganismos, cada uma delas com interações diferentes e características; processos que conectam os seres vivos, tanto uns aos outros como também ao meio físico que os circunda” (HESSE, 2008, p. 24-25).

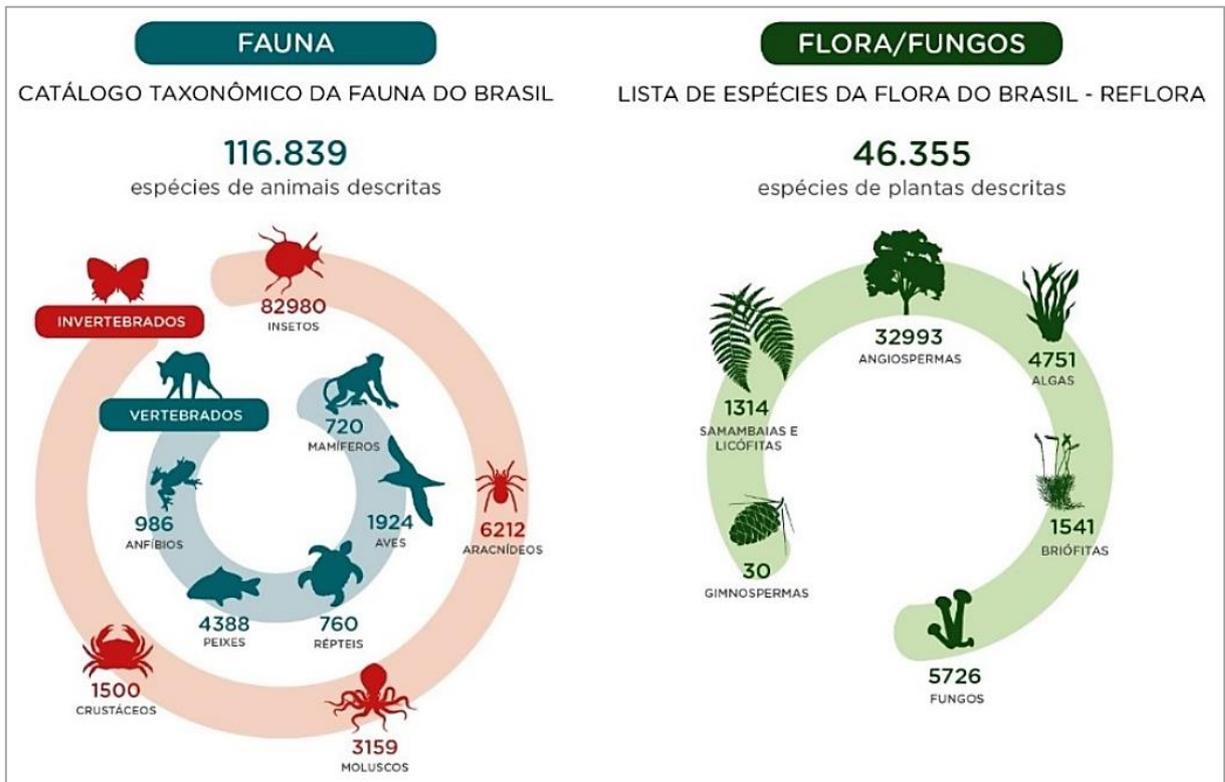
Apesar de sua importância para a manutenção da vida na terra, a diversidade biológica tem sido degradada em ritmo alarmante nas últimas décadas. A Avaliação global realizada pelo Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES, 2019) concluiu que em torno de 25% das espécies animais e vegetais do planeta encontram-se ameaçadas, o que significa que em torno de um milhão de espécies já estão em processo de extinção.

O Brasil é um dos 17 países megadiversos, de acordo com a classificação da United Nations Educational, Scientific and Cultural Organization (UNESCO, 2017) e informações do Ministério do Meio Ambiente (2019).

O Catálogo Taxonômico da Fauna do Brasil indica 116.839 espécies animais já registradas, entre vertebrados e invertebrados, e a Lista de Espécies da Flora do Brasil já conta com 46.355 espécies registradas. No país, ocorre o maior número de espécies de plantas do mundo, das quais mais de 40% são endêmicas, sendo que o grupo das angiospermas possui maior representatividade de endemismo, com 56%. A cada dia, novas espécies são descobertas e descritas no Brasil, o que torna possível afirmar que esses números sejam ainda mais elevados (MINISTÉRIO DO MEIO AMBIENTE, 2019, *online*).

A Figura 1 ilustra a riqueza dessas espécies.

Figura 1 – Riqueza de espécies do Brasil



Fonte: Ministério do Meio Ambiente (2019)

Infelizmente, essa riqueza é ameaçada pela extinção de espécies da biodiversidade, que cresceu com a superexploração de habitats e recursos naturais nas últimas décadas, especialmente no Brasil.

Os fatores de perda de biodiversidade podem ser divididos em cinco categoriais, de acordo com informações do IPBES (BALVANERA, 2019), órgão vinculado à ONU:

- conversão de terras para usos diversos (como agricultura, pecuária e expansão urbana) e degradação dos oceanos provocadas por atividades como a pesca e aquacultura;
- extração de recursos vivos (como a pesca, caça e extração madeireira) e de recursos não-vivos (como a mineração e extração de petróleo);
- poluição do ar, solos e cursos hídricos;
- espécies exóticas invasoras que, sem predadores naturais, se instalam nos habitats e provocam o colapso das espécies nativas; o aumento das espécies exóticas invasoras nos ecossistemas globais se deve a alguns fatores:

“expansão das redes comerciais, maior mobilidade humana, contínua degradação do habitat e mudanças climáticas” (BALVANERA, 2019, p. 119);

e) mudança do clima do planeta.

Esses fatores têm causado a extinção de espécies na Mata Atlântica, por exemplo, que é o bioma³ brasileiro com maior quantidade de espécies de fauna e flora por área, e é também um dos mais ameaçados; aproximadamente 61% das espécies de fauna ameaçadas de extinção no Brasil pertencem a esse bioma, juntamente com 1.544 espécies de flora ameaçadas, o que torna a Mata Atlântica o bioma brasileiro com o maior número de espécies ameaçadas (MARQUES *et al.*, 2016). De acordo com Marques *et al.* (2016), 70% da população brasileira vive em regiões de ocorrência de Mata Atlântica, resultado de anos de devastação e ocupação de áreas de floresta, que atualmente tem sua área original reduzida a um percentual entre 11% e 16%.

Outros biomas no Brasil também enfrentam ameaças. O Cerrado, bioma que “[...] abriga 10.400 espécies de plantas, das quais 50 são endêmicas, ou seja, só ocorrem nesta região [...]”, assim como importante diversidade de fauna, “[...] incluindo 180 espécies de répteis, 113 espécies de anfíbios, 837 de pássaros e 195 espécies de mamíferos”, já teve mais de 45% de sua área original modificada para produção agropecuária, de acordo com dados da World Wide Fund for Nature Brasil (WWF-BRASIL, 2018, *online*). Dados do Instituto Sociedade, População e Natureza (2018, *online*) também confirmam esse problema:

o Cerrado enfrenta fortes pressões humanas que põem seu futuro em risco. O desmatamento no bioma é alarmante, chegando a 1,5% ao ano, ou seja, três milhões de hectares/ano, conforme as estimativas mais conservadoras. Isso equivale a 2,6 campos de futebol desmatados a cada minuto. O desmatamento no Cerrado, portanto, é maior que na Amazônia.

A Amazônia é outro ambiente que enfrenta grande perda de biodiversidade, o que provoca impactos ao nível global devido à influência desse ecossistema no clima do

³ “Um bioma é uma área do planeta que pode ser classificada de acordo com as plantas e animais que a habitam. [...] Um bioma é diferente de um ecossistema. Um ecossistema é a interação de coisas vivas e não vivas em um ambiente. Um bioma é uma área geográfica específica notável pelas espécies que vivem lá. Um bioma pode ser composto de muitos ecossistemas. Por exemplo, um bioma aquático pode conter ecossistemas como recifes de corais e florestas de algas. (RUTLEDGE *et al.*, 2011, *online*).

planeta⁴.

Os ecossistemas da Amazônia abrigam cerca de 10 a 15% da biodiversidade terrestre; sua precipitação abundante de cerca de 2,2 m·y – 1 torna a região uma importante fonte de calor para a atmosfera, gerando uma estimativa de 210.000 m³·s⁻¹ a 220.000 m³·s⁻¹ de descarga fluvial, o que é 15% da água doce entrada nos oceanos; armazena cerca de 150 bilhões a 200 bilhões de toneladas de carbono; e apresenta um mosaico de diversidade etnológica e linguística (NOBRE *et al.*, 2016, p. 10759, tradução do proponente).

Logo, a devastação dos biomas brasileiros provoca impactos ambientais, sociais e econômicos. A degradação biótica, por exemplo, implica em perda de possibilidades de uso sustentável de recursos derivados de espécies em inovações no setor de biotecnologia, indústria farmacêutica/química, e até mesmo na agropecuária (cujas atividades estão entre os principais responsáveis pela degradação ambiental), devido à “possibilidade de obtenção de material genético nativo resistente ou de maior produtividade” (MAY; VEIGA NETO; POZO, 2000, p. 18).

O uso determina o valor econômico da biodiversidade. A extração de madeira, caça e pesca, por exemplo, são atividades de uso direto da biodiversidade, pois possuem valor econômico imediato. O uso indireto da biodiversidade refere-se aos benefícios e produtos oferecidos por um ecossistema equilibrado, como proteção da bacia hidrográfica, redução da erosão dos solos, ciclo de nutrientes, polinização, controle climático e outros benefícios. O valor econômico do uso indireto da biodiversidade deve ser maior que do uso direto, entretanto sociedades em desenvolvimento como o Brasil não são eficientes em afirmar esse valor como superior (PEARCE; MORAN, 1994).

Com o objetivo de reduzir os impactos da ação humana sobre o planeta, a Organização das Nações Unidas (ONU) lançou uma série de ações para promover o desenvolvimento sustentável. Uma delas é a Agenda 2030, que estabelece 17 objetivos de desenvolvimento sustentável a serem seguidos por todas as nações do mundo. Dentre esses objetivos, no que tange especificamente à biodiversidade terrestre, o objetivo 15 da Agenda é “Proteger, recuperar e promover o uso sustentável dos ecossistemas terrestres, gerir de forma sustentável as florestas, combater a

⁴ Posição atual que tem consenso científico de cerca de 97% dos pesquisadores que publicam sobre o tema, segundo dados da National Aeronautics and Space Administration (NASA), no documento Scientific consensus: Earth's climate is warming (NASA, 2019).

desertificação, deter e reverter a degradação da terra e deter a perda de biodiversidade” (NAÇÕES UNIDAS BRASIL, 2015, *online*). Para cumprir tais metas, o Brasil deve desenvolver mecanismos de gestão da informação sobre a biodiversidade, que integrem dados científicos e técnicos sobre as espécies disponíveis em sistemas de informação bem estruturados. É importante que estes sistemas de informação disponibilizem informações verdadeiras e verificadas por especialistas, de forma a garantir segurança aos usuários que os utilizam.

2.1.1 Iniciativas para organização de informação sobre biodiversidade

Para produzir e gerir informações sobre a biodiversidade é necessário classificar e descrever as espécies que são observadas em seus habitats ou em ambientes controlados (laboratórios). A classificação facilita o estudo das espécies e gera repositórios de conhecimento sobre a biodiversidade. A descrição taxonômica das espécies e outros dados, como data de registro do evento, local, comportamento, sexo, número de indivíduos, entre outros, são combinados de forma a compor o que é chamado registro de ocorrência.

Diversas organizações desenvolvem iniciativas ao redor do mundo para inventariar a biodiversidade global e tornar público o conhecimento sobre ela, tal como a International Union for Conservation of Nature (IUCN), criada em 1948, que foi a primeira organização global para preservação do meio ambiente (IUCN, 2018).

Iniciativas como a da IUCN geram repositórios de conhecimento sobre espécies, que são fontes de informação importantes para a preservação e conservação da biodiversidade, pois permitem que pesquisadores, governos e cidadãos de diversas partes do mundo tenham acesso a informações atualizadas para pesquisa e iniciativas comuns (entre governo e sociedade civil), visando à preservação das espécies e dos ecossistemas. Dentre as ferramentas desenvolvidas pela IUCN para sustentar as ações de preservação do meio ambiente, está a lista de espécies ameaçadas de extinção, conhecida como IUCN Red List of Threatened Species™, “fonte de dados global mais abrangente sobre espécies ameaçadas de extinção” (IUCN, 2018, *online*, tradução do proponente).

Bases internacionais como a da IUCN utilizam padrões de metadados para representar as informações sobre espécies, o que permite a interoperabilidade com

outros sistemas de informação e importar dados sobre espécies de repositórios locais. No Brasil, muitos sistemas de informações sobre a biodiversidade não fazem uso de padrão de metadados, o que impede o rastreamento de dados por outros sistemas de informação. Segundo Canhos (2003, p. 45),

A situação da biodiversidade neste início do século XXI pode ser caracterizada pela combinação de processos acelerados de destruição de ecossistemas primários associados a esforços mobilizadores para a conservação e uso sustentável da biodiversidade e de grandes avanços em tecnologia de informação e comunicação de dados.

Esses avanços tecnológicos têm produzido ferramentas para gestão da informação sobre biodiversidade, que envolvem desde programas de computador (software) a padrões e protocolos para transferência de dados, o que, segundo Canhos (2003) e Santos *et al.* (2011), dá forma a um novo campo de estudo denominado informática para biodiversidade.

A informática para biodiversidade (BI) desempenha um papel importante nos ajudando a conhecer, proteger e usar a biodiversidade de maneira sustentável. Abrange atividades desde a digitalização, padronização, compartilhamento e agregação de dados até o suporte à tomada de decisões e políticas (SARAIVA, 2019, p. e37373, tradução do proponente).

A preocupação com a publicação de informações sobre a biodiversidade brasileira não é recente. Em 2002 foi decretada a Política Nacional da Biodiversidade. Entre suas disposições, algumas estão diretamente relacionadas à publicação de informações sobre a biodiversidade:

10.1.12. Promover a recuperação e a síntese das informações existentes no acervo científico brasileiro, principalmente teses e dissertações.

10.1.13. Promover o mapeamento da biodiversidade em todo o território nacional, gerar e distribuir amplamente mapas da biodiversidade brasileira, resguardando-se o devido sigilo de informações de interesse nacional (BRASIL, 2002, *online*).

O mapeamento da biodiversidade consiste em registrar eventos de ocorrência de espécies. As espécies existem em determinados locais devido às relações que estabelecem com os ecossistemas e com outras espécies. Esses eventos registrados podem ser publicados das mais diversas maneiras, sendo as mais comuns na forma de artigo científico, tese e dissertação. Uma ocorrência é uma evidência da existência

de “uma espécie (ou outro táxon) em um lugar particular em um momento específico” (GBIF, 2018, *online*).

Posteriormente, em 20 de maio de 2015, foi estabelecido o Marco Legal da Biodiversidade, lei nº 13.123, que, dentre as disposições sobre bens, direitos e obrigações, apresenta importantes premissas relativas ao acesso ao conhecimento sobre a biodiversidade nacional, a saber (BRASIL, 2015, *online*):

I - ao acesso ao patrimônio genético do País, bem de uso comum do povo encontrado em condições *in situ*, inclusive as espécies domesticadas e populações espontâneas, ou mantido em condições *ex situ*, desde que encontrado em condições *in situ* no território nacional, na plataforma continental, no mar territorial e na zona econômica exclusiva;

II - ao conhecimento tradicional associado ao patrimônio genético, relevante à conservação da diversidade biológica, à integridade do patrimônio genético do País e à utilização de seus componentes;

III - ao acesso à tecnologia e à transferência de tecnologia para a conservação e a utilização da diversidade biológica;

IV - à exploração econômica de produto acabado ou material reprodutivo oriundo de acesso ao patrimônio genético ou ao conhecimento tradicional associado;

VII - à implementação de tratados internacionais sobre o patrimônio genético ou o conhecimento tradicional associado aprovados pelo Congresso Nacional e promulgados.

Esta lei oficializa a criação do Conselho de Gestão do Patrimônio Genético (CGen), órgão que passa a ser responsável por regulamentar a sua implementação. Uma das obrigações do CGen é coordenar a produção e o compartilhamento de conhecimentos sobre a diversidade genética nacional, atendendo a “critérios para a criação de banco de dados para o registro de informação sobre patrimônio genético e conhecimento tradicional associado” (BRASIL, 2015).

A informação genética faz parte do grande *corpus* de informação sobre a biodiversidade. Essas informações devem estar disponíveis para acesso público e, para que sejam recuperadas, é necessário que sejam criados coleções e repositórios de conhecimento. O problema da recuperação da informação sobre a biodiversidade já foi relatado por diversos autores, como Canhos (2003, p. 45):

Apesar dos avanços decorrentes da implementação da Convenção

sobre a Diversidade Biológica e dos esforços de governos e sociedade civil, a base de conhecimento sobre a diversidade biológica é incipiente e desagregada. As informações existentes estão dispersas e, via de regra, inacessíveis. Temos uma enorme quantidade de dados em bibliotecas tradicionais (papel) ou em bases de dados digitais não integradas. Temos um volume enorme de informações associadas a amostras (espécimes) de material biológico depositadas em herbários e coleções zoológicas.

Segundo Rodrigues *et al.* (2010), os repositórios surgem neste cenário de necessidade de organização e curadoria de dados científicos, por meio de esforços de atores de diferentes setores, privado, acadêmico e governamental. Como fontes de informação científica, os repositórios devem apresentar:

solução segura e amigável para armazenamento de grandes volumes de dados e para a sua partilha controlada; b) infraestrutura sustentável para publicação e preservação a longo termo de dados de investigação; c) instruções para questões práticas relacionadas com a gestão dos dados ao longo do seu ciclo de vida; d) orientações no que se refere à publicação de dados e à sua preservação; e) apoio (competências, ferramentas e normas) durante todo o ciclo de vida da investigação, tendo em vista o armazenamento, a preservação e a garantia de que os dados recolhidos poderão vir a ser acedidos e reutilizados (RODRIGUES *et al.*, 2010, p. 30).

Para cumprir tais funções, deve haver, por trás dos repositórios, curadoria de dados acurada. Essa curadoria deve ter como “meios de apoio e suporte [o] [...] ponto de vista político, legal e ético, relacionados com o acesso e reutilização dos dados científicos” (RODRIGUES *et al.*, 2010, p. 34). Para que o armazenamento eficiente e reuso de dados sejam possíveis, os repositórios contam com “infraestrutura técnica, como sistemas, normas e protocolos, necessária para garantir a recolha, a preservação e o acesso aos dados” (RODRIGUES *et al.*, 2010, p. 34). A infraestrutura técnica diz respeito, basicamente, a tecnologias e plataformas digitais, como o software DSpace, ao passo que as normas e protocolos são os padrões de metadados e de interoperabilidade (RODRIGUES *et al.*, 2010).

Uma questão que aparenta estar ainda pouco desenvolvida nas diversas iniciativas relativas aos repositórios de dados científicos é a normalização dos metadados. Esta situação não é de estranhar, atendendo a que, por um lado, falta ainda resolver muitos dos problemas básicos, e por outro é de prever que seja muito difícil uniformizar a descrição entre domínios científicos (RODRIGUES *et al.*, 2010, p. 27).

A necessidade de normalização dos metadados apontada pelos autores se aplica ao

campo das coleções de espécimes, de forma que as informações possam ser eficientemente recuperadas e rastreadas por outros sistemas de informação.

O parágrafo I do Art. 3º da Instrução Normativa Nº 160, de 27 de abril de 2007, promulgada pelo Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis (IBAMA, 2007, p. 404) apresenta uma definição importante:

coleção biológica científica: coleção de material biológico devidamente tratado, conservado e documentado de acordo com normas e padrões que garantam a segurança, acessibilidade, qualidade, longevidade, integridade e interoperabilidade dos dados da coleção, pertencente à instituição científica com objetivo de subsidiar pesquisa científica ou tecnológica e a conservação *ex situ*.

A Instrução Normativa se refere, majoritariamente, às coleções físicas de espécimes vivos ou não-vivos, mas também torna obrigatório que essas coleções estejam devidamente documentadas e que a interoperabilidade dos dados das mesmas seja assegurada, de acordo com normas e padrões internacionais, que são, basicamente, os princípios indicados por Rodrigues *et al.* (2010) para a construção de repositórios de dados científicos.

Para tratar especificamente sobre a publicação, disponibilização, acesso e uso dos dados sobre a biodiversidade, o Instituto Chico Mendes de Conservação da Biodiversidade (ICMBio, 2015) instituiu a Política de Dados e Informações sobre Biodiversidade para o tratamento das coleções produzidas no âmbito do instituto. Esta instrução normativa também inclui importantes definições, as quais:

III - dado: sequência de símbolos quantificados ou quantificáveis referentes a um objeto ou evento, podendo consistir em textos, números, datas, imagens, arquivos vetoriais, entre outros;

IV - informação: afirmação realizada a partir da organização, análise ou interpretação de um conjunto dados;

VI - dados ou informações sensíveis: são aqueles para os quais a disponibilização pode comprometer a proteção de espécies ou ecossistemas (ICMBio, 2015, *online*).

A instrução normativa trata, ainda, sobre outras definições relativas à informação sobre biodiversidade, mas não indica um padrão para a representação da informação de espécies. A utilização de padrões de metadados consolidados como base para a descrição de espécies é essencial para que haja interoperabilidade entre os modelos

de sistemas de informação.

O padrão DwC⁵, por exemplo, que reúne metadados para descrição de registro de ocorrência de espécies em meio digital, é o padrão utilizado internacionalmente neste contexto, e seus campos de metadados foram traduzidos para o português pelo Sistema de Informação sobre a Biodiversidade Brasileira (SIBBR, 2018).

Foi constatado, durante exploração preliminar nesta pesquisa, que muitas bases de dados brasileiras ainda não utilizam o padrão DwC, e nem outro qualquer, para descrever seus dados. Também se verificou que alguns dos bancos de dados brasileiros que adotam o DwC o utilizam de maneira incorreta, ou incompleta, conforme verificado em estudo realizado por Soares, Hamanaka e Maculan (2018). Qualquer das duas situações impossibilita a interoperabilidade de dados com outras bases de dados nacionais e estrangeiras. Esse problema ocorre porque, segundo Silva, Farinelli e Almeida (2014, p. 4077), “a interoperabilidade entre sistemas é possível apenas se as linguagens subjacentes aos modelos possuem conceitualizações que se sobrepõem em algum nível”. Ou seja, é necessário que os sistemas de informação adotem um mesmo modelo para representação da informação, de forma que cada campo ou conceito deste modelo seja único e elimine a ambiguidade semântica. Esse problema de representação dos registros dificulta a recuperação da informação sobre espécies, afetando tanto os pesquisadores que utilizam essas bases de dados quanto usuários não especialistas desses sistemas, resultando na dificuldade em implementar a Política Nacional de Biodiversidade (PNB).

Discorrendo sobre a integração de dados, Santos *et al.* (2011, não paginado) afirmam que “Um dos problemas mais frequentes é como integrar dados das múltiplas fontes disponíveis. A integração de dados requer alguns cuidados, tais como, integração semântica de dados, interoperabilidade etc.” A dificuldade na integração de dados dificulta a criação de uma fonte-referência, onde seria possível centralizar todos os dados sobre a biodiversidade.

Ressalta-se que, sistemas que manipulam dados de uma forma geral, mas especificamente sobre biodiversidade requerem interoperabilidade (capacidade de um sistema se comunicar com outro

⁵ Detalhes específicos sobre o padrão DwC estão descritos na Seção 2.6.

sistema) das diversas fontes utilizadas. Desta forma, heterogeneidade semântica, onde um único fato pode ter mais de uma descrição, dependendo de quem vê/descreve, deve ser considerada. Por exemplo, tipo de vegetação também pode ser entendido como bioma (SANTOS *et al.*, 2011, não paginado).

Segundo os autores, profissionais de diversas áreas estão trabalhando em conjunto para desenvolver ferramentas e modelos que resolvam os problemas de interoperabilidade, tais como profissionais das ciências biológicas e ciência da computação. Os profissionais da Biblioteconomia e CI também podem contribuir com tal empreendimento, com as expertises de modelagem conceitual de domínios e aplicação das teorias da classificação consolidadas na Biblioteconomia e CI.

A gestão de dados sobre a biodiversidade é um recurso importante para setores da economia, como a agricultura. Diversas pesquisas têm sido realizadas no sentido de aproveitar os recursos ecológicos como insumos sustentáveis para os sistemas de cultivo, de forma a aumentar o volume da produção das lavouras com impacto ambiental reduzido. Dá-se forma, então, a um campo de estudo: a agrobiodiversidade, foco desta investigação.

2.2 Agrobiodiversidade

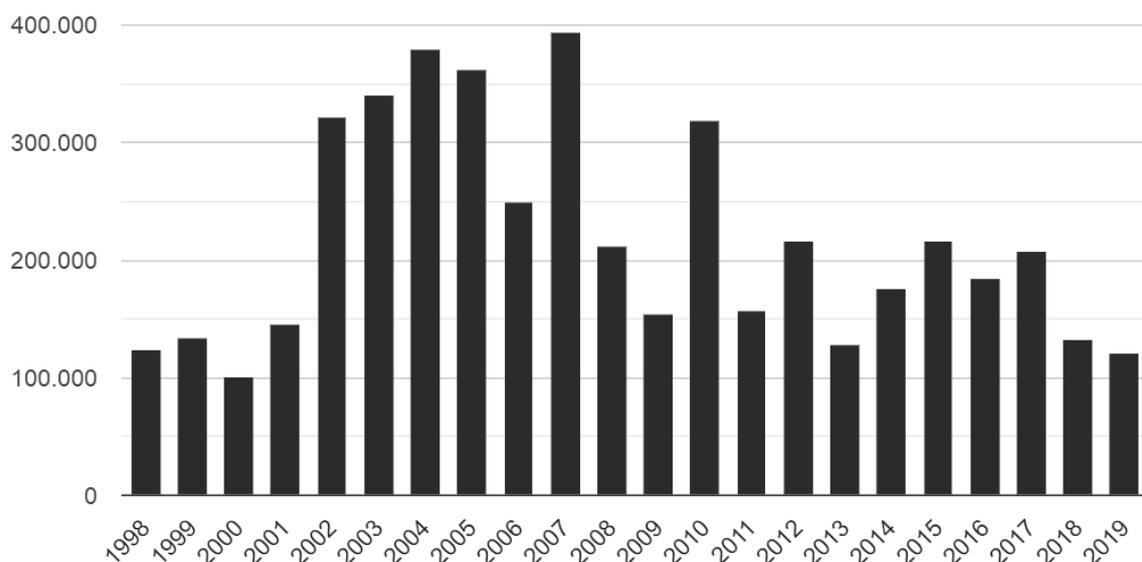
A agrobiodiversidade “inclui todos os componentes da biodiversidade que têm relevância para a agricultura e alimentação”, que são “as variedades e a variabilidade de animais, plantas e de microrganismos, nos níveis genético, de espécies e de ecossistemas os quais são necessários para sustentar as funções chaves dos agroecossistemas, suas estruturas e processos.” (CBD, 2000, p. 85, tradução do proponente).

Tendo conhecimento sobre a aplicação de recursos da biodiversidade para melhorias nas práticas agrícolas, as fazendas devem atuar utilizando esses recursos como estratégia para “resiliência [...] a eventos climáticos e patologias de plantas, e [...] opções para estratégias adaptativas às mudanças ambientais e econômicas” (ARNAUD *et al.*, 2016, p. 7, tradução do proponente). Esse tipo de uso da biodiversidade cria uma relação benéfica entre fazendeiros e os ecossistemas, e ambas as partes são beneficiadas: “isso apoia a restauração dos serviços ecossistêmicos e fornece um reservatório genético de novas características e espécies para agricultura” (ARNAUD *et al.*, 2016, p. 7, tradução do proponente).

Tendo conhecimento do potencial da biodiversidade e dos ecossistemas, os fazendeiros devem trabalhar de maneira a preservar os ecossistemas, e não o contrário, como tem sido observado em diversos casos no Brasil. O Gráfico 1, que contém dados do Instituto Nacional de Pesquisas Espaciais (INPE), mostra o histórico de focos de queimada no Brasil (em todo o território nacional) no período de 1998 a 2019.

De acordo com estudos sobre queimadas no Brasil, realizados por Alencar *et al.* (2004), Torres *et al.* (2016) e Silvério *et al.* (2019), o principal motivo para os focos de queimadas no Brasil, apresentados em números no Gráfico 1, é a limpeza de terras para conversão em pastagens para produção pecuária ou campos de produção de grãos.

Gráfico 1 – Número de focos de incêndio no Brasil detectados por satélite



Fonte: INPE (2019)

As queimadas apresentam impactos negativos para a saúde dos ecossistemas e para a saúde das populações humanas, tanto em áreas rurais como em áreas urbanas (STEINBERGER, 2002). As nuvens de fumaça que escureceram o céu de diversas cidades brasileiras no ano de 2019 resultaram do transporte de fuligem das queimadas em região amazônica, principalmente, por massas de ar para os principais centros urbanos do país, sendo que a concentração de fuligem no ar representa risco em potencial para a saúde humana (ALVES *et al.*, 2015; SILVÉRIO *et al.*, 2019; STEINBERGER, 2002). Segundo a meteorologista Pegorim (2019) e uma nota técnica

emitida pelo Instituto Nacional de Meteorologia (INMET, 2019), o escurecimento do céu foi uma combinação de fumaça advinda de focos de queimadas em região amazônica da Bolívia, em Rondônia, no Acre, no Paraguai e no norte da Argentina, juntamente com muita umidade trazida por uma frente fria no litoral paulista, atingindo também a região sul de Mato Grosso, Mato Grosso do Sul e Paraná.

Além do impacto no clima global, as queimadas nos ecossistemas também acarretam completa perda de biodiversidade nas regiões de floresta e mudança do microclima regional, podendo alterar a sazonalidade de chuvas na região e temperatura, por exemplo (GUIMARÃES *et al.*, 2014). Outros impactos nos ecossistemas apontados por Guimarães *et al.* (2014, p. 45-46) são:

- a) “aumento na velocidade do vento”, provocada pela retirada de vegetação e que contribui para a destruição da vegetação remanescente que escapou à queimada;
- b) “liberação de carbono na atmosfera – o CO₂”;
- c) “redução da qualidade da água”;
- d) “mortalidade de animais”;
- e) interferência na reconstituição vegetal, uma vez que a queimada favorece o brotamento de espécies resistentes ao fogo e pode provocar a extinção ou desaparecimento progressivo de espécies não resistentes.

Já a limpeza de campos de cultivo por fogo, provoca a morte da microbiota do solo que contém, por exemplo, bactérias úteis à fixação de nutrientes para as plantas (CAPECHE, 2012, p. 17; GUIMARÃES *et al.*, 2014).

Sabe-se que a queimada altera, direta ou indiretamente, as características físicas, químicas, morfológicas e biológicas dos solos, como o pH, teor de nutrientes e carbono, biodiversidade da micro, meso e macrofauna, temperatura, porosidade e densidade. Sem falar no aumento do efeito estufa, na redução da qualidade do ar e da água, e da saúde. (CAPECHE, 2012, p. 17).

No caso das pastagens, a queimada favorece o brotamento de certas espécies de gramíneas utilizadas como pasto (GUIMARÃES *et al.*, 2014). Entretanto, os impactos para o solo retratados por Capeche (2012) tendem a prejudicar a produção das

pastagens a longo prazo.

Apresentado este panorama, destaca-se de real importância o estudo e o incentivo à aplicação da biodiversidade, denominada neste contexto como agrobiodiversidade, de maneira sustentável nas práticas agrícolas para que os maus usos dos ecossistemas sejam evitados. É preciso dar acesso a esse conhecimento aos principais produtores do meio em território nacional, mas também aos pequenos agricultores dedicados à agricultura familiar, que também podem desenvolver práticas prejudiciais e pouco sustentáveis (ALENCAR *et al.*, 2004).

Seguindo este direcionamento, em 2016 foi publicado pelo Task Group on Data Fitness for Use in Agrobiodiversity (Task Group) o *Final Report of the Task Group on GBIF Data Fitness for Use in Agrobiodiversity*, que propõe diretrizes para melhorar as práticas de produção e compartilhamento de dados sobre a agrobiodiversidade no mundo, com foco em lavouras e campos cultiváveis (ARNAUD *et al.*, 2016). O foco do documento é sugerir práticas para a gestão de dados sobre a agrobiodiversidade que fomentem soluções sustentáveis de uso de recursos biológicos em agroecossistemas. Este documento é a principal referência sobre gestão de dados da agrobiodiversidade.

O Task Group on Data Fitness for Use in Agrobiodiversity foi estabelecido pela secretaria da GBIF e pela Bioversity International para ajudar a ajustar o tratamento dos dados relacionados à agrobiodiversidade à variedade de usos importantes requeridos e solicitados pela comunidade de pesquisa e política. O grupo de trabalho tem analisado as principais ações para a criação de interoperabilidade de dados sobre conservação *ex situ*, *in situ* e *on farm* da agrobiodiversidade, com foco nas plantas (ARNAUD *et al.*, 2016, p. 6, tradução do proponente).

Dentro da presente pesquisa, a agrobiodiversidade representa um amplo escopo de estudos. Por sua vez, o Task Group tem seu escopo de pesquisa atualmente limitado ao cultivo de plantas, mas reconhece a importância de receber especialistas de outros domínios de pesquisa da agrobiodiversidade, como a pecuária, para ampliar o escopo das recomendações do relatório (ARNAUD *et al.*, 2016).

Arnaud *et al.* (2016) definem 53 recomendações de melhores práticas para tratamento e gestão de dados da agrobiodiversidade, que “cobrem todo o fluxo de dados, desde a publicação até uso de dados com foco na agrobiodiversidade” (ARNAUD *et al.*, 2016, p. 6, tradução do proponente). Dentre essas recomendações, está presente a

indicação de integração do Multi Crop Passport Data standard (MCPD), um antigo padrão de metadados muito utilizado por bancos de genes agrícolas, como uma extensão⁶ do DwC⁷ (ARNAUD *et al.*, 2016).

Outro tema das recomendações são as interações ecológicas. Na Seção 6.5 do relatório, Arnaud *et al.* (2016) abordam as interações ecológicas no contexto da agrobiodiversidade, afirmando que esse estudo é importante para

prever, para uma determinada área, os possíveis fluxos genéticos entre as culturas e as espécies selvagens parentes das espécies cultivadas, tirar conclusões sobre as oportunidades para a evolução da diversidade nas fazendas e adaptação das culturas e identificar trade-offs e riscos nas intervenções para ações de conservação, manejo da terra ou restauração. São necessários dados sobre a presença de gado, pragas, doenças, insetos úteis como polinizadores ou predadores de pragas. Além disso, o risco de uma espécie se tornar invasora em um determinado ambiente é uma informação importante para a tomada de decisões sobre restauração da paisagem, gerenciamento de recursos naturais e conservação de espécies ameaçadas (ARNAUD *et al.*, 2016, p. 23, tradução do proponente).

Os autores ressaltam, de tal maneira, a importância de se gerenciar os dados sobre as interações ecológicas para garantir a manutenção de diversos serviços ecossistêmicos, como o controle de pragas e doenças e polinização. Para que esses serviços sejam mantidos em equilíbrio, é necessário que os atores que lidam com o cultivo da terra adotem práticas que não provoquem distúrbios nas populações dos organismos que desenvolvem os serviços ecossistêmicos, logo, é preciso que os dados sobre as interações estejam disponíveis.

Além de Arnaud *et al.* (2016), a CBD (2011) e a Food and Agriculture Organization of the United Nations (FAO, 2011) também reconhecem a importância das interações ecológicas para a produção de alimentos. Eles recomendam a prática integrada entre agricultura e conservação dos ecossistemas, utilizando-se para tal a técnica de manejo conhecida como mosaicos sustentáveis, em que plantações são cultivadas em meio a corredores ecológicos.

A agricultura em mosaicos paisagísticos complexos e diversificados é considerada uma maneira de integrar a necessidade de conservação da biodiversidade e produção de alimentos, sendo que práticas

⁶ Para saber mais sobre extensões de metadados, leia a Seção 2.6.1.

⁷ Para saber mais sobre o padrão de metadados DwC, leia a Seção 2.6.

agrícolas que se baseiam no conhecimento da biodiversidade e na interação entre espécies podem aumentar consideravelmente a produtividade” (FAO, 2011 citado por CBD, 2011, tradução do proponente).

Arnaud *et al.* (2016) apresentam quatro recomendações sobre a gestão de dados para favorecer as práticas sustentáveis da agrobiodiversidade, que são:

6.5.1 Atributos adicionais são necessários sobre espécies relacionadas a culturas ou entre espécies no nível do nome do táxon, como pragas, predadores, polinizadores etc.

6.5.2 Atributo adicional "patógeno" com o nome científico dos patógenos e nomes vernáculos das doenças devem ser disponibilizados.

6.5.3 O portal GBIF deve permitir a seleção e o download de ocorrências de culturas juntamente com ocorrências de pragas, doenças, polinizadores, gado etc.

6.5.4 As informações sobre o risco de uma espécie se tornar invasora devem ser disponibilizadas por meio de um link entre o portal GBIF e outros bancos de dados com essas informações, como o banco de dados do CABI sobre espécies invasoras (ARNAUD *et al.*, 2016, p. 23).

As recomendações não são muito objetivas, pois apontam o que deve ser feito, mas não indicam ferramentas ou padrões de metadados para todas as situações. A recomendação 6.5.1 aponta a necessidade de se adotar atributos adicionais que descrevam as relações entre espécies com as culturas (ou seja, com espécies cultiváveis), e entre espécies que não fazem parte da lavoura em si, ou que não são cultivadas, entretanto, não há detalhamento sobre como isso deve ser feito ou sobre quais ferramentas devem ser utilizadas.

O mesmo acontece com as recomendações 6.5.2 a 6.5.4, o que leva à conclusão de que a ausência do detalhamento das recomendações se deve à inexistência de ferramentas para tal intento. Desta forma, o enquadramento desta pesquisa busca atender exatamente à recomendação 6.5.1, uma vez que se busca apresentar princípios para a criação de uma extensão de metadados DwC para as interações ecológicas no contexto da agrobiodiversidade.

2.2.1 Interações ecológicas

As interações ecológicas se desenvolvem no ambiente natural como uma complexa

teia de ligações, sendo que a maioria dos seres vivos desenvolve muitas delas ao longo da vida, muitas vezes ao mesmo tempo (BEGON; TOWNSEND; HARPER, 2007).

A organização da análise das interações ecológicas implica em classificação, do mesmo modo que ocorre com diversos outros campos científicos. De acordo com Bronstein (2015), essa classificação influencia a maneira de se perceber um fenômeno, o que pode levar “certo fenômeno a ser desproporcionalmente estudado, ou ser injustificadamente ignorado; [a classificação] chama a atenção para as semelhanças e diferenças, mas também pode levá-lo ao esquecimento” (BRONSTEIN, 2015, p. 4, tradução do proponente). De acordo com a autora, a classificação tem esse efeito sobre as interações ecológicas.

Indivíduos de diferentes espécies podem interagir entre si, podendo haver afetação ou não a ambas as partes (CAIN; BOWMAN; HACKER, 2018; BEGON; TOWNSEND; HARPER, 2007; ODUM; BARRET, 2006). O conceito apresenta-se na literatura representado por diversos termos, que são considerados equivalentes: interações entre espécies (ARNAUD *et al.*, 2016; CBD, 2011; ODUM; BARRET, 2006; PINTO-COELHO, 2007; ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT, 2001); interações entre organismos (CAIN; BOWMAN; HACKER, 2018); interações (BEGON; TOWNSEND; HARPER, 2007); interações biológicas (BIOLOGICAL..., 2011); interações ecológicas (STEIN, 2018). Considera-se, na presente pesquisa, que o termo ‘interações ecológicas’ é o mais adequado para representar o conceito. Apesar de ser muito utilizado, o termo ‘interações entre espécies’ pode conter ambiguidade semântica e ser entendido apenas como a relação entre duas espécies diferentes. Todavia, como será explorado adiante nesta seção, as interações podem também ocorrer entre indivíduos de uma mesma espécie, em uma população.

Para sobreviver em um ecossistema, os organismos devem encontrar as condições favoráveis e recursos necessários à sua sobrevivência, sendo que a dependência desses fatores surgiu a partir de um longo processo adaptativo da espécie (STEIN, 2018). Entende-se por condições favoráveis os fatores químicos e físicos presentes nos diferentes ecossistemas da Terra, que não podem ser consumidos pelos organismos e podem sofrer variações de acordo com pressões diversas, como por

exemplo o efeito estufa (STEIN, 2018). São exemplos de condições: “a) temperatura; b) umidade do ar; c) pH de lagoas, oceanos, lagos, rios etc.; d) pressão atmosférica” (STEIN, 2018, p. 85). Já os recursos são elementos bióticos ou abióticos suscetíveis de serem consumidos ou utilizados de outra maneira pelos organismos para garantir “sua manutenção, seu crescimento e reprodução” (STEIN, 2018, p. 85). São recursos: “a) alimento [...]; b) luminosidade [...]; c) abrigos [...]; d) território [...]; e) água; f) parceiros sexuais” (STEIN, 2018, p. 86).

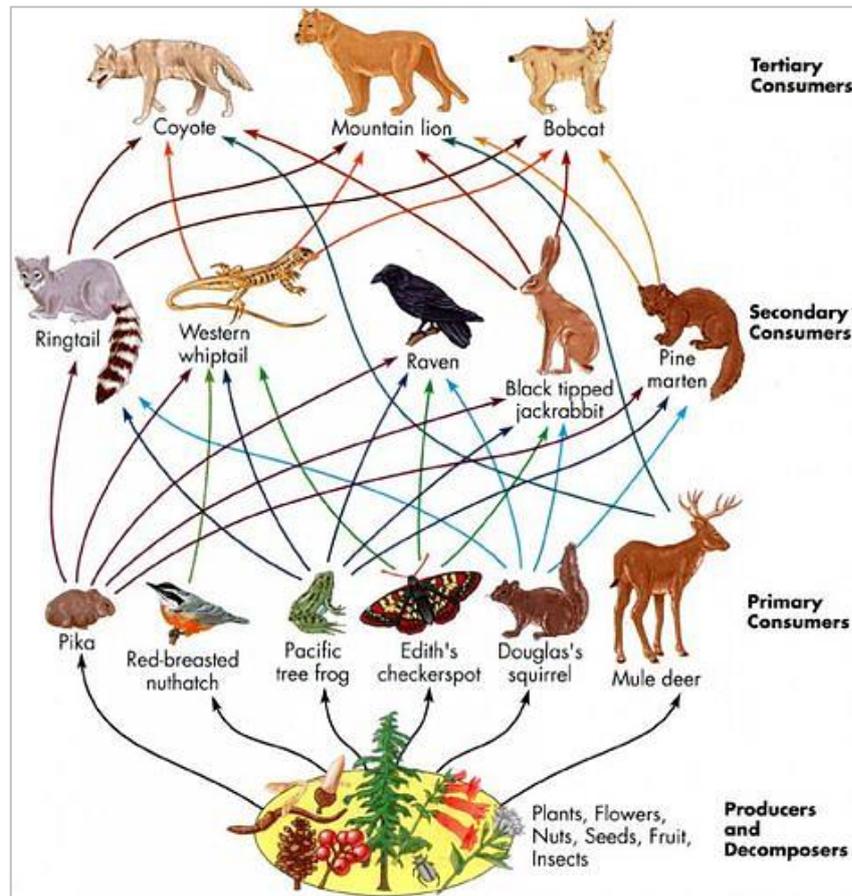
As interações ecológicas podem ter diferentes funções, como alimentação, abrigo e reprodução. Tratando-se da alimentação, há dois tipos de organismos: autotróficos e heterotróficos (ODUM; BARRET, 2006; STEIN, 2018). Os organismos autotróficos produzem seu próprio alimento a partir de matéria inorgânica presente no meio onde se desenvolvem, seja por meio do processamento da luz (fotossíntese⁸), ou por reações químicas em ambientes sem presença de luz (quimiossíntese⁹) (MOYES; SCHULTE, 2010; STEIN, 2018). Os organismos heterotróficos, por sua vez, são incapazes de produzir seu próprio alimento, nutrindo-se, pois, às custas de outros organismos (STEIN, 2018).

Os organismos autotróficos são denominados produtores primários da cadeia alimentar. Já os organismos heterotróficos são sempre consumidores, sendo considerados consumidores primários aqueles que se alimentam de organismos produtores, consumidores secundários aqueles que se alimentam de consumidores primários, e assim por diante (STEIN, 2018). A Figura 2 ilustra os ciclos, de maneira muito simplificada, de uma teia alimentar ou teia trófica, como também é referenciada na literatura.

⁸ Organismo denominado fotoautotrófico (MOYES; SCHULTE, 2010).

⁹ Organismo denominado quimioautotrófico (MOYES; SCHULTE, 2010).

Figura 2 – Teia alimentar



Fonte: Descubra... (2019)

O papel de cada organismo na teia alimentar depende, basicamente, de seu hábito alimentar. O Quadro 1 descreve e exemplifica os hábitos alimentares dos organismos na teia alimentar.

Quadro 1 – Os tipos de hábitos alimentares

HÁBITOS ALIMENTARES	FONTE DE ENERGIA	EXEMPLOS
Fotossintetizantes (produtores primários)	Energia solar	Plantas, bactérias e protistas fotossintetizantes
Herbívoros (consumidores primários)	Tecidos de produtores primários	Cupins, gafanhotos, anchovas, veados e gansos
Carnívoros primários (consumidores secundários)	Herbívoros	Aranhas, mariquitas, lobos e copépodos
Carnívoros secundários (consumidores terciários)	Carnívoros primários	Atuns, falcões e orcas
Onívoros	Vários níveis tróficos	Homens, gambas, caranguejos e tordos
Detritívoros (decompositores)	Cadáveres e excretas de outros seres vivos	Fungos, muitas bactérias, urubus e minhocas

Fonte: Saudava *et al.* (2009)

Os hábitos alimentares de cada organismo descritos no Quadro 1 são responsáveis por manter o ciclo de transferência de matéria e energia no ecossistema, que se inicia com os produtores primários e termina com os decompositores, para ter novo início (ODUM; BARRET, 2006; SADAVA *et al.*, 2009; STEIN, 2018). Esses processos dão origem a diversas interações entre organismos, a fim de que cada um consiga os recursos necessários para sobreviver. Os hábitos alimentares dos organismos podem mudar ao longo da vida e/ou conforme o ambiente, o que pode levá-lo a ocupar diferentes posições na teia alimentar.

Os organismos que desempenham o papel de consumidores na teia alimentar da Figura 2 representam um tipo de interação ecológica, que é a predação. Existem oito categorias principais de interações ecológicas, e a afetação para os organismos envolvidos pode ser neutra, positiva ou negativa (ODUM; BARRET, 2006), conforme descrito no Quadro 2.

Quadro 2 – Grade de interações ecológicas

Tipo de interação	Espécie		Natureza da interação
	A	B	
Competição	-	-	Inibição mútua
Neutralismo	0	0	Sem efeitos
Mutualismo	+	+	É obrigatória
Protocooperação	+	+	Facultativa
Predação	+	-	B é destruído por A
Parasitismo	+	-	B é explorado por A
Comensalismo	+	0	Hospedeiro não é afetado
Amensalismo	-	0	A população A é inibida; a população B não é afetada

Fonte: Pinto-Coelho (2007)

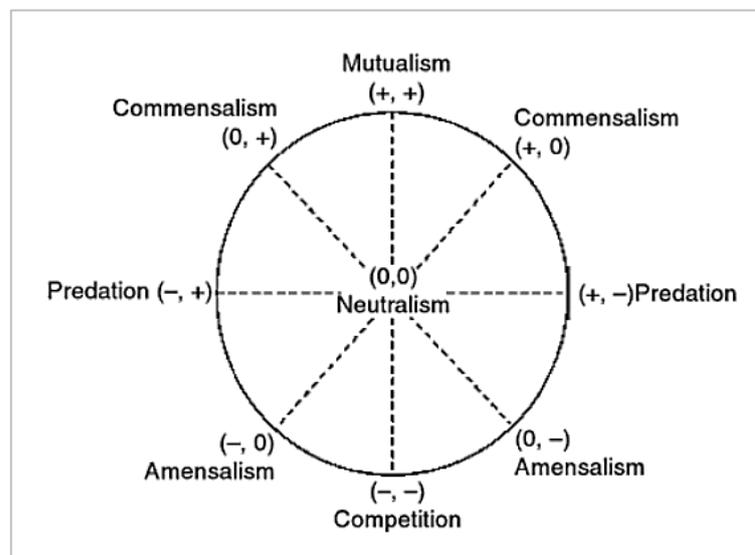
Legenda: 0: ausência de benefício ou prejuízo para a espécie em questão; +: associação traz benefício para a espécie considerada; -: associação provoca prejuízos para a espécie considerada.

As interações relacionadas no Quadro 2 são classificadas como interespecíficas ou heterotípicas, pois ocorrem entre indivíduos de espécies diferentes (ODUM; BARRET, 2006; STEIN, 2018). Já as interações intraespecíficas ou homotípicas, não

compreendidas pelo Quadro 2, são: sociedade (Seção 2.2.1.7), colônia (Seção 2.2.1.8), competição (Seção 2.2.1.2) e canibalismo (Seção 2.2.1.9). E a relação pode, também, ser classificada quanto ao tipo de afetação dos indivíduos envolvidos, sendo harmônica ou positiva quando pelo menos um dos indivíduos é beneficiado sem causar prejuízo ao outro (+/+ ou +/0) e desarmônica ou negativa quando pelo menos um dos indivíduos envolvidos é prejudicado (-/-, +/- ou -/0) (CAIN; BOWMAN; HACKER, 2018; STEIN, 2018).

As interações interespecíficas presentes no Quadro 2 são comumente apresentadas em forma radial, como a Figura 3.

Figura 3 – Bússola de interação



Fonte: Holland e Bronstein (2008)

Nota: Mover-se do centro em direção à periferia da bússola aumenta a força ou magnitude do resultado da interação, mas não altera o sinal do efeito da interação para nenhuma das espécies que interagem. Por outro lado, mover-se pela periferia da circunferência altera o sinal e o tipo de interação interespecífica (HOLLAND; BRONSTEIN, 2008, p. 2485).

É possível observar que a Figura 3 exibe apenas seis das oito interações apontadas no Quadro 2. Isso acontece porque há certa divergência conceitual no campo da Ecologia quanto à definição de algumas interações, como parasitismo e protocooperação, que são incluídas por alguns autores nas categoriais predatismo e mutualismo, respectivamente (BRONSTEIN, 2015; HOLLAND; BRONSTEIN, 2008; MORIN, 2011). Além disso, o Quadro 2 e a Figura 3 não apresentam as interações intraespecíficas, ou seja, que acontecem entre indivíduos de uma mesma espécie,

que são a sociedade, colônia, canibalismo e competição intraespecífica.

A seguir, define-se as interações ecológicas que são consideradas no desenvolvimento desta pesquisa: a) predatismo; b) parasitismo; c) competição (intra e interespecífica); d) neutralismo; e) mutualismo; f) protocooperação; g) comensalismo; h) amensalismo; i) sociedade; j) colônia; k) canibalismo.

2.2.1.1 Predatismo e parasitismo

Conforme mencionado anteriormente, a teia trófica da Figura 2 representa a interação do tipo predação, que é definida no Quadro 2 como uma relação +/-, em que o predador (espécie A) destrói a presa (espécie B). No exemplo da cadeia alimentar da Figura 2, essa interação pode ser observada desde a base, uma vez que a herbivoria (o consumidor primário que se alimenta do produtor, que é uma planta) também é considerada um tipo de interação de predação (ODUM; BARRET, 2006). A predação é considerada uma relação desarmônica interespecífica. A Figura 4 representa um caso de predatismo que acontece em lavouras.

Figura 4 – Exemplo de predatismo



Fonte: Sanches (2016)

Na Figura 4, a joaninha da espécie *Cryptolaemus montrouzieri* (Mulsant, 1853) se alimenta da cochonilha-rosada, espécie que ataca mais de 200 espécies de plantas cultiváveis (VALE, 2016). Esta espécie de joaninha, que é exótica, foi introduzida no Brasil pela Embrapa e é utilizada como controle biológico em diversos tipos de plantações (SANCHES; CARVALHO, 2010).

O parasitismo segue o mesmo princípio da predação: um indivíduo (denominado

hospedeiro) sofre danos e nenhum benefício, enquanto o indivíduo parasitário se beneficia do primeiro (relação +/-) (CAIN; BOWMAN; HACKER, 2018; ODUM; BARRET, 2006). O organismo parasito desenvolve comumente uma relação de dependência para com o hospedeiro, vivendo dentro ou sobre ele (vide Figura 5) (CAIN; BOWMAN; HACKER, 2018, p. 273; ODUM; BARRET, 2006).

Figura 5 – Exemplo de parasitismo



Fonte: Secretaria de Educação (2013)

Na Figura 5 observa-se as raízes da erva-de-passarinho em tronco de uma laranjeira, situação muito comum de ser observada até mesmo em quintais de residências. A erva-de-passarinho se instala no tronco da árvore e ocupa boa parte da copa, o que reduz a capacidade de fotossíntese da laranjeira; as raízes da erva parasita adentram o tronco da laranjeira e sugam-lhe nutrientes vitais, o que prejudica seu desenvolvimento e pode leva-la à morte no auge da infestação (GONÇALVES *et al.*, 2012).

2.2.1.2 Competição

Retomando os outros tipos de interações do Quadro 2, a competição (- -) é um tipo de interação em que ambas as partes são prejudicadas, portanto é considerada desarmônica. Dois organismos competem por um mesmo recurso disponível no ambiente, afetando “seu crescimento, sobrevivência ou reprodução” (CAIN; BOWMAN; HACKER, 2018, p. 273). A competição pode ser de dois tipos: por interferência direta ou por uso de recurso (CAIN; BOWMAN; HACKER, 2018; ODUM; BARRET, 2006; PINTO-COELHO, 2007; STEIN, 2018).

A competição por interferência direta ou explorativa acontece quando um organismo

inibe diretamente o desenvolvimento de outro, mesmo que não haja escassez de recursos no ambiente (PINTO-COELHO, 2007). Pode-se citar, como exemplo, o eucalipto, cujas folhas contêm substâncias fitotóxicas que ao serem decompostas liberam toxinas que podem matar ou causar danos a outras plantas, como se pode observar na Figura 6 (SAITO, 2015).

Figura 6 – Exemplo de competição por interferência direta



Fonte: Estudo... (2019)

A Figura 6 representa um exemplo de competição interespecífica por interferência direta, uma vez que o eucalipto elimina outras espécies de planta e não há competição entre os indivíduos da mesma espécie.

Já a competição por uso de recurso “ocorre quando um certo número de organismos de uma dada espécie ou de espécies diferentes usa um recurso que esteja sendo limitante” (PINTO-COELHO, 2007), afetando, portanto, o outro indiretamente, “de forma adversa, na disputa por recursos escassos” (STEIN, 2018, p. 88). Pode-se citar, como exemplo, a competição de plantas cultiváveis em uma lavoura, como o milho, e ervas daninhas, que pode resultar em redução do “crescimento, a massa vegetal e o rendimento de grãos das culturas” (RIZZARDI *et al.*, 2001, p. 708), conforme ilustrado na Figura 7.

Figura 7 – Exemplo de competição por recurso



Fonte: Plantas... (2019)

O milho, planta com folhas verdes mais destacadas, e diversas espécies de ervas daninhas, que são as plantas de menor porte na Figura 7, competem por luz, nutrientes e espaço, o que representa uma competição interespecífica.

Tratando-se de competição entre indivíduos de uma mesma espécie, essa pode ocorrer por motivos de ocupação de espaço, alimentação e reprodução, por exemplo. A Figura 8 ilustra o exemplo de competição entre leões¹⁰, que pode acontecer pela dominação de um território, reprodução, alimentação ou outros fatores. No contexto da agricultura, pode ocorrer competição intraespecífica em plantações, caso os limites de espaço não sejam adequados entre uma planta e outra. No caso do milho, por exemplo, o rendimento dos grãos é alterado “de acordo com o grau de competição intraespecífica proporcionado pelas diferentes densidades de planta” (CRUZ; PEREIRA FILHO; ALBUQUERQUE FILHO, 2011, *online*).

¹⁰ *Panthera leo* (Linnaeus, 1758).

Figura 8 – Exemplo de competição intraespecífica



Fonte: Wessels (2019)

Portanto, a competição é um tipo de interação em que ambas as espécies ou indivíduos sofrem prejuízos (desarmônica), podendo ser intra ou interespecífica. O mais adaptado ao ambiente e às condições conseguirá vencer a competição e se sobressair, levando a outra espécie ou organismo a extinção, o que é denominado exclusão competitiva (CAIN; BOWMAN; HACKER, 2018).

2.2.1.3 Neutralismo

O neutralismo é um tipo de interação que não pode ser classificada como harmônica ou desarmônica, visto que não há benefícios nem prejuízos a qualquer das partes envolvidas (0/0, vide Quadro 2). Caracteriza-se pela coexistência de organismos em um mesmo espaço, sem interferência de um sobre o outro, conforme o exemplo:

Duas espécies próximas de aves aquáticas, o corvo-marinho-de-crista (*Phalacrocorax aristotelis*) e o corvo-marinho-de-faces-brancas (*P. carbo*), são encontradas juntas na Inglaterra durante a estação de procriação, mas se alimentam de tipos de peixes diferentes. Dessa forma, elas não competem diretamente por fontes de alimento (isto é, o nicho das duas espécies é diferente). (ODUM; BARRET, 2006, p. 296).

Outro exemplo, segundo a Universidade do Vale do Itajaí (2017), é a relação entre zebras e girafas, conforme ilustrado na Figura 9. Ambas as espécies convivem no mesmo ambiente, “mas não dependem uma da outra e não afetam uma à outra em questão de habitat, alimentos, abrigo” (UNIVERSIDADE DO VALE DO ITAJAÍ, 2017, p. 3).

Figura 9 – Exemplo de neutralismo



Fonte: Pixabay (2016)

Segundo Gelambi (2019), não é comum que as espécies interajam sem que haja efeito para pelo menos uma das partes, por isso é raro observar este tipo de interação na natureza. Segundo a autora, é mais comum ocorrerem interações em que uma das espécies é afetada ou se beneficia e a outra permanece na neutralidade: “Existem mais dois modelos, comensalismo e amensalismo, que propõem neutralidade para uma das espécies participantes e a outra terá um efeito positivo ou negativo – respectivamente” (GELAMBI, 2019, *online*, tradução do proponente).

2.2.1.4 Mutualismo e protocooperação

O mutualismo é uma interação ecológica harmônica interespecífica, na qual os indivíduos de ambas as espécies envolvidas se beneficiam da relação (+/+), vide Quadro 2). É uma das interações ecológicas mais presentes nos ecossistemas, já que a “a maioria das espécies participa de interações de mutualismo” (SADAVA *et al.*, 2009, p. 826). As relações mutualísticas podem desempenhar diferentes funções:

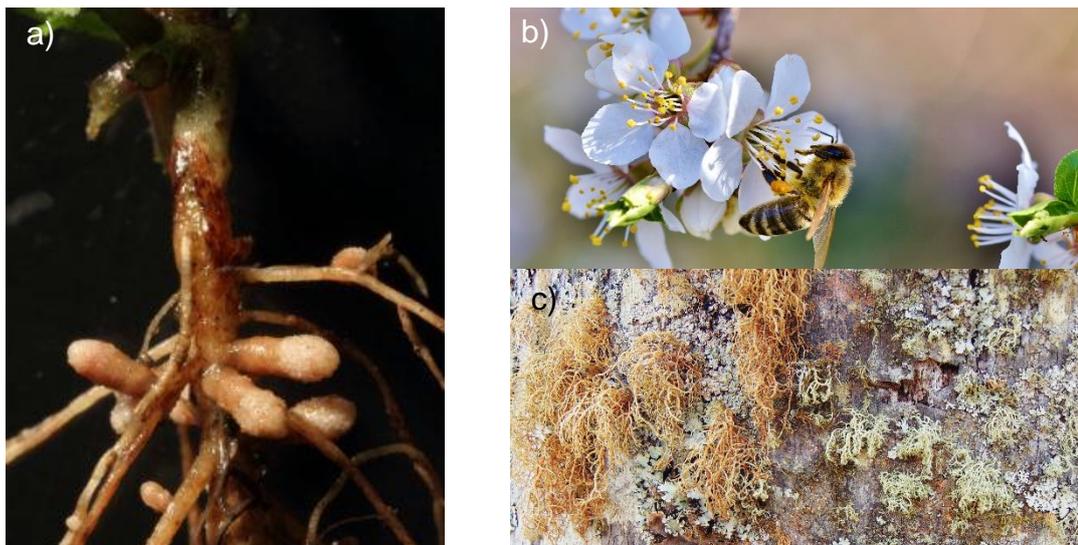
- a) nutricional: fornecimento de alimento ou síntese de nutrientes (BOUCHER; JAMES; KEELER, 1982; BRONSTEIN, 2015; CAIN; BOWMAN; HACKER, 2018;
- b) proteção: anula ou ameniza o impacto de fatores ambientais, como estresse

causado pela escassez de água e calor, e de predadores e/ou parasitos sobre o parceiro, além de fornecer substrato para o desenvolvimento do mesmo (BOUCHER; JAMES; KEELER, 1982; BRONSTEIN, 2015; CAIN; BOWMAN; HACKER, 2018);

- c) transporte: dispersão de gametas (polinização) e de sementes (BOUCHER; JAMES; KEELER, 1982; BRONSTEIN, 2015; CAIN; BOWMAN; HACKER, 2018).

O termo protocooperação (também conhecido como mutualismo facultativo) teve origem, de acordo com Bronstein (2015), como uma categoria de interação ecológica devido à relutância de muitos pesquisadores em denominar como mutualística uma interação para a qual não houvesse “evidências para comprovar que as espécies compartilharam longa história de coevolução, ou teve pelo menos algum nível de dependência obrigatória” (BRONSTEIN, 2015, p. 7, tradução do proponente). Desta forma, seria considerado mutualismo apenas as relações as quais os indivíduos não podem viver de maneira independente, ou seja, estão anatomicamente conectados, o que é chamado também de simbiose (BRONSTEIN, 2015). A Figura 10 apresenta exemplos dos três tipos de mutualismo.

Figura 10 – Exemplos de interações mutualísticas



Fonte: a) Young (2015); b) Pixabay (2019); c) banco de imagens do proponente (2019)

Legenda: a) nódulos de bactéria fixadoras de nitrogênio na raiz de uma leguminosa; b) abelha polinizando uma flor; c) líquen (ou fungo liquenizado), que é uma associação entre fungos e algas.

Os exemplos A e C da Figura 10 representam associações obrigatórias, ou seja, um organismo está fisicamente interligado ao outro e não pode se desassociar; já o exemplo B mostra que os organismos, apesar de dependerem um do outro para alimentação e perpetuação da espécie, não estão fisicamente ligados, sendo que a abelha é livre para se deslocar para outras flores. Quanto às funções, pode-se dizer que o exemplo A representa uma relação por nutrição, na qual as bactérias se beneficiam da seiva elaborada produzida pela planta, sendo que a planta se beneficia da fixação de nutrientes absorvidos do solo pelas bactérias; o exemplo B representa uma relação de transporte/nutrição, na qual a abelha transporta o pólen de uma flor a outra, possibilitando a troca de gametas e, conseqüentemente, a reprodução, ao passo que se alimenta do néctar produzido pela flor; o exemplo C representa uma relação de nutrição/proteção, na qual a espécie de fungo fornece abrigo e proteção para a espécie de alga, que realiza a fotossíntese e encarrega-se de alimentar o fungo (SPIELMANN, 2006).

2.2.1.5 Comensalismo

É um tipo de interação interespecífica harmônica onde apenas um dos organismos envolvidos se beneficia (relação + 0 no Quadro 2), mas sem causar prejuízo para o outro (PINTO-COELHO, 2007). “As plantas epifíticas de florestas tropicais (bromélias e orquídeas) e suas árvores hospedeiras constituem exemplos desta interação ecológica” (PINTO-COELHO, 2007, p. 37).

Figura 11 – Exemplo de comensalismo



Fonte: Ferreira (2015)

A Figura 11 demonstra uma orquídea sobre o tronco de uma árvore, relação denominada epifitismo e que é um tipo de comensalismo. A orquídea consegue seus nutrientes a partir da casca da árvore, mas não causa nenhum prejuízo ou benefício para a árvore.

2.2.1.6 Amensalismo

É uma interação interespecífica desarmônica em que apenas um organismo é prejudicado por outro (- 0 no Quadro 2), sem haver motivo aparente que justifique o dano causado ao organismo. Cassini (2005) cita como exemplo o fenômeno conhecido como maré vermelha, ilustrado na Figura 12, que ocorre devido à liberação de toxinas na água por microalgas do tipo pirrófitas, quando essas encontram-se em estado de superpopulação. Esse fenômeno causa a morte por intoxicação de muitos seres aquáticos (CASSINI, 2005).

Figura 12 – Exemplo de amensalismo



Fonte: Schumann (2018)

A Figura 12 mostra o fenômeno da maré vermelha que ocorreu na costa de San Diego, no estado da Califórnia.

2.2.1.7 *Sociedade*

É uma interação intraespecífica harmônica na qual os indivíduos não se encontram anatomicamente ligados entre si, ou seja, podem se locomover de forma livre e independente um do outro (BLUME; REINIGER, 2007). São exemplos de espécies que se organizam em sociedades as abelhas (vide Figura 13), os cupins e as formigas.

Figura 13 – Exemplo de sociedade



Fonte: Maggydurch (2016)

De acordo com Blume e Reinger (2007), nas sociedades os indivíduos manifestam polimorfismo (formas diferentes) de acordo com as castas que ocupam dentro da sociedade. Na Figura 13 é possível observar que a abelha no centro, que é a rainha, é maior e possui tamanho e coloração diferente em relação às outras ao redor, que são operárias.

2.2.1.8 Colônia

É uma interação intraespecífica harmônica na qual os indivíduos encontram-se anatomicamente ligados entre si e há profunda interdependência fisiológica (BLUME; REINIGER, 2007). Em outras palavras, não conseguem sobreviver de maneira desassociada. Pode-se citar como exemplos as caravelas e os recifes de coral (BLUME; REINIGER, 2007).

Figura 14 – Exemplo de colônia



Fonte: Doubilet (2017)

As colônias de diferentes espécies de corais, juntas, formam a estrutura vista na Figura 14, que é denominada recife. Os recifes são *hotspots* de biodiversidade, o que significa que são ambientes com grande diversidade de espécies.

2.2.1.9 Canibalismo

É uma interação intraespecífica desarmônica em que indivíduos da mesma espécie predam um ao outro (BLUME; REINIGER, 2007). Assim como o predatismo, poderia

esta interação poderia ser representada pelos símbolos + -, uma vez que um dos indivíduos se beneficiará da morte do outro. O canibalismo é muito comum nos ecossistemas e acontece, por exemplo, entre indivíduos jovens de trairão¹¹, conforme relatado por Ramos *et al.* (2018). Pode-se citar também como exemplo a viúva-negra¹², ilustrado pela Figura 15.

Figura 15 – Exemplo de organismo que pratica canibalismo



Fonte: Rocha (2017)

A Figura 15 exibe uma viúva-negra, aranha venenosa que devora o macho após a cópula – e por esse motivo recebeu o nome popular de viúva-negra (BLUME; REINIGER, 2007).

Tendo sido expostos os conceitos e as interações ecológicas que envolvem a temática da agrobiodiversidade, passa-se às questões da representação da informação na forma de metadados, que são relevantes ao desenvolvimento desta pesquisa.

2.2.2 Considerações sobre as interações ecológicas

É preciso ampliar a discussão sobre as interações, além do que já foi discutido nesta dissertação, para envolver outras classificações adotadas na literatura para as interações ecológicas, como os conceitos de simbiose, esclavagismo, detritivoria, inquilinismo e foresia, apresentados por autores como Begon, Townsend e Harper

¹¹ *Hoplias intermedius* (Günther, 1864).

¹² *Latrodectus mactans* (Fabricius, 1775).

(2007), Bronstein (2015), Cain, Bowman e Hacker (2018) e Cassini (2005). Há questionamentos na literatura sobre como categorizar certos tipos de interações entre os seres vivos, e se as classificações consolidadas na literatura estão corretas, como é o caso do líquen (vide Figura 10), sobre o qual há dúvidas se é de fato uma relação de mutualismo ou uma relação de esclavagismo (SPIELMANN, 2006).

Além disso, o grau de afetação aos indivíduos envolvidos na interação pode variar e não há conhecimento preciso sobre isso na literatura. A classificação das interações facilita o entendimento, mas deve-se entender que, como citado por Bronstein (2015), essa classificação pode ser limitadora e levar ao entendimento que os domínios de cada tipo de interação estão delineados e bem definidos, o que não é verdade.

Na Seção 2.3 foram relacionados os conteúdos sobre representação da informação e metadados que foram aplicados ao longo do desenvolvimento desta pesquisa.

2.3 Representação da informação em metadados

A informação é elemento intrínseco às relações humanas e, quando veiculada ou armazenada em meio digital, necessita ser bem descrita para que possa ser eficientemente recuperada e acessada.

A representação da informação é um campo de estudo da Biblioteconomia e da CI, que nos últimos tempos tem sido estudada também por campos relacionados às Tecnologias de Comunicação e Informação e Linguística Computacional (LOURENÇO, 2016). Para a autora, inicialmente os estudos sobre metadados estavam direcionados à descrição bibliográfica, mas, com a criação dos repositórios e bibliotecas digitais, novos padrões de metadados começaram a ser desenvolvidos para atender a diferentes áreas especializadas.

De acordo com Lourenço (2016), os estudos no campo da representação da informação têm buscado alcançar a interoperabilidade entre repositórios de informação, aplicando para isso “metadados tanto descritivos quanto estruturais e de conteúdo” (LOURENÇO, 2016, p. 98), auxiliando na recuperação dos recursos informacionais.

A recuperação da informação depende, principalmente, do uso correto de metadados. De acordo com Zeng e Qin (2008), metadados descrevem o que a informação é, sobre

o que trata e onde pode ser localizada. Em outras palavras, metadados são informações criadas para descrever e definir dados, de acordo com a International Organization for Standardization (ISO, 2004) e Zeng (2015), e permitir que pessoas manipulem esses dados para obter conhecimento (RILEY, 2017). “Metadados são a chave para a funcionalidade dos sistemas que proveem conteúdo, possibilitando aos usuários encontrar itens de interesse, registrar informações essenciais sobre eles e compartilhar informações com outros” (RILEY, 2017, p. 2, tradução do proponente). Em suma, os metadados são um meio de interação entre o usuário e o sistema.

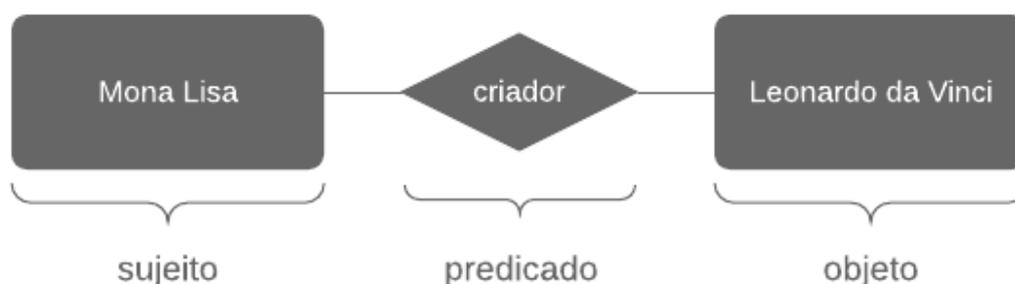
A trivial definição de metadados como *dados sobre dados* é insuficiente para definir o conceito de metadados, segundo Pomerantz (2015), uma vez que “os dados são apenas informação potencial, bruta e não processada, antes de alguém ser realmente informado por ela” (POMERANTZ, 2015, p. 26, tradução do proponente). Esta definição ‘dados sobre dados’ aparece frequentemente na literatura, inclusive na norma ISO/IEC 11179-4, com um acréscimo: “dados que definem e descrevem outros dados” (ISO; IEC, 2004, p. 2, tradução do proponente). Considerando-se o argumento de Pomerantz (2015), esta definição da ISO/IEC (2004) contém um erro semântico: dados são informação não processada, bruta, portanto, incapaz de definir e informar. Os usos e definições do termo *metadata* foram apontados por Zeng (2015) como variantes em diferentes comunidades. A autora define o termo de uma maneira que mais se enquadra ao perfil desta pesquisa: “informação sobre coisas específicas” (ZENG, 2015, *online*, tradução do proponente). E acrescenta: “De um modo geral, os metadados encapsulam as informações que descrevem qualquer entidade portadora de informações” (ZENG, 2015, *online*, tradução do proponente). A definição de Zeng (2015) é menos ambígua que a definição da ISO/IEC (2004), entretanto, a definição mais completa e coerente foi estabelecida pela National Information Standards Organization - NISO (2004):

Os metadados são informações estruturadas que descrevem, explicam, localizam ou facilitam a recuperação, o uso ou o gerenciamento de um recurso de informações. Os metadados são chamados de dados sobre dados ou informações sobre informações (NISO, 2004, p. 1, tradução do proponente).

Considerando-se os dados como informações potenciais ou como objetos potencialmente informativos, infere-se que “metadados são declarações sobre um objeto potencialmente informativo” (POMERANTZ, 2015, p. 26, tradução do

proponente). Para que essas declarações sejam feitas, é necessário que haja duas condições básicas: “alguma coisa para se fazer uma declaração sobre” e “algo para se dizer sobre essa coisa” (POMERANTZ, 2015, p. 27, tradução do proponente). Este “objeto potencialmente informativo” é algo sobre o qual [está sendo feita] uma declaração. Powell *et al.* (2007) denomina este objeto como recurso. Uma ou várias declarações sobre o recurso, compõe a descrição (POMERANTZ, 2015; POWELL *et al.*, 2007). De acordo com a NISO (2004), Pomerantz (2015) e Powell *et al.* (2007), uma declaração possui três partes básicas, conforme a Figura 16.

Figura 16 – As três partes de uma declaração



Fonte: Pomerantz (2015, p. 27, tradução e design do proponente)

A primeira parte da declaração é o sujeito, ou seja, a entidade que está sendo descrita, representado na Figura 16 pela obra de arte Mona Lisa. A segunda parte é o relacionamento (predicado) que há entre o recurso informacional e o objeto, identificado na Figura 16 como a relação de criador, ou seja, o criador da obra Mona Lisa. A terceira parte é o objeto que está vinculado ao sujeito por meio do predicado. “O objeto é uma entidade utilizada para descrever o sujeito”. (POMERANTZ, 2015, p. 28, tradução do proponente). Na Figura 16, o objeto é Da Vinci, o criador da obra de arte Mona Lisa. Por ser composta por três partes, esta estrutura é denominada *tripla* (sujeito-predicado-objeto) e representa a semântica dos elementos de metadados (NISO, 2004).

Os metadados desempenham as seguintes funções, de acordo com a NISO (2004 citado por ZENG; QIN, 2008, p. 4, tradução do proponente):

- a) descrever o que são os recursos informacionais e sobre o que eles tratam, e organizá-los de acordo com critérios controláveis;
- b) permitir que recursos informacionais possam ser encontrados por meio de critérios relevantes, agregando recursos semelhantes, e provendo caminhos para se encontrar a informação desejada;
- c) facilitar a troca de dados e permitir a interoperabilidade;
- d) prover identificação digital e descrição para arquivar e preservar o

recurso informacional.

Quanto às funções que exercem, os metadados podem ser classificados em três tipos: descritivos, administrativos, estruturais (NISO, 2004; ZENG, 2015; ZENG; QIN, 2008).

De acordo com Riley (2017), os metadados descritivos permitem a descrição de um recurso informacional para que seja possível recuperá-lo ou entendê-lo, sendo possível compartilhar os dados de forma automática, ou seja, que haja interoperabilidade. São exemplos de metadados descritivos: título, autor, assunto, tipo de documento e data de publicação (RILEY, 2017).

Os metadados administrativos se subdividem em três categorias (RILEY, 2017, p. 6, tradução do proponente): a) técnicos, “para decodificação e renderização de arquivos”, como os metadados *tipo_de_arquivo*, *tamanho_do_arquivo* e *data_de_criação*; b) de preservação, aplicados à “gestão de arquivos a longo prazo”, como os metadados *soma_de_verificação* e *evento_de_preservação*; c) e de direitos, que descrevem “propriedade intelectual sobre o conteúdo”, como *status_de_copyright*, *termos_de_licença*, *detentor_de_direitos*. Já os metadados estruturais descrevem os componentes do recurso informacional e as relações entre as partes desse recurso, como páginas sequenciais, capítulos, sumários, listas de ilustrações e quadros, índice, entre outros. Esses metadados podem dar suporte a pesquisa avançada em bases de dados e permitem a navegação dentro dos recursos informacionais (METADATA..., 2017; RILEY, 2017).

Por último, os metadados estruturais “indicam como objetos compostos são colocados juntos, por exemplo, como as páginas são ordenadas para formar capítulos” (NISO, 2004, p. 1, tradução do proponente). Em documentos, por exemplo, palavras podem ser marcadas para indicar certa semântica que poderia passar despercebida, caso não houvesse a marcação. São exemplos: o uso das aspas para destacar trechos de citação direta, o uso do itálico para dar ênfase ao significado de alguma palavra ou para apontar que aquela palavra, naquele contexto, possui função sintática diferente da usual (RILEY, 2017).

Os metadados podem ser armazenados e compartilhados. Segundo Riley (2017), as formas mais comuns de se armazenar e compartilhar metadados são bases de dados relacionais, *eXtensible Markup Language* (XML), *Linked Data* e *Resource Description*

Framework (RDF). Essas três estratégias de armazenamento e compartilhamento de metadados estão descritas nas Seções 2.3.1 a 2.3.3.

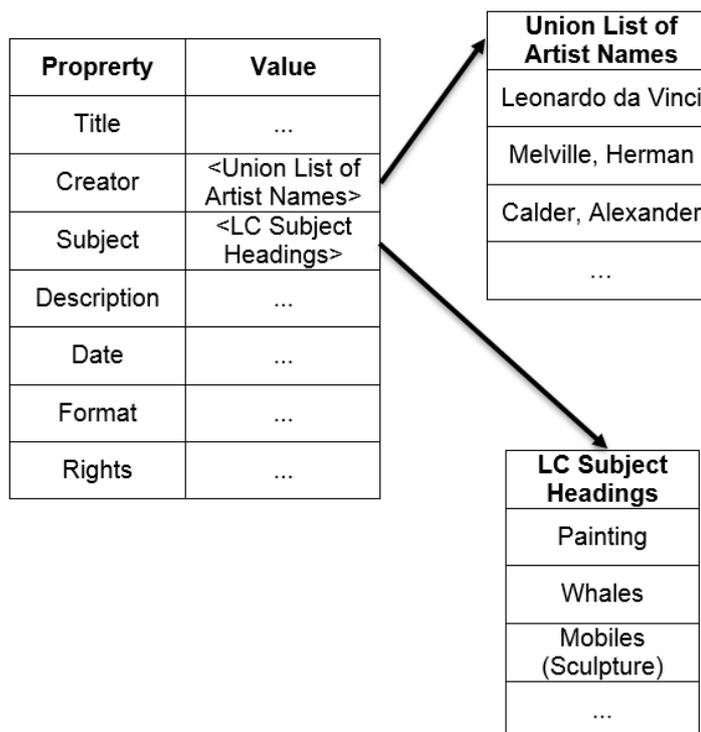
2.3.1 Bases de dados relacionais

As bases de dados relacionais são as formas mais comuns de aplicação de metadados em sistemas, de acordo com Riley (2017). Nas bases de dados, os metadados são organizados em coleções, denominadas *registro*, e “podem ser armazenados como campos em tabelas de base de dados relacional” (RILEY, 2017, p. 8, tradução do proponente). O uso de metadados em bases de dados relacionais deve “maximizar a eficiência do armazenamento” e ao mesmo tempo “otimizar a performance da *query*”, ou seja, deve tornar o armazenamento de dados mais eficiente de forma a poupar espaço no sistema e evitar duplicatas desnecessárias e permitir a aplicação de estratégias de busca que permitam recuperar informações de maneira mais eficiente (RILEY, 2017, p. 8, tradução do proponente).

Em um banco de dados relacional, um relacionamento pode ser estabelecido entre um campo e um conjunto de valores tabulados, a fim de controlar quais valores podem ser atribuídos a esse campo. Em outras palavras, a tabela à qual um campo se refere se torna um vocabulário controlado e os valores atribuídos às células nesse campo só podem ser atribuídos a partir desse vocabulário controlado. Os bancos de dados relacionais são especialmente úteis para garantir a qualidade dos dados: os arquivos de autoridade de nomes, por exemplo, evitam que os nomes sejam escritos de maneira incorreta, eliminam a ambiguidade sobre diferentes indivíduos com o mesmo nome etc. (POMERANTZ, 2015, p. 138, tradução do proponente).

A Figura 17 é um exemplo de uma tabela de uma base de dados relacional.

Figura 17 – Tabela de uma base de dados relacional



Fonte: Pomerantz (2015, p. 139)

No exemplo da Figura 17, o campo *creator* só pode receber valores da *Union List of Artist Names*, que é um esquema de codificação de nomes de artistas. Dessa maneira, todos os nomes de artistas em registros da base de dados estarão padronizados de acordo com esse esquema de codificação. Esse mesmo efeito se aplica ao campo *subject*, cujos valores devem ser padronizados de acordo com a *Library of Congress Subject Headings*, que é um vocabulário controlado de assuntos que abrange todas as áreas do conhecimento.

2.3.2 eXtensible Markup Language (XML)

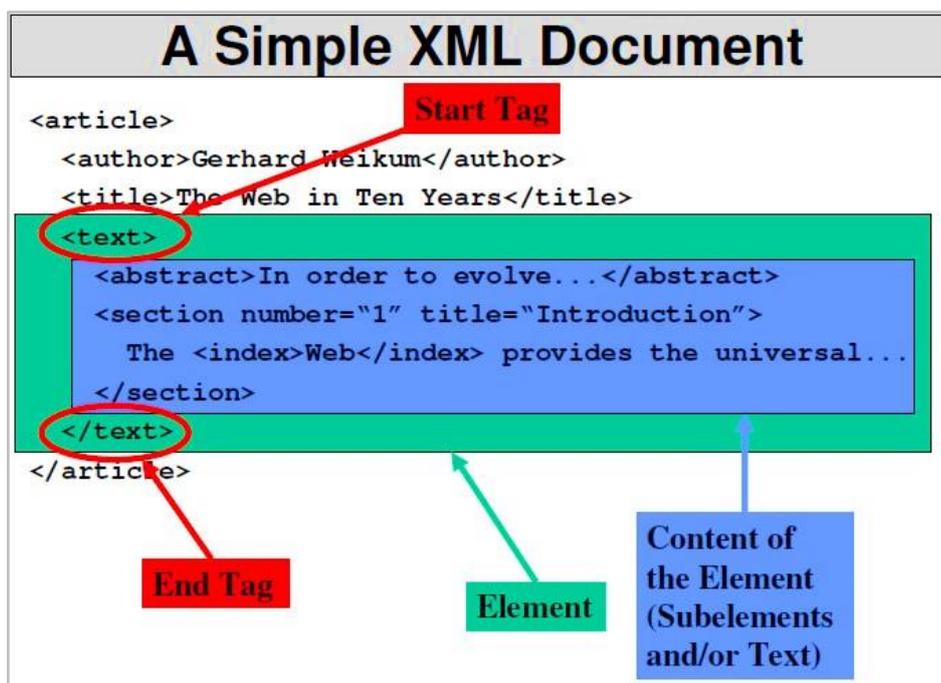
XML é uma linguagem de marcação de documentos, ou seja, tem a função de identificar as estruturas de um documento, desenvolvida como um mecanismo para codificação, transferência e armazenamento de metadados em sistemas, tendo como base a norma ISO 8879 (BRAY *et al.*, 2008; ISO, 1986; RILEY, 2017; WALSH, 1998). Foi criada em 1996 pelo Standard Generalized Markup Language (SGML) Editorial Review Board, grupo dedicado à revisão da norma ISO 8879 (1986), grupo de trabalho inserido no World Wide Web Consortium (W3C); atualmente, o grupo é denominado XML Working Group (BRAY *et al.*, 2008). Uma linguagem de marcação não é uma linguagem em seu sentido literal: é “[...] um vocabulário controlado que permite que

as instruções sejam incorporadas no texto de um documento, para que haja uma separação clara entre o texto e as instruções” (POMERANTZ, 2015, p. 146, tradução do proponente).

XML define elementos, tags que significam que os valores dentro deles têm um certo significado. Os elementos também podem ter outros elementos dentro deles, e é a partir desse recurso que os documentos XML ganham sua estrutura. Um documento XML é uma árvore que começa com um único elemento raiz. Outros elementos e valores se ramificam dessa raiz original, construindo uma estrutura aninhada que contribui para o significado dos valores de metadados no documento. Os elementos XML podem receber atributos, que normalmente também possuem seus próprios valores. Um atributo XML e seu valor refinam o significado do elemento no qual eles aparecem. O XML oferece suporte ao multilinguismo¹³ de metadados, fornecendo um atributo predefinido para indicar o idioma em que o valor de um elemento é exibido. Como nos bancos de dados relacionais, um documento XML descrevendo uma coisa definida é conhecido como um registro de metadados (RILEY, 2017, p. 8, tradução do proponente).

A Figura 18 apresenta de maneira esquemática a semântica da estrutura XML.

Figura 18 – A estrutura semântica do XML



Fonte: Schenkel (2003)

De acordo com Schenkel (2003), um campo de um registro XML sempre possui uma *start tag* e uma *end tag* como delimitadores (marcam o início e o fim do campo). Essas

¹³ Coexistência de valores em dois ou mais idiomas em um registro de metadados.

tags representam o elemento propriamente dito, e os valores compreendidos entre a *start* e *end tag* representam o conteúdo do elemento, ou seja, os dados.

O XML é uma das formas de se publicar registros de metadados DwC, de acordo com o Darwin Core Task Group (2009e). O Esquema 1 tem em si um exemplo de registro DwC no formato XML.

Esquema 1 – Registro de metadados XML

```
<?xml version="1.0"?>
<dwr:SimpleDarwinRecordSet
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://rs.tdwg.org/dwc/xsd/simpledarwincore/ http://rs.tdwg.org/dwc/xs
  xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:dwc="http://rs.tdwg.org/dwc/terms/"
  xmlns:dwr="http://rs.tdwg.org/dwc/xsd/simpledarwincore/">
  <dwr:SimpleDarwinRecord>
    <dcterms:type>PhysicalObject</dcterms:type>
    <dcterms:modified>2009-02-12T12:43:31</dcterms:modified>
    <dcterms:rightsHolder>Museum of Vertebrate Zoology</dcterms:rightsHolder>
    <dcterms:rights>Creative Commons License</dcterms:rights>
    <dwc:institutionCode>MVZ</dwc:institutionCode>
    <dwc:collectionCode>Mammals</dwc:collectionCode>
    <dwc:occurrenceID>urn:catalog:MVZ:Mammals:14523</dwc:occurrenceID>
    <dwc:basisOfRecord>PreservedSpecimen</dwc:basisOfRecord>
    <dwc:country>Argentina</dwc:country>
    <dwc:countryCode>AR</dwc:countryCode>
    <dwc:stateProvince>Neuquén</dwc:stateProvince>
    <dwc:locality>25 km al NNE de Bariloche por Ruta 40 (=237)</dwc:locality>
  </dwr:SimpleDarwinRecord>
</dwr:SimpleDarwinRecordSet>
```

Fonte: Darwin Core Task Group (2009e)

O Esquema 1 ilustra a definição de Riley (2017) para XML: o elemento `SimpleDarwinRecordSet` é a raiz do esquema, e tudo o que está compreendido entre a *start tag* e a *end tag* `</dwr:SimpleDarwinRecordSet>` é o conteúdo do elemento, que é formado por subelementos e texto. Os atributos `<xmlns:>` são declarações do tipo XML *namespace*, ou seja, são o código de um namespace válido. O *namespace* é um URI, que serve para evitar a ambiguidade entre os nomes dos elementos no esquema (DARWIN CORE TASK GROUP, 2009a). Já o atributo `xsi:` indica a localização dos esquemas de elementos e atributos especificados pelo namespace e utilizados no registro XML em questão, que são três: o Simple Darwin Core (`dwr:`), Darwin Core Terms (`dwc:`) e Dublin Core Terms (`dcterms:`). Em seguida no Esquema 1, nota-se o elemento *SimpleDarwinRecord*, que inclui subelementos. O conteúdo desse elemento é delimitado pelas *tag* e a *end tag* `<dwr:SimpleDarwinRecord>`. Para destacar a

hierarquia que existe entre os elementos e melhorar a legibilidade, é aplicado um recuo que aumenta da esquerda para a direita conforme o grau de subordinação do elemento, o que é conhecido como indentação (GUIMARÃES, 2005). Os subelementos com seus respectivos valores “refinam o significado do elemento no qual eles aparecem” (RILEY, 2017, p. 8, tradução do proponente).

2.3.3 *Linked data e RDF*

Linked data é um conceito que foi introduzido em 2006 por Tim Berners-Lee. *Linked data* (dados ligados, em português) consiste em estabelecer conexões entre dados na web para que, ao procurar certo dado, seja possível encontrar dados relacionados (BERNERS-LEE, 2006). Os dados ligados podem ser formatados em Hypertext Markup Language (HTML) ou RDF, e para ambos os casos é recomendável que sejam seguidas as seguintes regras:

use URIs como nomes para coisas; b) use HTTP URIs para que as pessoas possam procurar esses nomes; c) quando alguém procura um URI, forneça informações úteis, usando os padrões (RDF, SPARQL); d) inclua links para outros URIs para que eles possam descobrir mais coisas (BERNERS-LEE, 2006, *online*, tradução do proponente).

De acordo com Riley (2017), as operações com *linked data* dependem muito do RDF. “O RDF é um conjunto de especificações [da] W3C projetado para o uso de metadados na Web Semântica” (RILEY, 2017, p. 10, tradução do proponente).

RDF é um modelo padrão para intercâmbio de dados na Web. O RDF possui recursos que facilitam a mesclagem de dados, mesmo que os esquemas subjacentes sejam diferentes, e suporta especificamente a evolução dos esquemas ao longo do tempo, sem exigir que todos os consumidores de dados sejam alterados (RDF WORKING GROUP, 2014, *online*, tradução do proponente).

O RDF Diferencia-se do XML quanto a estruturação e semântica:

[...] enquanto o XML modela a informação como uma árvore, o RDF a modela como um grafo, com pequenos bits de informação, cada um conectado a outros pequenos bits de informação. [...] Um grafo RDF é [...] [composto] de triplas individuais, em que um sujeito (a entidade sobre a qual a tripla está) está conectado a um objeto (a entidade à qual está relacionado) com um predicado (um descritor de um relacionamento). Uma tripla pode representar conteúdo como: “A Tempestade” (sujeito) “foi escrito por” (predicado) “William Shakespeare” (objeto) (RILEY, 2017, p. 10, tradução do proponente).

Desta forma, as triplas do esquema RDF podem ser ilustradas como a Figura 16, apresentada anteriormente (para entender a lógica de representação completa do RDF, veja a Figura 25).

A lógica dos dados ligados é aplicada no DwC o tempo todo, para ligar as classes do padrão de metadados, elementos e remeter a fontes externas. Portanto, é de fundamental importância que ao se criar uma extensão de metadados DwC, aplique-se a lógica do *linked data* para vincular a extensão às demais classes de metadados do DwC.

2.4 Princípios teóricos para a criação de metadados

Os metadados podem ser criados automaticamente por computador, como índices de palavras, ou por seres humanos, como catálogos de registros desenvolvidos nas bibliotecas (DUVAL *et al.*, 2002; GILLILAND, 2016).

Os mecanismos de pesquisa na Web coletam e indexam uma parte significativa da Internet e fornecem acesso a índices de baixo custo, geralmente em um modelo suportado por anunciantes. Essa indexação pode ser vista como um tipo de metadados e, para muitas necessidades de informações, fornece uma solução surpreendentemente econômica para a descoberta de recursos (DUVAL *et al.* 2002, p. 12, tradução do proponente).

Nesta dissertação serão discutidos os princípios em torno da criação humana de metadados.

Antes de adentrar na discussão sobre os princípios de criação de metadados, o entendimento de alguns conceitos básicos é indispensável:

- a) esquema de metadados: estrutura formal composta por regras sobre como o padrão de metadados deve ser aplicado, definindo, dentre outras coisas, quais vocabulários controlados utilizar na descrição de valores, (NISO, 2007; POMERANTZ, 2015; TASK FORCE ON METADATA, 2000);
- b) elementos: são os predicados que se pode fazer sobre o sujeito e que cumprem o papel de definir, identificar e representar os valores atribuídos ao sujeito (ISO; 2004; POMERANTZ, 2015);
- c) valores: são os dados propriamente ditos (POMERANTZ, 2015; POWELL *et*

al., 2007).

Para se criar metadados, é necessário que sejam seguidos alguns princípios de maneira a garantir a funcionalidade dos metadados. Segundo Zeng e Qin (2008), são três os princípios fundamentais para a criação de metadados:

- a) simplicidade: selecionar apenas os dados indispensáveis a serem representados e criar metadados de fácil implementação;
- b) extensibilidade: esquemas de metadados devem permitir a incorporação de novos metadados, que podem ser provenientes de padrões de metadados já existentes ou criados para uso local;
- c) interoperabilidade: possibilitar a leitura dos dados representados no padrão para a troca de dados entre diferentes sistemas com a menor perda de conteúdo possível.

Já a NISO (2007) define seis princípios fundamentais que qualquer esquema de metadados deve ter como embasamento:

- a) os metadados devem estar de acordo com as necessidades de uso da comunidade, possibilitando descrição eficiente da coleção de forma a facilitar o uso da mesma pelos usuários;
- b) os metadados devem possibilitar a interoperabilidade entre sistemas de informação de diferentes instituições que armazenam coleções de registros;
- c) os padrões de metadados devem usar controle de autoridade e vocabulários controlados (também denominados esquemas de codificação) como ferramentas para se atribuir valores aos campos;
- d) um registro de metadados deve incluir elementos que descrevam o status de direitos autorais do objeto digital descrito;
- e) os metadados devem facilitar a “gestão, curadoria e preservação dos objetos de uma coleção” (NISO, 2007, p. 83, tradução do proponente);
- f) e por último, “os registros de metadados são objetos em si e, portanto, devem

ter as qualidades de bons objetos, incluindo autoridade, autenticidade, arquivabilidade, persistência e identificação única” (NISO, 2007, p. 84, tradução do proponente).

Após análise desses princípios, constata-se que eles são uma versão mais detalhada dos princípios apontados por Zeng e Qin (2008).

A NISO (2007) recomenda os princípios que devem ser seguidos para a construção de metadados, mas não apresenta informações concretas sobre como os metadados devem ser estruturados. Isso fica a cargo da ISO/IEC 11179-4 (2004), que correlaciona recomendações em duas categorias. A primeira, determina o que a definição de um elemento de um esquema de metadados deve ter (ISO; IEC, 2004, p. 3, tradução do proponente):

- a) ser indicado no singular;
- b) declarar o que é o conceito, não apenas o que não é;
- c) ser indicado como frase ou sentença descritivas [que esclareçam o significado do termo usado como elemento];
- d) conter apenas abreviações comumente entendidas;
- e) ser expresso sem incorporar definições de outros metadados ou conceitos subjacentes.

Já a segunda, indica o que a definição de elemento de um esquema de metadados deveria ter, ou seja, o que é recomendável (ISO; IEC, 2004, p. 3, tradução do proponente):

- a) estabelecer o significado essencial do conceito;
- b) ser preciso e inequívoco;
- c) ser conciso;
- d) ser capaz de permanecer sozinho [e ser compreendido sem demais explicações];
- e) ser expresso sem incorporar justificativa, uso funcional ou informações processuais;
- f) evitar raciocínio circular;
- g) usar a mesma terminologia e estrutura lógica consistente para definições relacionadas;
- h) ser apropriado para o tipo de item de metadados que está sendo definido.

Desta forma, a ISO/IEC 11179-4 (2004) apresenta recomendações de sintaxe e semântica para a construção de metadados que, se seguidas, o resultado alcançado serão metadados que atendem aos princípios de simplicidade e interoperabilidade propostos por Zeng e Qin (2008). Essas recomendações serão aplicadas na proposta elaborada nesta pesquisa.

Além dos princípios que orientam a criação de esquemas de metadados, há um formato básico a ser seguido para a construção desses esquemas. Cada elemento de metadados deve possuir, segundo Zeng e Qin (2008), um conjunto de atributos que o identifique, com um nome, uma *label* e uma definição. Alguns esquemas incluem também uma nota para registrar as mudanças do conceito do elemento ao longo do tempo. Por exemplo, cada elemento do padrão DC contém o seguinte conjunto de atributos: nome, URI, *label*, definição, comentário e referências (DCMI USAGE BOARD, 2012; ZENG; QIN, 2008). A estrutura de um elemento do padrão DC pode ser observada no Quadro 3.

Quadro 3 – Conjunto de atributos de um elemento do Dublin Core

Term Name: date	
URI:	http://purl.org/dc/elements/1.1/date
Label:	Date
Definition:	A point or period of time associated with an event in the lifecycle of the resource.
Comment:	Date may be used to express temporal information at any level of granularity. Recommended best practice is to use an encoding scheme, such as the W3CDTF profile of ISO 8601 [W3CDTF].
References:	[W3CDTF] http://www.w3.org/TR/NOTE-datetime

Fonte: DCMI Usage Board (2012)

O Quadro 3 apresenta a estrutura do elemento *date*. O primeiro atributo é o nome do termo *date*; o segundo atributo é o URI, identificador que torna único o elemento; o terceiro atributo é uma *label*, etiqueta atribuída ao elemento para ser usada como nome de campo em um registro (geralmente é igual ou semelhante ao nome do termo); o quarto atributo é a definição, que especifica o tipo de dado que o elemento *date* representa e *comment* especifica as melhores práticas recomendadas para registro de dados do tipo *date*; e o quinto atributo apresenta uma referência do esquema de codificação de datas recomendado no atributo *comment* para formatar os dados para o campo de data.

Como mencionado anteriormente nesta seção, o esquema de metadados determina um conjunto de regras e diretrizes para padronizar e normalizar o registro de dados. Os atributos *definição* e comentários (vide *definition* e *comments* no Quadro 3)

descrevem, geralmente, o significado do elemento e as práticas de registro de dados, apontando normas e vocabulários controlados a serem seguidos para a descrição dos dados (ZENG; QIN, 2008).

Os esquemas de metadados são formalizados em padrões, que atendem determinados domínios.

Os metadados são agrupados em esquemas, que os organizam, normalizam e os descrevem, criando padrões. Esses esquemas permitem o melhor entendimento da finalidade de cada metadado e são, geralmente, mantidos por instituição ou organização, permitindo assim o melhor uso e, principalmente, a troca de informação entre iniciativas que utilizam o mesmo esquema de metadados (INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA, 2013, *online*).

Os padrões de metadados contêm as regras básicas sobre como os metadados devem ser aplicados por todo mundo para que haja interoperabilidade entre sistemas de diferentes instituições, quando aplicando o mesmo padrão de metadados.

Há diversos padrões de metadados, empregados em diferentes ambientes e disciplinas (ZENG, 2015). O DC foi o primeiro padrão de metadados desenvolvido para descrição de recursos digitais na Web. Surgiu a partir de um congresso sediado na cidade de Dublin, Ohio, onde bibliotecários elaboraram uma proposta de metadados usando o Código de Catalogação Anglo-Americano (AACR2) como norma e o Machine Readable Cataloging (MARC21) como formato de padronização de dados (DCMI USAGE BOARD, 2012; ZENG; QIN, 2008).

O resultado da aplicação de um padrão de metadados é denominado registro. Um registro é um conjunto de declarações sujeito-predicado-objeto (vide Figura 16) feito sobre um recurso informacional (POMERANTZ, 2015; POWELL *et al.*, 2007). Em um registro, os metadados podem ser também denominados campos de dados, campos de metadados ou simplesmente campos (BACA, 2016). Uma característica importante do registro de metadados, segundo Pomerantz (2015), é que deve ser sempre elaborado com base em uma relação de um-para-um (1-1): um registro para um recurso informacional.

Com base nessa regra, deve existir apenas um registro para um recurso informacional. Entretanto, é possível que uma mesma entidade seja replicada em diversos suportes,

e isso preconiza a criação de diferentes registros para uma mesma entidade. Pomerantz (2015) indica como exemplo a obra Mona Lisa, de Leonardo da Vinci: o registro de metadados sobre a obra original, exposta no Louvre, deve ser diferente do registro de metadados de uma fotografia da mesma obra, exposta em um site ou em uma base de dados, uma vez que o registro de metadados é um objeto informacional tanto quanto a obra que ele descreve (NISO, 2007). A Figura 19 é um registro de metadados da obra original Mona Lisa.

Figura 19 – Registro de metadados da obra Mona Lisa

Domaine	peinture
Dénomination	tableau
Titre	PORTRAIT DE MONA LISA (1479-1528) ; DITE LA JOCONDE
Auteur/exécutant	LEONARDO DI SER PIERO DA VINCI ; VINCI Léonard de (dit)
Précision auteur/exécutant	Vinci, 1452 ; Amboise, 1519
Ecole	Italie
Période création/exécution	1er quart 16e siècle
Millésime création/exécution	1503 entre ; 1506 et
Genèse oeuvre en rapport ; reproduit en gravure	
Sources complémentaires	INHA
Historique commandé par le florentin Francesco del Giocondo, époux de Mona Lisa entre 1503 et 1506 ; nombreuses copies dont une conservée au Louvre ; gravé par Fauchery, par Filhol, par Landon	
Matériaux/techniques	peinture à l'huile ; bois
Dimensions	77 H ; 53 L
Sujet représenté	portrait (Mona Lisa, femme, à mi-corps, de trois-quarts, assis, accoudé, sourire, loggia, Italien) ; fond de paysage (montagne, rocher, cours d'eau, pont, plaine, route)
Lieu de conservation	Paris ; musée du Louvre département des Peintures
	 Musée de France
	au sens de la loi n°2002-5 du 4 janvier 2002
Statut juridique	propriété de l'Etat ; musée du Louvre département des Peintures
Anciennes appartenances	François 1er ; Couronne de France
Numéro d'inventaire	INV 779
Commentaires	légère diminution du tableau sur les côtés (environ 7 mm) ; acheté vraisemblablement vers 1519, après la mort de l'artiste
Bibliographie	HEYDENRICH 6 ; OTTINO DELLA CHIESA 31 ; VILLOT I 484 ; HAUTECOEUR 1601 ; C.S.I. 1981, P 192
Copyright notice	© Musée du Louvre, © Direction des Musées de France, 1999
Crédits photographiques	© Réunion des musées nationaux - utilisation soumise à autorisation ; © Hervé Lewandowski ; © Thierry Le Mage
	Demande de photographie et/ou de conditions d'utilisation
	Renseignements sur le musée
	Contact musée
	000PE025604

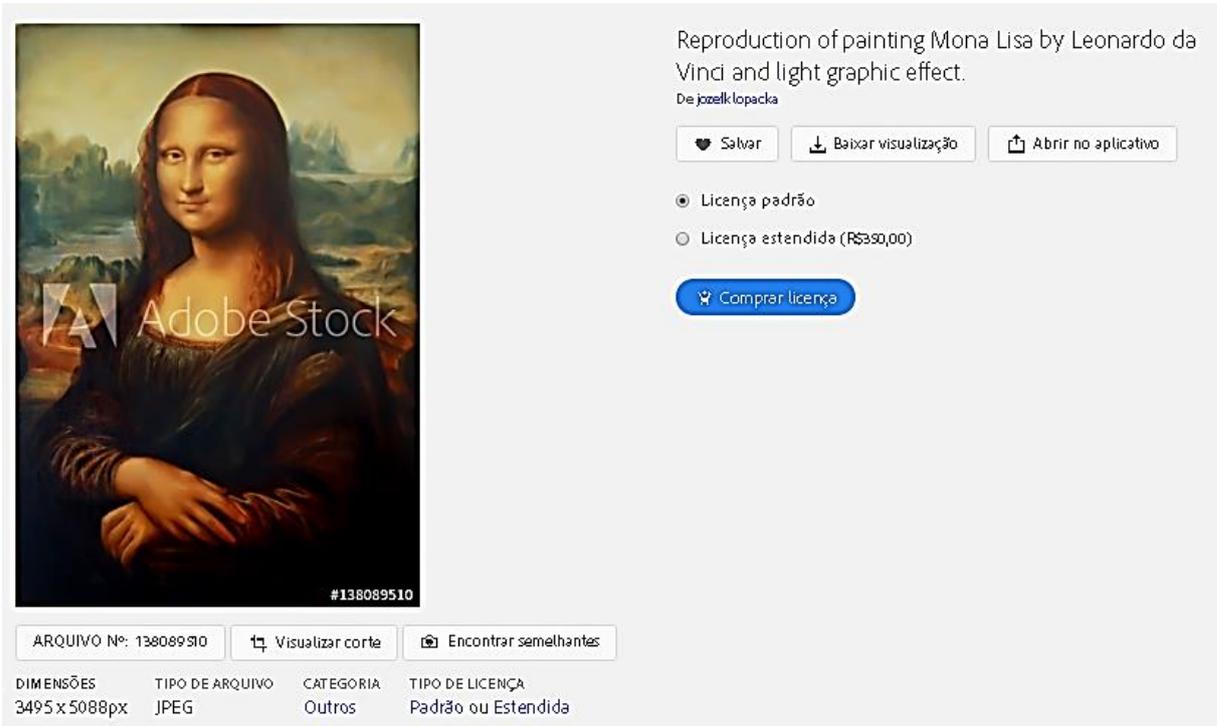
Fonte: Portail... (1999)

A Figura 19 mostra um registro detalhado de metadados sobre as dimensões da pintura, história, localização, autor e outros aspectos da obra. Os metadados permitem identificar que esse registro representa um objeto com as seguintes características: a) *dimensions*: 77H (altura) e 53L (largura); b) *matériaux/techniques*: *peinture à l'huile/bois*; c) *domaine*: *peinture*; d) *titre*: Portrait de Mona Lisa (1479-1528); dite la

Joconde; e) *auteur*. Leonardo di Ser Piero da Vinci. Esses metadados demonstram que o objeto representado é uma pintura a óleo em madeira, com 77 cm de altura por 53 cm de largura, cujo título é Portrait de Mona Lisa (1479-1528); dite la Joconde e o autor é Leonardo di Ser Piero da Vinci. Aliados aos outros metadados presentes na Figura 19, fica evidente que o registro de metadados representa a pintura original da Mona Lisa.

A Figura 20 também contém metadados para a entidade Mona Lisa.

Figura 20 – Metadados de uma foto da Mona Lisa



Reproduction of painting Mona Lisa by Leonardo da Vinci and light graphic effect.
De jozeklopacka

Licença padrão
 Licença estendida (R\$350,00)

ARQUIVO Nº: 138089510

DIMENSÕES	TIPO DE ARQUIVO	CATEGORIA	TIPO DE LICENÇA
3495 x 5088px	JPEG	Outros	Padrão ou Estendida

Fonte: Lopacka (2019)

Os metadados da Figura 19 e Figura 20 descrevem a mesma entidade: a obra Mona Lisa. Entretanto, os metadados da Figura 20 descrevem uma fotografia digital, e não a pintura original, caso da Figura 19. Nota-se que os valores atribuídos aos metadados são diferentes, com o título *Reproduction of painting Mona Lisa by Leonardo da Vinci and light graphic effect*, autor Jozeklopacka, tipo de arquivo JPEG com dimensões 3495x5088px.

Nesse caso, apesar de o sujeito ser o mesmo (Mona Lisa), muitos valores são diferentes, por se tratar de duas manifestações distintas de uma mesma entidade. Por isso é necessário que sejam criados registros de metadados diferentes que permitam

a identificação dos dois objetos: da obra original e da fotografia da obra.

Destaca-se que é importante que seja preservada a relação entre os dois objetos informacionais, e essa relação pode ser preservada por meio de *linked data* (POMERANTZ, 2015). Logo, a relação de um-para-um é diretamente influenciada pela localização do recurso informacional que, segundo Pomerantz (2015), pode ser *inside* ou *outside*.

A localização *inside* (dentro) refere-se aos metadados que estão presentes no interior do recurso informacional, ou seja, são intrínsecos ao objeto e permanecem estáticos, pois permanecem inalterados enquanto o objeto existir (POMERANTZ, 2015). O autor cita o exemplo do recurso informacional do tipo livro: os metadados presentes em um livro físico em uma biblioteca são do tipo *inside*, pois estão dentro do recurso informacional; os metadados presentes em um registro criado por uma base de dados que comercializa livros são exemplos de metadados *outside*, pois o registro existe fora do recurso informacional.

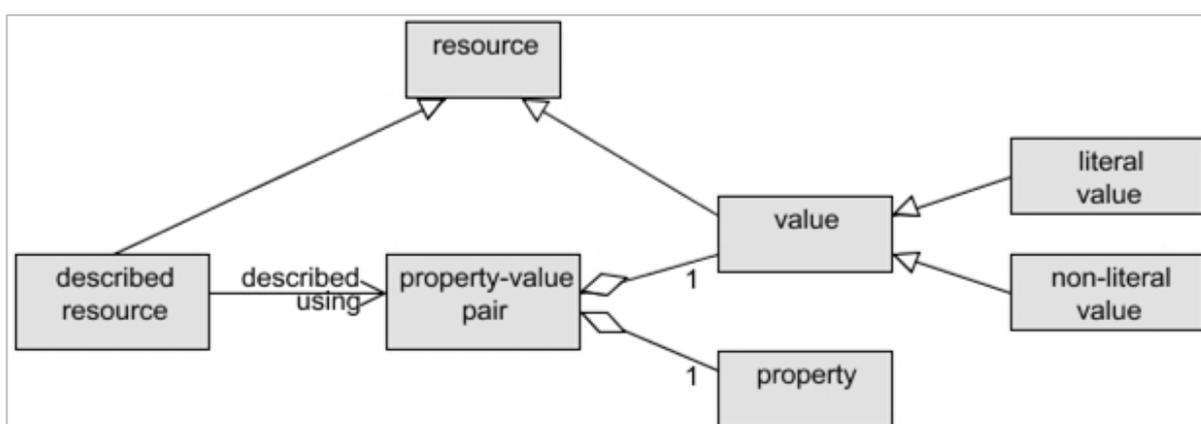
Pomerantz (2015) afirma que não é possível qualificar quais metadados são melhores, pois o contexto de uso determinará quais metadados devem ser utilizados. Entretanto, o autor reitera que registros *outside* podem conter metadados diferentes e, algumas vezes, ter valores incorretos para representar o recurso informacional. Justifica-se essa ocorrência pelo fato de que os metadados *inside* são criados juntamente com o recurso informacional, logo, sempre representam o conteúdo real e fidedigno, e os metadados *outside* são criados externamente ou podem ser compartilhados entre diferentes bases de dados, o que pode resultar em preenchimento de campos com valores incorretos.

O Dublin Core Abstract Model (DCAM) é a base do padrão de metadados DC e foi desenvolvido pela DCMI. Esse modelo, também denominado DCMI Abstract Model, é composto por três partes: a) DCMI Resource Model da Figura 21; b) DCMI Description Set Model da Figura 23; c) The DCMI Vocabulary Model da Figura 24. De acordo com Powell *et al.* (2007, *online*, tradução do proponente) esse modelo permite “obter uma melhor compreensão dos tipos de descrições [...] [que estão sendo codificadas] e facilita o desenvolvimento de melhores mapeamentos e traduções de sintaxe cruzada.”

De acordo com Coyle (2008), o DCAM esquematiza a estrutura de um elemento de metadados e o termo 'abstrato' presente no nome não significa que o modelo seja subjetivo ou confuso; significa, pois, que DCAM é um modelo conceitual e não um registro de metadados propriamente dito. A autora reitera que o modelo possui conceitos de difícil compreensão, o que dificulta sua aplicação na criação de metadados.

O primeiro submodelo DCMI Resource Model está representado na Figura 21.

Figura 21 – The DCMI resource model



Fonte: Powell *et al.* (2007)

Legenda: conectores que terminam em seta: leia-se *é* ou *é um*; conectores que terminam em losango: leia-se *contém* ou *tem um*.

O modelo da Figura 21 pode ser compreendido da seguinte maneira, de acordo com instruções prescritas por Powell *et al.* (2007, *online*): um recurso informacional descrito é um recurso (*resource*¹⁴) descrito por meio de pares propriedade-valor (*property-value-pairs*¹⁵). Um par propriedade-valor é a combinação de uma propriedade (*property*) e um valor (*value*). Um valor é sempre um recurso (ou um objeto, conforme o modelo de triplas ilustrado na Figura 16), e um par propriedade-valor é uma relação de um-para-um, conforme o modelo de triplas da Figura 16.

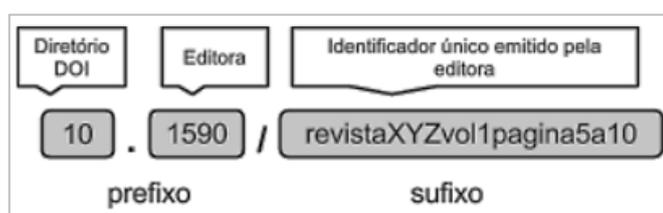
Um valor pode ser *literal*, quando é representado por *strings*, ou seja, conjuntos de

¹⁴ Pode ser qualquer objeto potencialmente informativo passível de identificação, como “um documento eletrônico, uma imagem, um serviço (por exemplo, ‘relatório meteorológico de hoje em Los Angeles’)” ou “seres humanos, corporações, livros encadernados em uma biblioteca. Nem todos os recursos informacionais podem ser recuperados por motores de busca eletrônicos” (POWELL *et al.* 2007, *online*, tradução do proponente).

¹⁵ “combinação de uma propriedade e um valor (*value*), usada para descrever uma característica de um recurso” (POWELL *et al.* 2007, *online*).

caracteres que representam um dado e que podem ser reconhecidos por humanos e máquinas, desde que haja uma *label* que o descreva e que certos padrões de formatação sejam seguidos (COYLE, 2008; POLLOCK, 2010; POWELL *et al.* 2007); ou *não literal*, dado definido em fonte externa ao registro de metadados e que deve seguir certos padrões de controle, podendo ser representado apenas por uma URI ou por termos provenientes de um vocabulário controlado (COYLE, 2008; POLLOCK, 2010; POWELL *et al.* 2007). De acordo com Coyle (2008), o dado do tipo não literal pode ser mais facilmente manipulado por máquina devido ao controle que esse tipo de dado exige. Por exemplo, o elemento *references* do padrão DwC, que indica referências de uma fonte de dados, poderia ser preenchido apenas com o Digital Object Identifier (DOI) da publicação fonte dos dados, que é um identificador único, padronizado e universal para textos científicos. Observe a Figura 22, desenvolvida pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) para ilustrar a estrutura do DOI.

Figura 22 – Componentes do nome DOI



Fonte: Ibict (2015)

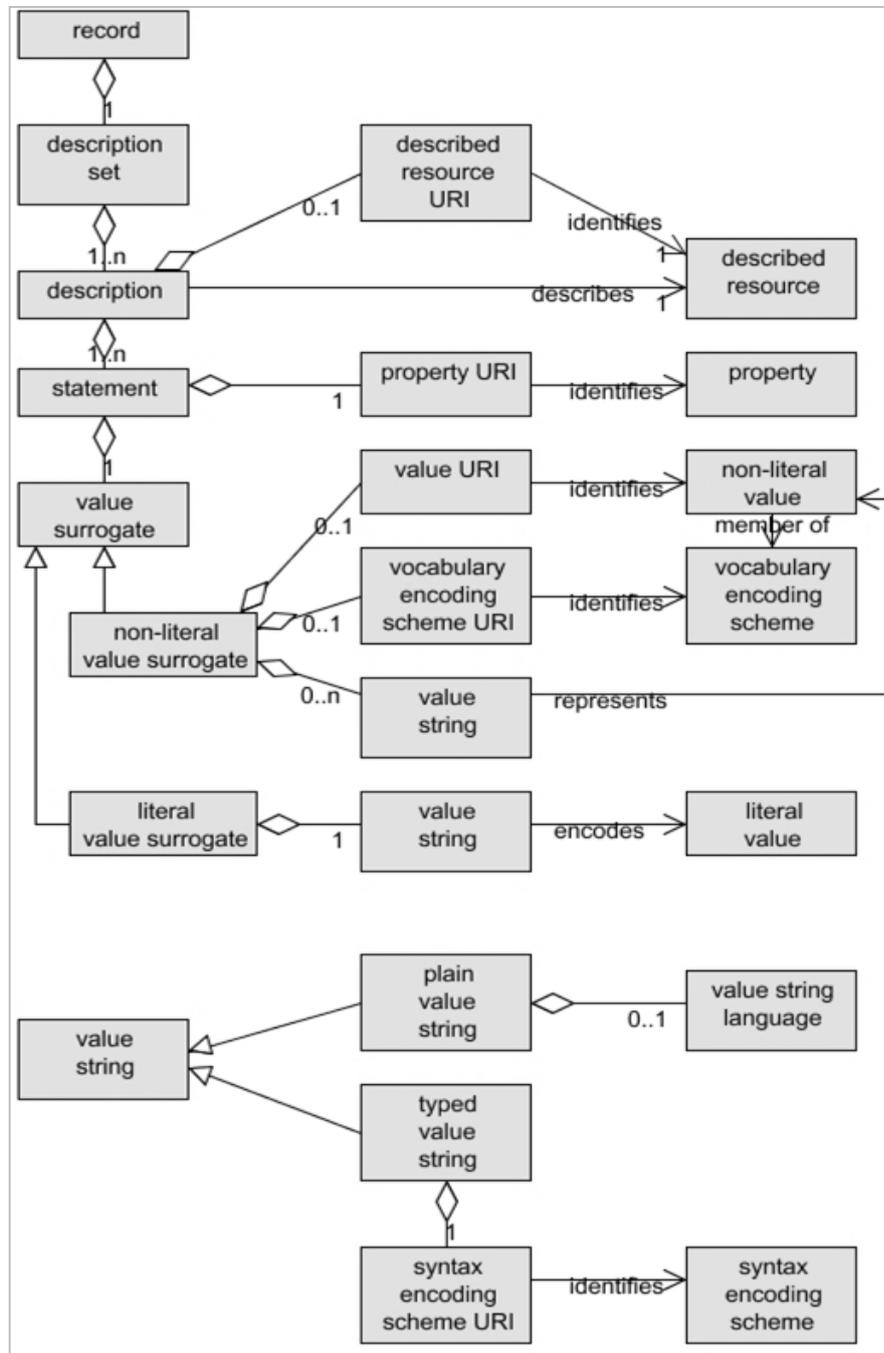
Como observado na Figura 22, o nome DOI, que é um URI atribuída ao artigo científico, apresenta sempre os mesmos elementos no mesmo formato, com valores diferentes para cada artigo científico. Desta forma, pode ser considerado um valor não literal.

O modelo da Figura 21 representa o conceito de recurso informacional aplicado no contexto da descrição por meio de metadados. Esse modelo apresenta as relações inerentes ao modelo de triplas representado pela Figura 16, que relaciona sujeito, predicado e objeto. O sujeito (recurso informacional descrito) e o objeto (valor) são ambos objetos informacionais, por isso o modelo apresenta ambos os conceitos com uma relação *é um* em relação ao conceito *recurso*.

O segundo submodelo DCMI Description Set Model representado na Figura 23 é mais específico e apresenta a estrutura do grupo de descritores. Está semanticamente

relacionado ao modelo da Figura 21, ou seja, são complementares.

Figura 23 – DCMI Description Set Model



Fonte: Powell *et al.* (2007)

Legenda: conectores que terminam em seta: leia-se *é* ou *é um*; conectores que terminam em losango: leia-se *contém* ou *tem um*.

O modelo do grupo de descritores presente na Figura 23 pode ser compreendido da seguinte maneira:

um *grupo de descritores* (*description set*) é um conjunto de um ou mais *descritores*, sendo que cada um descreve apenas um *recurso*

informacional. Um *descriptor* é composto por uma ou mais *declarações* (sobre um e apenas um *recurso informacional*) e zero ou um *URI de um recurso informacional descrito*¹⁶. Cada *declaração* instancia um *property-value-pair* e é composta por uma *property URI*¹⁷ e um *valor substituto (value surrogate)*. Cada *valor substituto* é um *valor literal substituto* ou um *valor não-literal substituto (non-literal value surrogate)*. Um *valor literal substituto* é um *valor substituto* para um *valor literal* e é composto por exatamente um string de valor (*value string*). Um *string de valor* é um *literal* que codifica o *valor literal*. Um *valor não-literal substituto* é um *valor substituto* para um *valor não-literal* e é composto de um ou zero *URI de valor* (um URI que identifica o *valor não-literal* associado à *propriedade*), zero ou um URI de vocabulário de esquema de codificação (um URI que identifica o vocabulário de esquema de codificação do qual o valor não literal é um membro) e zero ou mais cadeias de valores. Cada *string de valor* é um literal que representa o valor não-literal. Uma *string de valor* é uma *string* de valor simples ou uma *string* de valor digitada. Uma cadeia de valor simples pode ter uma linguagem de cadeia de valor associada que é uma tag de idioma ISO (por exemplo, en-GB). *Strings* de valor simples devem ser legíveis por humanos. Uma cadeia de valor digitado possui um URI de esquema de codificação de sintaxe associado que identifica um esquema de codificação de sintaxe (POWELL *et al.*, 2007, *online*, tradução do proponente)

Percebe-se que a descrição do autor, no entanto, é pouco clara e deixa subentendido o significado de diversos conceitos. Mesmo o glossário disponível ao final do documento de Powell *et al.* (2007) não é claro o suficiente para descrever alguns conceitos e deixa dúvidas sobre o seu significado. De forma mais objetiva, pode-se explicar o referido modelo da seguinte maneira: um registro de metadados representa um e apenas um recurso informacional; um registro é composto por um conjunto de descritores, sendo que cada um desses descritores é formado por uma ou mais declarações; cada declaração é formada por duas partes, que são uma propriedade e um valor (par propriedade-valor), sendo a propriedade o termo ou o nome do metadado (como título, data, autor e outros) identificada por um URI (COYLE, 2008).

O modelo da Figura 23 torna-se mais difícil de ser compreendido quando apresenta as relações do conceito valor. O conceito *valor surrogate*, por exemplo, não possui significado muito bem delimitado. Ao consultar o glossário do documento, verifica-se a seguinte definição para o conceito: “Valor surrogate: um valor substituto literal ou um valor substituto não-literal” (POWELL *et al.* 2007, *online*, tradução do proponente).

¹⁶ “Um URI que identifica um recurso informacional descrito” (POWELL *et al.* 2007, *online*, tradução do proponente).

¹⁷ “Um URI que identifica uma *property*” (POWELL *et al.* 2007, *online*, tradução do proponente).

A definição não é autoexplicativa e faz remissiva a outros dois conceitos do glossário, que são:

Valor *surrogate* literal: um valor substituto para um valor literal, composto de exatamente uma *string* de valor (um literal que codifica o valor).

Valor *surrogate* não-literal: um valor substituto para um valor não-literal, composto por um URI de uma propriedade (um URI que identifica uma propriedade) zero ou um URI de valor (um URI que identifica o valor não-literal associado com a propriedade, zero ou um URI de esquema de codificação de vocabulário (um URI que identifica o esquema de codificação de vocabulário do qual o valor faz parte), e zero ou mais *strings* de valor (literais que representam o valor. (POWELL *et al.* 2007, *online*, tradução do proponente).

Essas definições dão pistas do significado dos conceitos relacionados a valor no modelo, mas não são autossuficientes para explicá-lo, o que o torna difícil de ser entendido. Isso acontece “porque [o DCAM] usa uma terminologia obscura e pensamentos distorcidos (e não contém exemplos e poucos textos explicativos)” (COYLE, 2008, *online*, tradução do proponente).

Com o objetivo de facilitar o entendimento e aplicação do DCAM na criação de metadados, Coyle (2008) apresentou definições mais objetivas e modelos simplificados para o uso do DCAM. A autora explica, por exemplo, o significado de valor *surrogate*:

O DCAM chama todos os valores de metadados de "*surrogate*". Na mente dos autores, o "valor" seria algo no mundo real, e o que está nos metadados é um substituto. Acontece que aceito que tudo em metadados é inventado e estou feliz em chamar um valor de metadados de um **valor**. É minha opinião que adicionar "*surrogate*" a "valor" apenas dificulta a compreensão. Você pode querer usar minha técnica de simplesmente ignorar o termo "*surrogate*" quando ele aparecer na documentação do DCAM. Eu sinceramente não acho que isso levará a qualquer erro na criação de metadados. Basta lembrar que o que há nos metadados representa algo que existe em outro lugar (COYLE, 2008, *online*, grifo da autora, tradução do proponente).

Pode-se inferir a partir da citação da autora que o qualificador *surrogate* (que pode ser traduzido como 'substituto') apenas serve para indicar que o valor em um registro de metadados é sempre um substituto para um valor que existe no objeto em si. Destarte, o termo 'substituto' não será mais empregado junto a valor neste trabalho, para evitar dificuldades de compreensão.

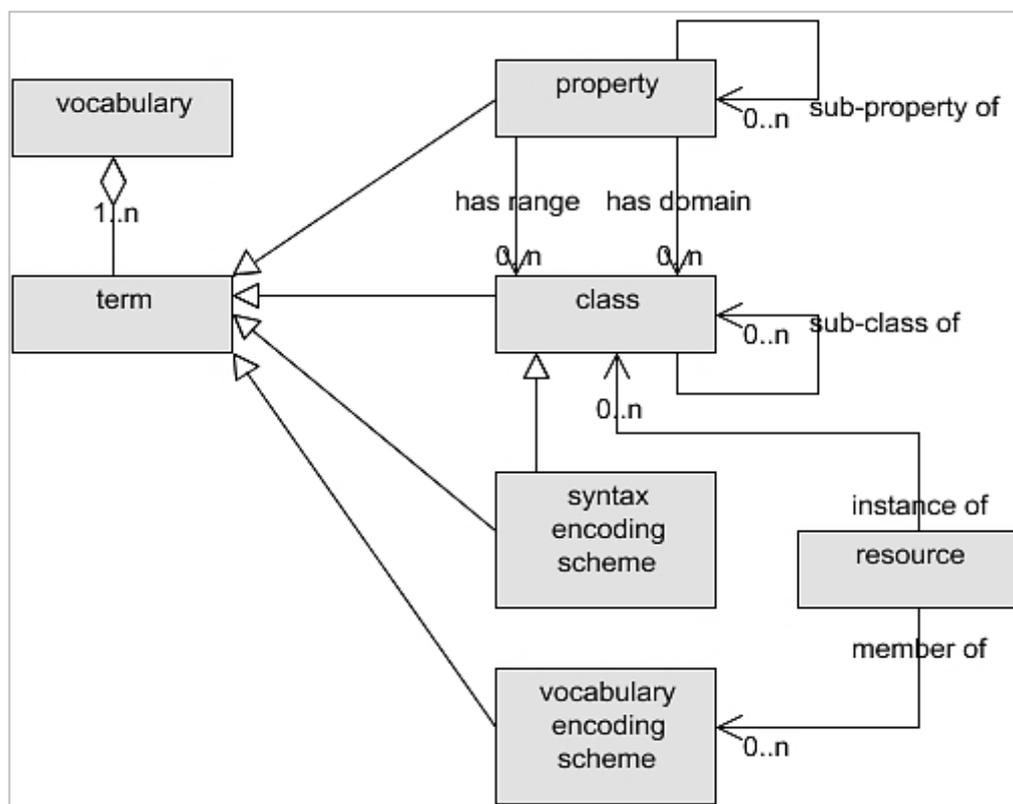
Seguindo o esquema da Figura 23, depara-se com os conceitos de valor literal substituto e valor não-literal substituto. De acordo com Coyle (2008), o valor literal é uma *string* de dados, e há dois tipos de valor literal:

string de valor simples: é basicamente o que diz: o valor do elemento de metadados será apenas uma cadeia simples de caracteres. Na programação, isso geralmente é chamado de "dados de caractere".

string de valor digitado: se você já fez alguma programação, provavelmente conhece os valores digitados. São coisas como datas, moeda, tipos numéricos (inteiro, longo). As *strings* digitadas têm a vantagem de poderem ser verificadas quanto à validade estrutural pelos programas. Portanto, se você tiver definido um elemento de metadados como um tipo de data no formato "aaaa-mm-dd", um programa poderá rejeitar quaisquer valores que não estejam em conformidade com esse formato [...] (COYLE, 2008, *online*, tradução do proponente).

O terceiro submodelo é o The DCMI Vocabulary Model da Figura 24, que apresenta a definição e os atributos do termo dentro do *abstract model*.

Figura 24 – The DCMI vocabulary model



Fonte: Powell *et al.* (2007)

Legenda: conectores que terminam em seta: leia-se *é* ou *é um*; conectores que terminam em losango: leia-se *contém* ou *tem um*.

O modelo da Figura 24 pode ser descrito da seguinte maneira:

[...] cada vocabulário é um grupo de um ou mais termos. [...] Um termo é uma propriedade (elemento), classe, esquema de codificação de vocabulário ou esquema de codificação de sintaxe. Cada propriedade pode estar relacionada a uma ou mais classes por um relacionamento de domínio. Onde é declarado que uma propriedade tem tal relação com uma classe e a propriedade é parte de um par de propriedade/valor, segue-se que o recurso descrito é uma instância dessa classe. Cada propriedade pode estar relacionada a uma ou mais classes por um relacionamento de intervalo. Onde é declarado que uma propriedade tem tal relação com uma classe e a propriedade é parte de um par propriedade/valor, é que o valor é uma instância dessa classe. Cada recurso pode ser uma instância de uma ou mais classes. Cada recurso pode ser um membro de um ou mais esquemas de codificação de vocabulário. Cada classe pode estar relacionada a uma ou mais outras classes por uma subclasse de relacionamento (em que as duas classes são definidas, ou seja, todas as instâncias que são instâncias da subclasse também são instâncias da classe relacionada). Cada propriedade pode estar relacionada a uma ou mais outras propriedades por uma subpropriedade de relacionamento. Onde é declarado que tal relacionamento existe, as duas propriedades são definidas que sempre que a subpropriedade fizer parte de um par propriedade/valor que descreve um recurso, segue-se que o recurso também é descrito usando um segundo par propriedade/valor feito acima da propriedade e do valor. Cada esquema de codificação de sintaxe é uma classe (de literais) (POWELL *et al.*, 2007, *online*, tradução do proponente).

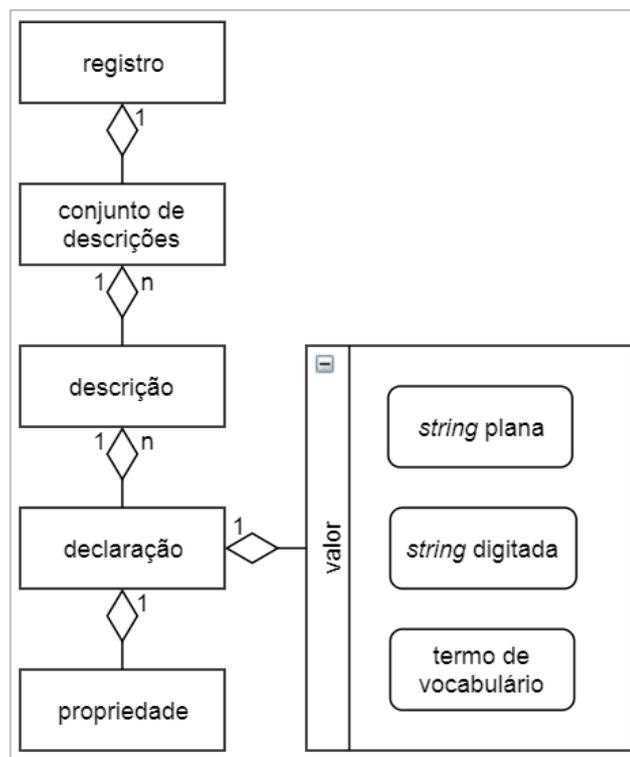
Desta forma, o DCAM determina dois tipos de valores válidos: valor literal (*string* plana ou *string* digitada) e valor não-literal (termo de vocabulário) (COYLE, 2008). Os valores do tipo não-literal são organizados em listas de termos ou vocabulários controlados (COYLE, 2008). Isso implica que esses valores podem ser considerados classes ou propriedades nos vocabulários controlados respectivos, ou seja, podem possuir estrutura hierárquica própria, exterior ao esquema de metadados ao qual foram adicionados (COYLE, 2008). Esses termos, por conseguinte, são formatados de acordo com esquemas de codificação sintáticos, que tem por função padronizar a forma como os caracteres do termo serão dispostos nos campos de metadados, e esquemas de codificação que vocabulário, que determinam quais conceitos poderão ser utilizados no preenchimento de determinados campos de metadados. De acordo com Coyle (2008) essa especificidade do DCAM é confusa, pois um valor pode ser uma propriedade em determinados casos ou outro tipo de valor.

Os três submodelos apresentados são complementares. De maneira resumida, um recurso informacional pode ser descrito por meio de pares propriedade-valor que são compostos por uma propriedade e um valor. O recurso informacional é um objeto potencialmente informativo, assim como o valor. O recurso descrito gera um registro.

O registro é composto por uma ou mais descrições; cada descrição é composta por uma ou muitas declarações; cada declaração contém uma propriedade e um valor. Os valores utilizados no preenchimento dos campos de metadados podem ser de dois tipos: *strings* ou termos de vocabulário controlado.

Para facilitar o entendimento da aplicação do DCAM na criação dos metadados para a agrobiodiversidade, optou-se por adotar o modelo simplificado elaborado por Coyle (2008), conforme Figura 25.

Figura 25 – Dublin Core Abstract Model simplificado



Fonte: Coyle (2008, tradução do proponente)

Pomerantz (2015, p. 145-146, tradução do proponente) destaca que “O DCMI Abstract Model, como modelo genérico, não especifica nenhuma sintaxe ou semântica específica para codificar as entidades e os relacionamentos em um esquema de metadados”. Em outras palavras, o DCAM não apresenta regras muito claras para a criação de metadados e dos relacionamentos entre eles. Essa limitação se aplica também, além do DC, ao DwC, conforme apontado por Baskauf e Webb (2016) (ver Figura 28). Entretanto, de acordo com Pomerantz (2015), a sintática e a semântica dos metadados podem ser codificadas e estruturadas em XML.

Em síntese, o DCAM é uma representação conceitual que descreve os componentes

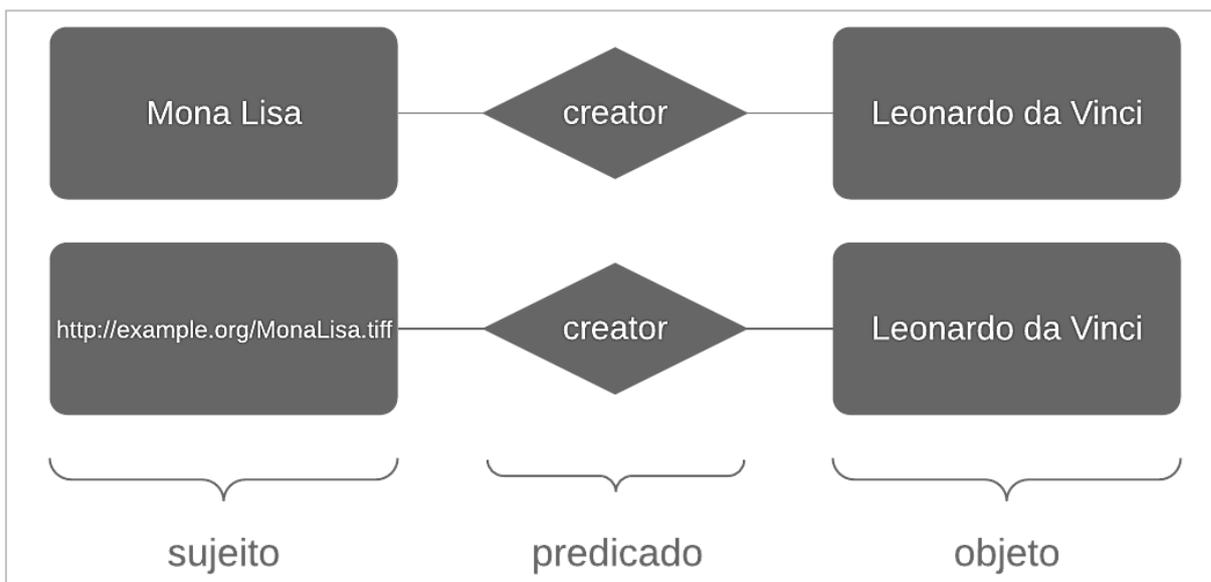
da tripla sujeito-predicado-objeto, definindo como os metadados devem ser criados atendendo a este princípio. A primeira parte do modelo, esboçada na Figura 21, descreve o elemento 'sujeito' da relação de triplas. Basicamente, demonstra que um recurso informacional é descrito por meio de um ou vários pares propriedade-valor.

A segunda parte do modelo, representada pela Figura 23, estipula os conceitos relacionados às declarações (predicados) que se pode fazer sobre o sujeito e como fazê-las formalmente, utilizando URIs para identificar as declarações.

A terceira parte do modelo, exibida na Figura 24, descreve o objeto (valor), que está ligado aos esquemas de codificação que padronizam a entrada de dados em registros de metadados.

O DCAM é base do RDF, que é um *framework*, um modelo de dados, para descrever recursos. O modelo de triplas, ilustrado na Figura 16, é o cerne do RDF, que determina que cada sujeito, predicado e objeto do modelo de triplas possuam um URI, de forma que possam ser identificados de forma inequívoca no próprio modelo de triplas e em sistemas de recuperação da informação (POMERANTZ, 2015; RDF WORKING GROUP, 2014). Pomerantz (2015) utiliza novamente a obra Mona Lisa como exemplo, na Figura 26, para explicar a relação de triplas no RDF.

Figura 26 – Relação de triplas no RDF



Fonte: Pomerantz (2015, p. 141, design e tradução do proponente)

No exemplo da Figura 26 o sujeito foi representado por um URI, seguindo os princípios

do RDF. O valor, Leonardo da Vinci pode ter um identificador único também, segundo Pomerantz (2015), que poderia ser representado em uma relação de tripla também, uma vez que, segundo o autor, todo recurso que possui uma URI pode ser representado no modelo de triplas.

DCAM, conforme descrito por Pomerantz (2015), representa a composição das triplas (sujeito-objeto-predicado). Logo, a criação de metadados deve atender a esses princípios de construção.

2.5 Padrões de metadados da biodiversidade

Diversos outros esquemas de metadados foram desenvolvidos para representar a informação em diferentes áreas de conhecimento tendo como base o DCAM. Nas Ciências biológicas, o uso de metadados é comum em muitas disciplinas, como a Zoologia Sistemática. A descrição de novas espécies deve estar de acordo com parâmetros dos códigos internacionais de nomenclatura e a representação desses registros deve ser feita de acordo com um conjunto de metadados.

Como exemplo, a Global Biodiversity Information Facility (GBIF) dá suporte a grupos de trabalho que desenvolvem e atualizam constantemente padrões de metadados devotados à representação de registros de ocorrência de espécies, coleções de espécimes e dados de biogeografia. Os padrões de metadados adotados pela GBIF são: o padrão DwC, o mais utilizado para o compartilhamento de dados no portal GBIF.org, por oferecer “uma estrutura estável, direta e flexível para compilar dados de biodiversidade a partir de fontes variadas e variáveis” (GBIF, 2019c, *online*, tradução do proponente); o EML, que é um instrumento para “formalizar e padronizar o grupo de conceitos essenciais para descrever dados ecológicos” (FEGRAUS *et al.*, 2005, p. 160, tradução do proponente); o ABCD, que “conecta dados de coleções biológicas de museus de história natural, de jardins botânicos/zoológicos e de instituições de pesquisa” (GBIF, 2013, *online*); e o Plinian Core, “um padrão orientado para compartilhar informações em nível de espécie” (PLINIAN CORE TASK GROUP, 2018, *online*).

A Seção 2.6 destaca a estrutura e composição do padrão de metadados DwC, que foi estudado neste projeto de pesquisa. Além das características desse padrão e definições, são oferecidos exemplos de uso do DwC em situações reais.

2.6 Padrão de metadados Darwin Core

A descrição normalizada de registros de espécimes permite que sistemas de informação sejam capazes de rastrear dados de outros sistemas que armazenam o mesmo tipo de dados. O padrão de metadados DwC foi concebido para facilitar a “descoberta, recuperação e integração de informação sobre espécimes biológicos modernos, suas ocorrências espaço-temporais, e suas evidências de comprovação hospedadas em coleções (físicas ou digitais)”. (DARWIN CORE TASK GROUP, 2009c, *online*, tradução do proponente).

O padrão foi criado em 2009 pelo Darwin Core Task Group, subdivisão do Biodiversity Information Standards (TDWG¹⁸), e tem sido aceito e aplicado por mais de 340 instituições, em 43 países, incluindo o Brasil (WIECZOREK *et al.*, 2012). Em 2012, estimava-se que a GBIF indexasse mais de 300 milhões de registros no formato DwC (WIECZOREK *et al.*, 2012). Hoje, o número de registros de ocorrência é de aproximadamente 1.302.735.625¹⁹, sendo desse montante 13.834.599²⁰ registros provenientes do Brasil.

A última versão do DwC, publicada em 2 de junho de 2015, foi atualizada em 26 de agosto de 2018²¹. Estes metadados estão divididos em duas categorias, *Record-level Terms* e *Auxiliary Terms*, sendo que a primeira agrupa metadados administrativos do registro de metadados e foram importados do DC, e a segunda possui campos para descrição do espécime, abrangendo dados taxonômicos, de geolocalização e outros, permitindo incluir também “informações da sequência genética associada com a ocorrência” (SiBBr, 2018, *online*). O DwC engloba diversos núcleos, que são classes de metadados que podem ser combinadas nos registros. São eles: *Record-level*, *Occurrence*, *Organism*, *MaterialSample*, *Event*, *Location*, *GeologicalContext*, *Identification*, *Taxon*, *MeasurementOrFact*, *ResourceRelationship*, *UseWithIRI*, *LivingSpecimen*, *PreservedSpecimen*, *FossilSpecimen*, *HumanObservation* e

¹⁸ O nome do grupo foi alterado de Taxonomic Databases Working Group para Biodiversity Information Standards, mas decidiram manter a sigla TDWG.

¹⁹ Dado obtido a partir da busca simples no repositório da GBIF. Disponível em: <https://www.gbif.org/occurrence/charts>. Acesso em: 28 abr. 2019.

²⁰ Dado obtido a partir da busca simples no repositório da GBIF. Disponível em: <https://www.gbif.org/occurrence/charts?country=BR>. Acesso em: 28 abr. 2019.

²¹ Disponível em: <<http://rs.tdwg.org/dwc/terms/index.htm>>. Acesso em: 9 ago. 2018.

MachineObservation.

Há dois tipos de metadados DwC: *class* (classe) e *property* (propriedade), de acordo com o Darwin Core Task Group (2009e) – vide Quadro 4 e Quadro 5, respectivamente. A documentação sobre o DwC e os projetos em andamento encontram-se disponíveis no repositório de acesso livre GitHub²².

Quadro 4 – Metadado do tipo *class*

Occurrence		Class
Identifier	http://rs.tdwg.org/dwc/terms/Occurrence	
Definition	An existence of an Organism (sensu http://rs.tdwg.org/dwc/terms/Organism) at a particular place at a particular time.	
Comments		
Examples	A wolf pack on the shore of Kluane Lake in 1988. A virus in a plant leaf in a the New York Botanical Garden at 15:29 on 2014-10-23. A fungus in Central Park in the summer of 1929.	

Fonte: Darwin Core Task Group (2009a)

Nota-se no Quadro 4 que o atributo *comments* está vazio, uma vez que se deve associar dados para metadados do tipo *class*. No atributo *examples* do Quadro 4 (tradução do proponente), há situações que ilustram o que pode ser um registro de ocorrência, como “Um vírus numa folha de planta no Jardim Botânico de Nova York às 15:29 em 23 de outubro de 2014”.

De acordo com Brickley, Guha e McBride (2014, *online*, tradução do proponente)

Recursos de informação podem ser divididos em grupos denominados classes. Os componentes de uma classe são conhecidos como instâncias da classe. Classes são recursos de informação. São frequentemente identificadas por IRIs e podem ser descritas utilizando-se propriedades RDF. A propriedade *rdf:type* por ser utilizada para estabelecer que um recurso é uma instância de uma classe.

Desta forma, as classes possuem instâncias ou propriedades, que são elementos subordinados. O elemento da Quadro 5 é uma instância da classe *Occurrence*, do

²² Disponível em: <https://github.com/tdwg/dwc>. Acesso em: 30 jan. 2019.

Quadro 4.

Quadro 5 – Metadado do tipo *property*

occurrenceID		Property
Identifier	http://rs.tdwg.org/dwc/terms/occurrenceID	
Definition	An identifier for the Occurrence (as opposed to a particular digital record of the occurrence). In the absence of a persistent global unique identifier, construct one from a combination of identifiers in the record that will most closely make the occurrenceID globally unique.	
Comments	Recommended best practice is to use a persistent, globally unique identifier.	
Examples	http://arctos.database.museum/guid/MSB:Mamm:233627 , 000866d2-c177-4648-a200-ead4007051b9 , urn:catalog:UWBM:Bird:89776	

Fonte: Darwin Core Task Group (2009a)

A partir do Quadro 5 nota-se que o atributo *comments* aponta a melhor prática de uso do elemento e em *examples* é possível observar exemplos concretos de dados que podem ser inseridos neste campo.

O DwC é baseado no DCAM, assim como o DC, “exceto a classe Darwin Core *record* que é equivalente ao Dublin Core *resource*” (DARWIN CORE TASK GROUP, 2009e, *online*, tradução do proponente). Seguindo os princípios do DCAM, o DwC estabelece os seguintes requisitos (DARWIN CORE TASK GROUP, 2009e, *online*, tradução do proponente):

- a) termos do Darwin Core são classes ou propriedades;
- b) cada propriedade tem no máximo uma classe como domínio (descreve não mais que uma classe);
- c) um registro Darwin Core é composto de zero ou mais classes e uma ou mais propriedades com seus valores associados.
- d) cada valor é uma string literal.
- e) os valores das propriedades em um registro Darwin Core descrevem esse registro;
- f) um registro Darwin Core deve incluir todas as propriedades requeridas, se houver, e seus valores associados.

Quanto à criação de registros de metadados, o DwC possui uma versão simples, o Simple DwC, que consiste em um grupo pré-definido de elementos (DARWIN CORE TASK GROUP, 2009d). Segundo o Darwin Core Task Group (2009d), essa versão

simples é a mais aplicada pelos repositórios. Há um conjunto de regras que orientam a sua aplicação, listadas no Quadro 6.

Quadro 6 – Regras de uso do Padrão Darwin Core

REGRA	DEFINIÇÃO
1	Qualquer nome de termo do Darwin Core pode ser usado como um nome de campo.
2	Nenhum nome de campo pode ser repetido em um registro.
3	Não utilizar uma classe (<i>Occurrence, Organism, MaterialSample, LivingSpecimen, PreservedSpecimen, FossilSpecimen, Event, HumanObservation, MachineObservation, Location, GeologicalContext, Identification, Taxon</i>) como campo.
4	Forneça dados em quantos campos você puder.
5	Use o campo <i>dcterms: type</i> para fornecer o nome da classe do tipo Dublin Core (<i>PhysicalObject, StillImage, MovingImage, Sound, Text</i>) que o registro representa.
6	Use o campo <i>basisOfRecord</i> para fornecer o nome da classe mais específica do Darwin Core (<i>LivingSpecimen, PreservedSpecimen, FossilSpecimen, MaterialSample, HumanObservation, MachineObservation, Event, Occurrence, Taxon, Identification, Organism, Location, GeologicalContext, MeasurementOrFact, ResourceRelationship</i>) que o registro representa.
7	Preencha os campos com dados que correspondam à definição do campo.
8	Use o vocabulário controlado para os valores dos campos que os recomendam.
9	Se os dados forem considerados sensíveis, use <i>informationWithheld</i> para dizer isso.
10	Se os dados forem compartilhados com qualidade inferior a original, use <i>dataGeneralizations</i> para isso.

Fonte: Darwin Core Task Group (2009d, *online*, tradução do proponente)

As regras do Quadro 6 orientam o uso do DwC. As regras de um a quatro são estruturantes, ou seja, referem-se às melhores práticas para elaboração do registro de dados; as regras cinco e seis determinam os campos que devem sempre ser utilizados; e as regras de sete a dez indicam as melhores práticas para a inserção de conteúdo nos metadados (DARWIN CORE TASK GROUP, 2009d).

Mais especificamente, a regra 1 determina que qualquer nome de termo do DwC pode ser usado como campo. Isso significa, basicamente, que o *term name* cumpre a função de *label* também. No DC, *term name* e *label* são diferentes atributos, mas no DwC possuem a mesma função. A aplicação da regra 1 pode ser conferida no exemplo

da Figura 27, que exibe, do lado esquerdo, os termos do DwC, e, do lado direito, um fragmento de um registro de metadados de uma ocorrência da espécie *Telenomus podisi* (Ashmead, 1893), indexado no repositório da GBIF (JOHNSON; CORA, 2019). Nota-se que o registro de ocorrência (lado esquerdo da Figura 27) adota os termos do DwC como nomes de campos, como recomendado pela regra 1.

Figura 27 – Regra 1: termo do Darwin Core como nome de campo

Term	Interpreted
Individual count	1
Life stage	ADULT
Catalogue number	OSUC 6278
Occurrence ID	urn:lsid:bioc..._627876
Recorded by	[redacted]
Sex	[redacted]

Fonte: elaborado pelo proponente (2019)

Nota: termos do lado esquerdo da figura recuperados de Wieczorek *et al.* (2009a) e lado direito capturado de Johnson e Cora (2019).

A regra 2 determina que nenhum nome de campo pode ser repetido no registro, uma vez que o nome de campo de um elemento “funciona como um identificador – deve ser único no registro” (JAVLIN, 2019, *online*, tradução do proponente).

A regra 3 determina que os metadados do tipo classe não podem ser usados como campos. As classes possuem um grupo de metadados subordinados, e apenas esses metadados podem ser utilizados como campos. Retomando o exemplo da Figura 27, é possível observar que o registro de metadados aplica o termo *Occurrence* como um cabeçalho, mas não há dados diretamente associados ao termo. A classificação dos metadados em classes ou propriedades é um atributo do esquema RDF, e “ao construir esquemas que tiram proveito da estruturação em classes, os implementadores são incentivados a manter os relacionamentos de propriedade/classe definidos pelos termos sempre que possível” (DARWIN CORE TASK GROUP, 2009e, *online*, tradução do proponente). Por exemplo, sistemas que indexam registros em XML devem manter a estruturação classe/propriedade, uma vez que essa linguagem de marcação preconiza o uso desse tipo de estrutura.

A regra 4 recomenda que os pesquisadores utilizem o máximo possível de campos para criação do registro de ocorrência. Quanto mais campos forem utilizados, maior será a quantidade de dados coletados e, proporcionalmente, o conhecimento sobre as espécies.

A regra 5 apresenta um código específico que deve ser acrescentado aos termos importados do padrão de metadados DC para diferenciá-los dos metadados do DwC. Observa-se a aplicação dessa regra no registro de metadados em XML apresentado no Esquema 2 que usa os códigos *dcterms* para o DC e *dwc* para o DwC.

Esquema 2 – Regra 5: usar códigos para metadados do DC e do DwC

```
<dwr:SimpleDarwinRecord>
  <dcterms:type>PhysicalObject</dcterms:type>
  <dcterms:modified>2009-02-12T12:43:31</dcterms:modified>
  <dcterms:rightsHolder>Museum of Vertebrate Zoology</dcterms:rightsHolder>
  <dcterms:rights>Creative Commons License</dcterms:rights>
  <dwc:institutionCode>MVZ</dwc:institutionCode>
  <dwc:collectionCode>Mammals</dwc:collectionCode>
  <dwc:occurrenceID>urn:catalog:MVZ:Mammals:14523</dwc:occurrenceID>
  <dwc:basisOfRecord>PreservedSpecimen</dwc:basisOfRecord>
  <dwc:country>Argentina</dwc:country>
  <dwc:countryCode>AR</dwc:countryCode>
  <dwc:stateProvince>Neuquén</dwc:stateProvince>
  <dwc:locality>25 km al NNE de Bariloche por Ruta 40 (=237)</dwc:locality>
</dwr:SimpleDarwinRecord>
```

Fonte: Darwin Core Task Group (2009e)

A regra 6 trata sobre a especificação da natureza do registro: o elemento *basisOfRecord* aponta o tipo de registro documental elaborado: pode conter dados observacionais sobre um espécime vivo, um espécime preservado, um animal observado na natureza, entre outros. Retomando o exemplo do Esquema 2, o par propriedade-valor `<dwc:basisOfRecord>PreservedSpecimen</dwc:basisOfRecord>` ilustra o cumprimento dessa regra, e determina que o registro de metadados representa um sujeito do tipo *PreservedSpecimen*.

A regra 7 reafirma a necessidade de se prestar atenção e compreender a semântica de cada elemento, para que os campos dos registros gerados não sejam preenchidos de maneira incorreta.

A regra 8 estipula que vocabulários controlados devem ser utilizados nos campos

quando recomendados, como, por exemplo, para os metadados da classe *Location*, que recomenda utilizar o Getty Thesaurus of Geographic Names (TGN) para inserir os nomes de lugares nos registros.

A regra 9 define que registros com dados incompletos por razões de restrição de dados, como, por exemplo, dados de localização de espécies ameaçadas, tenham essa informação apontada por meio do elemento *informationWithheld*. Isso deixa claro que há outros dados sobre aquela espécie, além daqueles que estão disponíveis para acesso público no registro.

A regra 10 aponta que dados compartilhados, ou seja, importados de outras bases de dados, e que foram generalizados ou alterados de alguma forma, devem ser indicados por meio do elemento *dataGeneralizations*. Isso acontece em casos que coordenadas geográficas precisas foram alteradas ou substituídas por uma área geográfica mais ampla, por exemplo.

O DwC importa o núcleo de metadados do padrão DC para a descrição de recursos informacionais e adota o conjunto de seus atributos para a definição de cada elemento. Entretanto, há algumas diferenças entre o conjunto de atributos do DwC e do DC. O atributo *label*, por exemplo, não aparece nos registros do DwC, e o atributo *Comment* do DC foi dividido em *Comment* e *Examples* no DwC.

Destaca-se que alguns autores detectaram algumas limitações no padrão de metadados DwC:

Enquanto DwC e ABCD representam um importante avanço na padronização de dados da biodiversidade, nenhum deles é desenhado para fornecer o tipo de semântica ou modelagem do conhecimento necessários para inferência lógica robusta. Em um nível mais básico, muitas definições de termos no vocabulário do DwC tem definições amplas que podem ser interpretadas de várias maneiras, limitando seriamente a habilidade de usar estes termos no raciocínio automatizado (e.g., *dwc:Taxon* é definido como “a categoria de informação referente a nomes taxonômicos, usos de nomes de táxons ou conceitos de táxons” e o DwC termo tipo *dwctype:Taxon* é definido como “um recurso que descreve uma instância da classe *Taxon*”) (WALLS *et al.*, 2014, p. 3-4, tradução do proponente).

Infere-se, assim, que o DwC possui duas grandes limitações:

- a) não é semântico o suficiente para permitir inferência lógica por máquina;

b) possui termos com definições ambíguas.

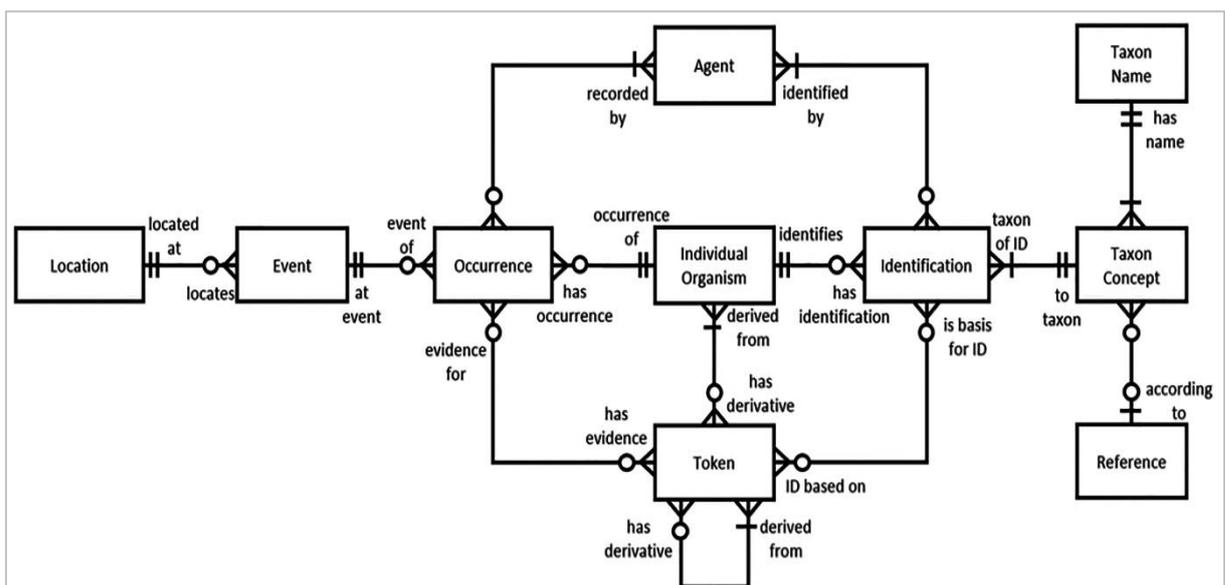
Essa ambiguidade prejudica a interoperabilidade de dados e torna a qualidade dos metadados baixa. Essas limitações do DwC requerem que seja executado o trabalho de modelagem semântica para que seja possível construir sistemas automatizados.

De acordo com Walls *et al.* (2014), esforços vêm sendo empregados tendo em vista resolver essas limitações: em 2006, o TDWG começou a desenvolver uma ontologia relacionada a Identificadores das Ciências da Vida (LSI), denominada *TDWG LSID Ontology*, mas o projeto foi encerrado. Em 2011, uma ontologia com foco mais específico no DwC, denominada *Darwin Core Semantic Web (DSW)*, foi desenvolvida com o objetivo de “facilitar a descrição das instâncias de dados da biodiversidade como RDF” (BASKAUF; WEBB, 2016, p. 630, tradução do proponente), considerando que

Embora a DSW use termos do OWL em suas definições, não é uma ontologia projetada para permitir raciocínio extensivo baseado em uma estrutura de classes hierárquica. No entanto, a estrutura da DSW e propriedades atribuídas a seus termos facilitam tarefas de raciocínio simples, mas úteis, que podem ser executadas usando consultas SPARQL (BASKAUF; WEBB, 2016, p. 632, tradução do proponente).

O modelo DSW, apesar de possuir propriedades ontológicas, não é uma ontologia (BASKAUF; WEBB, 2016). Pode ser expresso como um modelo de entidades e relacionamentos, conforme representado na Figura 28.

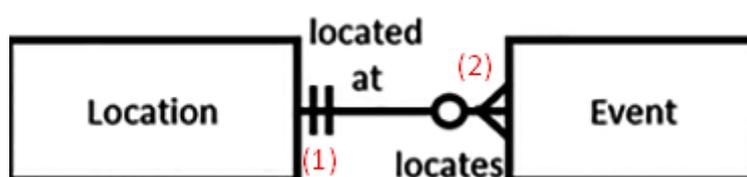
Figura 28 – Diagrama de relação de entidade do modelo de DSW



Fonte: Baskauf e Webb (2016)

A Figura 28 esquematiza como as classes do DwC poderiam ser relacionadas em um modelo de entidades e relacionamentos (ER) (BASKAUF; WEBB, 2016). Permite fazer inferências quanto aos relacionamentos que podem ser criados entre as diversas classes do DwC, mas que não podem ser expressos em um simples registro de dados, como planilhas em formato *Comma-separated values* (CSV), por exemplo. O ER da Figura 28 foi elaborado de acordo com o *Crow's foot notation*, esquema que, segundo Teorey (1999), liga entidades por meio de relacionamentos representados por linhas, com terminações de dois tipos, conforme a Figura 29:

Figura 29 – Modelo Crow's foot notation



Fonte: extrato do modelo de Baskauf e Webb (2016)

A terminação (1) da Figura 29 significa um e somente um e a terminação (2) significa zero, um ou muitos (TEOREY, 1999). Dessa forma, a partir da Figura 29 é possível interpretar que uma localização pode ser o espaço para muitos eventos, enquanto um evento acontece em uma, e apenas uma, localização.

Algumas iniciativas foram desenvolvidas com o objetivo de criar padrões para a representação do conhecimento da biodiversidade de forma a permitir inferências lógicas, ou seja, reconhecível por sistemas automáticos. Em 2010, a GBIF criou um grupo de trabalho para desenvolver Sistemas de Organização do Conhecimento (SOC) para os padrões de dados da biodiversidade (WALLS *et al.*, 2014). O trabalho desse grupo deveria “iniciar uma integração mais próxima entre os padrões do TDWG e a estrutura Open Biological and Biomedical Ontologies (OBO) *Foundry*, propondo especificamente adotar alguns princípios da OBO *Foundry*” (WALLS *et al.*, 2014, p. 4, tradução do proponente). Esse processo resultou na criação da Biological Collections Ontology (BCO), “um recurso semântico que representa as noções centrais de amostragem, coleção de espécimes e observações” (WALLS *et al.*, 2014, p. 2, tradução do proponente). A BCO, ainda segundo os mesmos autores, relaciona-se com as ontologias *Environment Ontology* (ENVO), e *Population and Community Ontology* (PCO).

Introduzidas as definições e conceitos sobre o DwC, a Seção 2.6.1 tratará sobre a criação de extensões para o padrão de metadados, que envolve os princípios e insumos teóricos para a criação de novos metadados, qualificadores e uso de esquemas de codificação.

2.6.1 Darwin Core Extensions

As extensões de metadados são, de acordo com a NISO (2004), conjuntos de elementos que não pertencem originalmente ao núcleo principal de termos do padrão de metadados, que são desenvolvidas para aplicação em domínios específicos.

A necessidade de uso de mais elementos além dos que compõem o núcleo básico do DC também ocorre com o DwC. Darwin Core Extensions são grupos de metadados criados para representar informações que não poderiam ser completamente representadas apenas pelo conjunto de termos do DwC. O TDWG permite que grupos de trabalho (*task groups*) desenvolvam extensões²³ que podem ser utilizadas de forma complementar ao DwC. Essa funcionalidade foi herdada do padrão de metadados DC.

O Dublin Core foi desenvolvido para ser o conjunto de metadados com o menor denominador comum. O problema com o menor denominador comum, no entanto, é que às vezes é muito baixo. Para alguns casos de uso, mais de 15 elementos podem ser necessários (POMERANTZ, 2015, p. 81, tradução do proponente).

No caso do DC, há três maneiras mais comuns de se estender o núcleo de metadados: a) definindo novos elementos; b) refinamento de elementos e c) utilizando-se esquemas de codificação e vocabulários controlados (BAKER *et al.*, 2005).

A extensão de metadados criada para o DC possui 40 elementos e 22 classes, contendo metadados que não existem no núcleo básico de 15 termos e qualificadores para alguns dos elementos do núcleo básico (DCMI USAGE BOARD, 2012). Dentre os novos metadados criados, destaca-se como exemplo os elementos *Replaces*, *Is Replaced By*, *Is Part Of*, *Latest Version*, *Status of Document* e *Description of Document*, expostos no Quadro 7.

Essa extensão inclui também qualificadores. Esses qualificadores tornam os

²³ As extensões aprovadas e as que estão em desenvolvimento podem ser conferidas no Darwin Core Archive Validator. Disponível em: <https://tools.gbif.org/dwca-validator/extensions.do>. Acesso em: 30 jan. 2019.

elementos do DC mais específicos. Por exemplo, o elemento *date* pode representar diferentes valores no contexto da publicação de arquivos em meio digital. Para especificar o tipo de valor representado, são aplicados os qualificadores para o elemento do núcleo principal *date* da seguinte forma: a) *Created*; b) *Valid*; c) *Available*; d) *Issued*; e) *Modified*; f) *Date Accepted*; g) *Date Copyrighted*; h) *Date Submitted* (HILLMANN, 2005). O Quadro 7 é um exemplo de uso do elemento *Date Issued*, que é uma versão qualificada do elemento *Date* e significa a data em que o recurso foi formalmente publicado (HILLMANN, 2005). Desta forma, é possível entender o significado de cada data no registro (data de criação, data de modificação, data de vigência, entre outros), o que seria impossível de se decifrar sem os qualificadores.

Quadro 7 – Exemplo de aplicação da extensão de metadados do Dublin Core

Using Dublin Core - The Elements	
Creator:	Diane Hillmann
Date Issued:	2005-11-07
Identifier:	http://dublincore.org/specifications/dublin-core/usageguide/2005-11-07/elements/
Replaces:	http://dublincore.org/specifications/dublin-core/usageguide/2005-08-15/elements/
Is Replaced By:	Not applicable
Is Part Of:	http://dublincore.org/specifications/dublin-core/usageguide/2005-11-07/
Latest Version:	http://dublincore.org/documents/usageguide/elements/
Status of Document:	DCMI Recommended Resource
Description of Document:	This document is intended as an entry point for users of Dublin Core. For non-specialists, it will assist them in creating simple descriptive records for information resources (for example, electronic documents). Specialists may find the document a useful point of reference to the documentation of Dublin Core, as it changes and grows.

Fonte: Hillmann (2005)

A terceira maneira de se estender o DC é utilizando esquemas de codificação (POMERANTZ, 2015). Esses esquemas podem ser vocabulários controlados que orientam e padronizam a entrada dos termos a serem inseridos como valores nos elementos de um registro, também denominados esquemas de codificação de

vocabulário, ou normas que definem formatos específicos de entrada de dados, denominados esquemas de codificação de sintaxe (DCMI USAGE BOARD, 2012). Os esquemas de codificação de vocabulário aplicados ao DC, de acordo com o DCMI Usage Board (2012), são:

- a) Vocabulário de Tipo DCMI (DCMIType): conjuntos de classes definidas pelo DC para descrever o tipo de recurso;
- b) Classificação Decimal de Dewey (CDD): o conjunto de elementos conceituais do sistema de classificação CDD, amplamente aplicado em bibliotecas para classificação de documentos bibliográficos;
- c) Internet Assigned Numbers Authority (IMT): termos para definir tipos de mídia;
- d) Library of Congress Classification (LCC): o conjunto de elementos conceituais do sistema de classificação LCC, amplamente aplicado em bibliotecas para classificação de documentos bibliográficos;
- e) Library of Congress Subject Headings (LCSH): o conjunto de conceitos do vocabulário controlado LCSH, amplamente utilizado em bibliotecas para a indexação de documentos bibliográficos;
- f) Medical Subject Headings (MESH): o conjunto de conceitos do vocabulário controlado MESH, amplamente utilizado para a indexação de documentos da área médica;
- g) National Library of Medicine Classification (NLM): o conjunto de elementos conceituais do sistema de classificação NLM, amplamente aplicado em bibliotecas para classificação de documentos da área médica;
- h) Getty Thesaurus of Geographic Names (TGN): vocabulário controlado de nomes de lugares;
- i) Classificação Decimal Universal (CDU): o conjunto de elementos conceituais do sistema de classificação CDU, amplamente aplicado em bibliotecas para classificação de documentos;

E os esquemas de codificação de sintaxe aplicados ao DC, ou seja, que orientam a

formatação dos valores a serem inseridos nos campos de metadados, de acordo com o DCMI Usage Board (2012), são:

- a) DCMI Box Encoding Scheme (DCMI Box): especificações dos limites de lugares no espaço geográfico por meio de coordenadas geográficas em formato textual;
- b) ISO 3166-1: lista os códigos dos nomes de países;
- c) ISO 639-2 e ISO 639-3: códigos de três letras para representação de idiomas;
- d) DCMI Period Encoding Scheme (Period): lista de intervalos de tempo;
- e) DCMI Point Encoding Scheme (Point): grupo de pontos demarcados no espaço geográfico de acordo com coordenadas geográficas;
- f) RFC1766²⁴, RFC3066, RFC4646 e RFC5646: grupo de tags para representação de idiomas, estando RFC 5646 em vigor e os demais obsoletos;
- g) Uniform Resource Identifiers (URI): especifica a sintaxe necessária para a construção de identificadores;
- h) W3C Date and Time Formats Specification (W3C-DTF): especifica o formato para entrada de valores de data e períodos;

Pomerantz (2015) ressalta que as extensões ao DC, assim como os metadados do núcleo básico, são feitas de acordo com algumas regras básicas, definidas no Dublin Core Abstract Model.

O Abstract Model é um modelo de dados para criar declarações do tipo sujeito-predicado-objeto, especificando os conceitos por trás desses sujeitos, predicados e objetos, e como esses podem ser combinados em grafos. Este modelo lógico é baseado no Resource Description Framework (RDF) (POMERANTZ, 2015, p. 83, tradução do proponente).

Logo, o modelo se aplica também ao padrão de metadados DwC, já que tem como base os mesmos princípios do DC.

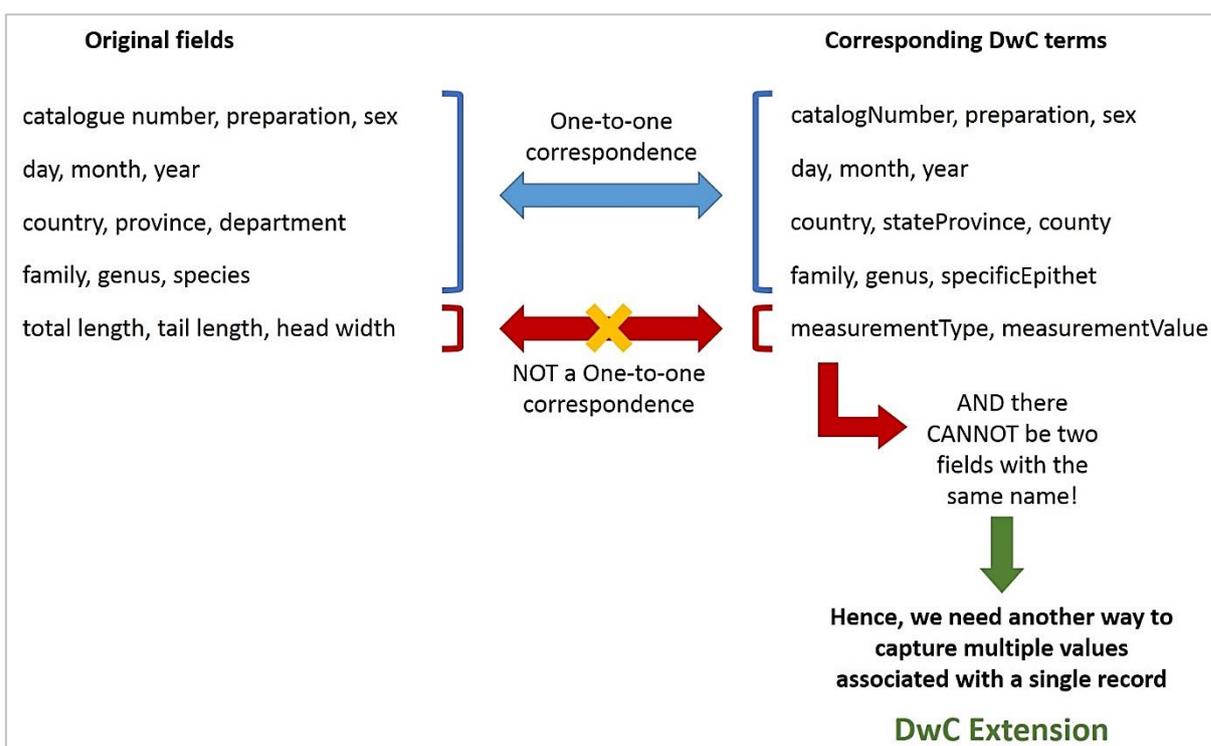
²⁴ Request for Comments (RFC).

Diversas extensões ao DwC já foram criadas e estão disponíveis para serem incorporadas a outros projetos, porém, nenhuma que atenda a agrobiodiversidade, o que, segundo Zermoglio (2018) é normal que aconteça, pois é provável que haja projetos de repositórios que não sejam atendidos pelas extensões disponíveis.

Se as extensões propostas atenderem aos requisitos estabelecidos pela TDWG, as mesmas são incorporadas ao repositório do DwC e ao Darwin Core Archive (DwC-A) Assistant²⁵, ferramenta que auxilia a criação de registros de dados no formato DwC por meio de um formulário digital para preenchimento de dados em formato CSV e a conversão automática dos arquivos para o formato XML (*metafile*).

As extensões desempenham também outra função: possibilitar a representação de valores diferentes que possuem apenas um metadado correspondente no DwC (ZERMOGLIO, 2018). Essa função evita que a regra 2 do Quadro 6 seja infringida, conforme ilustra a Figura 30.

Figura 30 – Aplicabilidade das extensões



Fonte: Zermoglio (2018)

Para explicar a Figura 30, Zermoglio (2018, *online*, tradução do proponente) utiliza o

²⁵ Disponível em: <http://tools.gbif.org/dwca-assistant/>. Acesso em: 9 ago. 2018.

seguinte exemplo:

- a) Suponha que tenhamos um registro que contenha uma medida do comprimento total de alguma criatura (comprimento total = 194 cm). Onde colocaríamos essa informação? Bem, existem alguns termos do Simple Darwin Core que atendem a esse propósito: *measurementType*, *measurementValue*, *measurementUnit*. Perfeito, então usaríamos esses três campos e os preencheríamos com o tipo de medida (“comprimento total”), o valor real (194) e as unidades (cm), respectivamente.
- b) Agora, suponha que tenhamos um registro que contenha ambas as informações sobre o comprimento total de uma criatura e o comprimento de sua cauda. No Simple Darwin Core, como não podemos usar dois campos chamados “*measurementType*”, não poderíamos inserir as duas medidas de comprimento (total e da cauda), e a mesma situação ocorreria com os valores e unidades correspondentes. Então, o que fazemos? Capturamos apenas um comprimento e esquecemos o resto? Não, claro que não! Precisamos do que chamamos de uma extensão.

Nota-se que na primeira situação, *measurementType*, *measurementValue* e *measurementUnit* são elementos que utilizam qualificadores para tornar o elemento *measurement* mais específico, para que seja possível distinguir claramente o significado de cada valor inserido. Conforme mencionado anteriormente nesta seção, esta é uma das formas de se criar extensões ao padrão de metadados DC, que se aplicam ao DwC, uma vez que este esquema de dados foi construído com base no *Dublin Core Abstract Model*, modelo que rege a criação do DC. Na segunda situação, Zermoglio (2018) afirma a necessidade da criação de um novo conjunto de elementos para representar a medida de partes de um organismo. Isso porque o DwC, assim como o DC, não permite a duplicação de campos para inserção de valores diferentes. Logo, a criação de novos elementos para a representação de informações específicas está de acordo com umas das três formas de se estender o DC, discutidas anteriormente nesta seção.

Até 31 de janeiro de 2019, estavam disponíveis, no site DwC-A Validator, as extensões descritas no Apêndice A. Essas extensões foram aprovadas e registradas pela GBIF como estáveis. “As extensões listadas como versões estáveis são imutáveis e retiradas do registro ativo da GBIF. Quaisquer alterações necessárias resultarão em uma nova versão com um novo *row type* e *namespace*, deixando o existente como estava.” (GBIF, 2013, *online*, tradução do proponente). O Apêndice B relaciona as extensões que estão em fase de desenvolvimento, que estão “sujeitas a alterações e

são principalmente destinadas a testes” (GBIF, 2013, *online*, tradução do proponente).

As extensões listadas nos Apêndices A e B são consideradas classes. Cada classe possui uma definição, campo que descreve brevemente a finalidade e o escopo dos metadados que são instâncias da classe. O campo seguinte representa o endereço eletrônico que remete ao repositório onde o esquema de metadados está armazenado ou em desenvolvimento. Logo em seguida, é possível verificar o número de propriedades (metadados) subordinadas a cada classe e o nome que designa a classe. A classe possui também um NameSpace, que “é uma coleção de nomes, identificados por um Uniform Resource Identifier (URI)” (ZENG; QIN, 2008, p. 136, tradução do proponente).

Na Seção 2.7 serão introduzidos alguns esquemas de codificação utilizados para estender o DwC. Esses esquemas são fundamentais para a padronização dos dados que podem ser atribuídos aos campos de metadados do DwC, assim como é aplicado no DC. O DwC é um padrão de metadados especializado na descrição de espécimes e por esse motivo aplica esquemas de codificação não prescritos pelo DC.

2.7 Esquemas de codificação

O padrão de metadados DwC, assim como outros esquemas de metadados, utiliza esquemas de codificação para padronizar a descrição dos valores. Esses esquemas de codificação são considerados uma forma de se estender o DwC. Para as propriedades cujos valores são nomes de táxons, o Darwin Core Maintenance Group (DwCMG, 2014) recomenda o uso dos códigos internacionais de nomenclatura para descrever esses nomes. Outros esquemas de codificação de sintaxe e de vocabulário também são recomendados nas definições dos metadados no Darwin Core Quick Reference Guide (WIECZOREK *et al.*, 2009a):

- a) ISO 8601 (2004): norma que padroniza o formato de escrita de data e horário;
- b) RFC 5646: grupo de tags para representação de idiomas, também recomendado pelo DC;
- c) Darwin Core classes: usar os nomes de classes do DwC como descritores no campo *basisOfRecord*;

- d) JavaScript Object Notation (JSON): esquema de codificação de valor recomendado para o campo *dynamicProperties*;
- e) TGN: vocabulário controlado de nomes de lugares, também recomendado pelo DC;
- f) ISO 3166-1: códigos dos países, também recomendado pelo DC;
- g) código do European Petroleum Survey Group (EPSG): sistema de coordenadas geográficas recomendado para o campo *geodeticDatum* e outros;
- h) International System of Units (SI): sistema de unidades de medida;
- i) Internationalized Resource Identifier (IRI): usar IRI para dados que representam ligações com fontes externas;
- j) Ontology of Units of Measure: esquema de unidades de medida recomendado para o campo *sampleSizeUnit*;
- k) International Code of Nomenclature for Algae, Fungi and Plants (ICN), International Code of Zoological Nomenclature (ICZN₂²⁶), International Code of Nomenclature for Cultivated Plants (ICNCP), BioCode: códigos de nomenclatura para descrição de nomes científicos (nomes de táxons) de seres vivos.

Os esquemas de codificação aplicados ao DwC podem ser adotados como forma de padronizar a descrição de valores do tipo nome de táxon nos novos metadados da extensão a ser desenvolvida para a área da agrobiodiversidade.

Os códigos internacionais de nomenclatura são sistemas complexos que padronizam a classificação e nomenclatura de espécies. No contexto da agrobiodiversidade, no qual o nome científico da espécie é utilizado como valor para representar a própria espécie em um registro de metadados, é necessário compreender como esses

²⁶ A sigla ICZN é a forma contracta tanto da International Commission on Zoological Nomenclature quanto do International Code of Zoological Nomenclature. Por essa razão, a sigla para a International Commission on Zoological Nomenclature será ICZN₁ (com o número 1 acrescido à sigla) e a sigla para o International Code of Zoological Nomenclature será ICZN₂ (com o número 2 acrescido à sigla).

códigos funcionam.

Os códigos ICN, ICZN₂ e ICNCP serão apresentados nas seções 2.6.1 a 2.6.3 porque são aplicados como esquemas de codificação de nomes de espécies que ocorrem nos sistemas agroecológicos, portanto foi necessário estudá-los nesta pesquisa.

2.7.1 *International Code of Zoological Nomenclature*

Um comitê da British Association for the Advancement of Science, que tinha entre os seus membros Charles Darwin e Richard Owen, foi o primeiro a apontar, em 1842, a necessidade de se corrigir as regras da nomenclatura zoológica existentes e estabelecer um conjunto de regras internacional (JAIN, 2018; MINELLI, 2013). Entretanto, este problema só se tornou internacional quando o “terceiro Congresso Internacional de Zoologia [(ICZ)], (em Leiden 1895) criou a Comissão Internacional de Nomenclatura Zoológica (ICZN₁), sob a presidência do zoólogo francês Raphaël Blanchard” (MINELLI, 2013, p. 2, tradução do proponente). Todas as atualizações do Código são publicadas no *Bulletin of Zoological Nomenclature* (BZN), criado pela ICZN₁ para este fim (MINELLI, 2013).

Após muitos anos de pesquisa e estudos de zoólogos em diversas partes do mundo para criação das regras, foi publicado durante o décimo quinto Congresso Internacional de Zoologia em Berlin (1901) as *Règles internationales de la Nomenclature zoologique*, em três idiomas: francês, inglês e alemão (ICZN, 1999; MINELLI, 2013). A aplicação das regras permaneceu sob responsabilidade da ICZN₁, que se tornou um corpo permanente do Congresso Internacional de Zoologia (MINELLI, 2013). As *Règles* foram reescritas e atualizadas em 1961, dando origem à primeira edição do ICZN₂; a segunda edição foi publicada não muito tempo depois, em 1964 (ICZN, 1999; JAIN, 2018). Em 1972, ocorreu a décima nona e última edição do ICZ, em Monte Carlo (ICZN, 1999; JAIN, 2018). A partir de então, a atuação da ICZN₁ tornou-se subordinada à União Internacional de Ciências Biológicas (IUBS) (MINELLI, 2013). Entretanto, o ICZ foi reestabelecido em 2000, em Atenas (MINELLI, 2013). Em fevereiro de 1985, foi publicada a terceira edição do código (ICZN, 1999). O comitê responsável pela publicação da terceira edição já havia notado que não seria possível lidar com uma gama de problemas presentes nesta edição, por isso apontaram a necessidade de se publicar uma quarta edição, com as soluções para os

problemas da edição anterior (ICNZ, 1999).

O projeto de formular a quarta edição foi formalmente iniciado pela Comissão em seu encontro em Canberra em outubro de 1988, juntamente com a XXIII Assembleia Geral da União Internacional de Ciências Biológicas (IUBS). Foi antecipado que a nova edição do Código incorporaria algumas mudanças maiores em relação às edições anteriores, embora, como essas, obviamente teria de ser compatível com as ações tomadas por zoólogos em tempos passados (ICNZ, 1999, p. 2, tradução do proponente).

O Código propriamente dito inclui o Preâmbulo, 90 artigos (agrupados em 18 capítulos) e um Glossário. Cada artigo é composto por uma ou mais disposições obrigatórias, que são por vezes acompanhadas de recomendações e / ou exemplos ilustrativos. Interpretando o Código, o significado de uma palavra ou expressão deve ser tomado como o Glossário (ver Artigo 89). As aplicações das disposições do Código podem ser dispensadas ou modificadas em caso particular quando a adesão estrita causaria confusão, mas isto só pode ser feito pela Comissão, atuando em nome de todos os zoólogos e utilizando o seu poder plenário (artigos 78 e 81), e nunca por um indivíduo (ICNZ, 1999, p. 1, tradução do proponente)

Além do próprio Código, o presente volume contém um Prefácio (pelo presente e Presidentes da Comissão) e uma Introdução (pelo Presidente da Comissão). Existem três apêndices; os dois primeiros têm o status Recomendações, e a terceira é a Constituição da Comissão (ICNZ, 1999, p. 1, tradução do proponente).

Com a propagação da Internet e das tecnologias da informação, pesquisadores em diferentes partes do mundo apontaram a necessidade de se mudar a regra do Código para o registro de novas espécies de forma que esse começasse a aceitar publicações exclusivamente em meio eletrônico (KNAPP; POLASZEK; WATSON, 2007; WHEELER, 2008; RIDE, 1999). Isso porque a quarta edição do código considerava como publicação “trabalho produzido por um método assegurando numerosas cópias físicas idênticas e duráveis, ou seja, impresso em papel usando tinta ou toner” (MINELLI, 2013, p. 3, tradução do proponente).

Em 2008, durante o vigésimo ICZ, a ICZN₁ iniciou “o processo necessário para alterar o Código Internacional de Nomenclatura Zoológica para permitir a publicação eletrônica de novos nomes científicos e outros atos nomenclaturais” (ICZN, 2008, p. 57, tradução do proponente). As alterações foram propostas em torno de três pontos principais:

Publicações somente eletrônicas devem ser permitidas, se forem encontrados mecanismos que forneçam garantia razoável da

acessibilidade a longo prazo das informações que contêm.

Algum método de registro deve fazer parte do mecanismo de permitir a publicação eletrônica de nomes e atos nomenclaturais.

Trabalhos físicos que não são baseados em papel (por exemplo, CD-ROMs, DVDs) devem ser proibidos. (ICZN, 2008, p. 57, tradução do proponente).

As alterações foram aceitas e aprovadas pela IUBS em 2009 durante o encontro da entidade em Cape Town (ICZN, 2012). Alguns pontos adicionais foram sugeridos durante a discussão, e as seguintes alterações foram aprovadas:

- a) As alterações relativas à publicação eletrônica devem ser efetivas a partir do início de 2012.
- b) A exigência de registro no ZooBank de novos nomes científicos em obras eletrônicas foi alterada para exigência de registro da própria obra.
- c) O requisito de que um trabalho eletrônico seja arquivado foi alterado para uma exigência de arquivamento, com este requisito sendo satisfeito pela declaração do arquivo pretendido no ZooBank.
- d) O requisito que um ISSN ou ISBN (International Standard Serial Number ou International Standard Book Number) seja incluído no registro ZooBank foi adicionado.
- e) O período durante o qual os discos ópticos, como o CD-ROM, foram aceitáveis, foi alterado de “depois de 2000 e antes de 2010” para “depois de 1985 e antes de 2013”. (ICZN, 2012, p. 1-2).

O Código é formado por 149 páginas contendo 90 artigos, os apêndices e o glossário oficial (ICZN, 2012). Atualmente, a quarta edição está em vigor, sendo válida desde 31 de dezembro de 1999, com emendas aos Artigos 8, 9, 10, 21 e 78, referentes à publicação eletrônica de registros de novas espécies, válidas desde 1º de janeiro de 2012 (ICZN, 2012). A quinta edição do código está em discussão²⁷, e o Código permanece sob responsabilidade da ICZN₁, subordinada à IUBS. O Código tem como objetivo

promover a estabilidade e a universalidade nos nomes científicos dos animais e assegurar que o nome de cada táxon seja único e distinto. Todas as suas disposições e recomendações são subservientes a esses fins e nenhuma restringe a liberdade de pensamento taxonômico ou ações (ICZN, 1999, p. 15, tradução do proponente).

Para cumprir tal objetivo, o código segue alguns princípios gerais, segundo Dubois

²⁷ Vide *5th edition Wiki*. Disponível em: <http://iczn.org/content/5th-edition-wiki>. Acesso em: 7 ago. 2018.

(2011):

- a) princípio da independência nomenclatural: diferentes paradigmas nomenclaturais podem ser utilizados como base para a construção de uma taxonomia, como o paradigma filogenético;
- b) princípio da fundação nomenclatural: o status nomenclatural²⁸ é fixado pela publicação original na qual o nome foi criado e registrado pela primeira vez. Isso significa que este nome não está mais disponível para ser atribuído a outras espécies e nenhuma ação pode mudar o status nomenclatural deste nome. Entretanto, em situações onde há ambiguidade terminológica é necessário que seja aplicado um princípio complementar, denominado princípio do primeiro revisor;
- c) princípio do primeiro revisor: aplica-se a situações onde há sinônimos e homônimos que não podem ser revolidos levando em consideração apenas o princípio de prioridade. Acontece com nomes com grafias diferentes atribuídos simultaneamente a uma mesma espécie, por exemplo. Neste caso, o primeiro autor a revisar o caso e a publicar uma solução aceitável de acordo com as regras do Código será considerado o primeiro revisor.

Além dos princípios gerais, os três estágios do código se orientam por princípios específicos. O primeiro estágio, denominado Disponibilidade de *nomen*, orienta-se pelos princípios de Série-nominal, Binomina, Coordenação e Neonímia; o segundo estágio, denominado Alocação de *nomen*, é regido pelo Princípio de Onomatóforos ou espécie-tipo; o terceiro estágio, denominado validade e extensão de *nomen*, é orientado pelo princípio da validade (DUBOIS, 2011).

2.7.2 International Code of Nomenclature for algae, fungi, and plants

Os estudos da nomenclatura botânica tiveram início no Século XVIII, notadamente em 1753, com a publicação da primeira edição de *Species Plantarum*, por Carl Linnaeus (KARTHICK; WILLIAMS, 2012). Em 1867, Alphonse de Candolle publicou um trabalho

²⁸ “[O status nomenclatural] De um nome, ato de nomenclatura ou obra: a sua nomenclatura (ou seja, a sua disponibilidade ou não, e no caso de um nome a sua ortografia, a tipificação que o *taxon* nominal denota, e sua precedência relativa a outros nomes)” (ICZN, 1999, p. 125, tradução do proponente).

intitulado *Lois de la nomenclature Botanique* (CANDOLLE, 1867) no Congresso Internacional de Botânica, em Paris, que consistiu em um código de nomenclatura botânica com 60 páginas (NICOLSON, 1991). De acordo com o mesmo autor, durante o congresso, uma comissão avaliou o código de Candolle e o aprovou, após sugerir algumas modificações. Foi considerado o primeiro código internacional de nomenclatura botânica, também conhecido como Código de Paris (NICOLSON, 1991; RODRIGUES, 2003). Tendo recebido inúmeras críticas ao código proposto, Candolle publicou o trabalho *Novas observações* em 1883 para corrigir alguns erros apontados por outros pesquisadores (NICOLSON, 1991; RODRIGUES, 2003).

Em 1892, durante o Meeting of Botanical Club, subordinado à American Association for the Advancement of Science (AAAS), em Rochester, NY, as *Lois* foram aceitas, mas modificadas em oito pontos, recebendo a designação de *Rochester resolutions* ou *Rochester Code* (NICOLSON, 1991; VAN RIJCKEVORSEL, 2014b).

A primeira edição do International Code of Botanical Nomenclature (ICBN), com o nome de Règles internationales de la Nomenclature botanique adoptées par le Congrès International de Botanique de Vienne 1905, foi aceita e publicada em 1905, no Congresso de Botânica de Viena, tendo sido desenvolvida em 1904 pelo Botanical Club of the American Association for the Advancement of Science (KARTHICK; WILLIAMS, 2012).

Na edição do Congresso em Estocolmo em 1950 foi fundada a International Association for Plant Taxonomy (IAPT), organização responsável por editar e dar suporte à aplicação do Código a partir de 1950. O Congresso resultou na publicação do Código de Estocolmo, sendo a primeira vez que o termo 'código' foi empregado no título da publicação (VAN RIJCKEVORSEL, 2014a).

As maiores e mais recentes mudanças no Código resultaram do Congresso de Melbourne, em 2011. O nome do código foi mudado de International Code of Botanical Nomenclature (ICBN) para International Code of Nomenclature for Algae, Fungi and Plants (ICN) e as publicações de registro de novas espécies apenas em meio eletrônico passaram a ser aceitas como válidas (KARTHICK; WILLIAMS, 2012; TURLAND *et al.*, 2018; VAN RIJCKEVORSEL, 2014a), desde que sejam publicadas em formato “[Portable Document Format] (PDF) com arquivamento (ISO 19005), ou

como um livro com o [Número Internacional de Série Padrão] (ISSN) ou um [International Standard Serial Number] (ISBN)” (KARTHICK; WILLIAMS, 2012, p. 551).

A sessão de Shenzhen foi a mais recente do congresso, tendo ocorrido em 2017. Cabe-se destacar uma grande mudança no código após o congresso:

A mudança mais significativa para o Código de Shenzhen é, sem dúvida, a criação de um capítulo especial no Código que reunirá todos os artigos que são específicos para fungos. Esta regra aparecerá como um Artigo, na próxima edição do Código (PRADO *et al.*, 2017, p. 1501).

O ICN é um conjunto de regras que governa o registro de nomes científicos

de todos os organismos tradicionalmente tratados como algas, fungos ou plantas, fósseis ou não fósseis, incluindo algas verde azuladas (*Cyanobacteria*), quitrídios, oomicetos, fungos e protistas fotossintéticos com seus grupos não fotossintéticos taxonomicamente relacionados (mas excluindo Microsporídia) (TURLAND *et al.*, 2018, *online*, tradução do proponente).

A versão mais recente do código é composta por três divisões, acrescidas de um glossário, um índice e apêndices (TURLAND *et al.*, 2018). Ainda de acordo com os mesmos autores, as três divisões principais englobam princípios, regras e recomendações e provisões para governança para nomenclatura de plantas, algas e fungos, com exceção das plantas cultivadas, que são nomeadas conforme código próprio (ver Seção 2.7.3). Segundo Turland *et al.* (2018, *online*, tradução do proponente):

A biologia exige um sistema preciso e simples de nomenclatura que seja usado em todos os países, lidando, por um lado, com os termos que denotam os ranques de grupos ou unidades taxonômicas e, por outro, com os nomes científicos que são aplicados a grupos taxonômicos individuais. O objetivo de dar um nome a um grupo taxonômico não é indicar seu caráter ou história, mas fornecer um meio de se referir a ele e indicar sua classificação taxonômica. Este Código visa a provisão de um método estável de nomeação de grupos taxonômicos, evitando e rejeitando o uso de nomes que possam causar erro ou ambiguidade ou lançar a confusão sobre a ciência. A próxima importância é evitar a criação inútil de nomes. Outras considerações, tais como correção gramatical absoluta, regularidade ou eufonia de nomes, costume mais ou menos prevalente, consideração por pessoas etc., apesar de sua importância inegável, são relativamente acessórias.

E, para cumprir tal objetivo, o ICN adota alguns princípios básicos, que são:

Princípio I: a nomenclatura de algas, fungos e plantas é independente da nomenclatura zoológica e procariótica. Este Código aplica-se igualmente a nomes de grupos taxonômicos tratados como algas, fungos ou plantas, independentemente de esses grupos terem ou não sido originalmente tratados como tal.

Princípio II: a aplicação de nomes de grupos taxonômicos é determinada por meio de tipos nomenclaturais²⁹.

Princípio III: a nomenclatura de um grupo taxonômico é baseada na prioridade de publicação.

Princípio IV: cada grupo taxonômico com uma circunscrição, posição ou ranque particulares pode suportar apenas um nome correto, o mais antigo de acordo com as regras, exceto em casos específicos.

Princípio V: nomes científicos de grupos taxonômicos são tratados como Latim, independentemente de suas derivações.

Princípio VI: as regras de nomenclatura são retroativas, a menos que sejam expressamente limitadas (TURLAND *et al.*, 2018, *online*, tradução do proponente).

2.7.3 International Code of Nomenclature for Cultivated Plants

Assim como os códigos de nomenclatura apresentados anteriormente nesta seção, o International Code of Nomenclature for Cultivated Plants (ICNCP) tem como precursor o sistema de classificação binominal desenvolvido por Linnaeus, porém plantas cultivadas ficaram subordinadas ao mesmo tipo de nomenclatura aplicada às plantas selvagens (HETTERSCHEID; VAN DEN BERG; BRANDENBURG, 1996).

Alphonse De Candolle, cujo trabalho foi fundamental para o desenvolvimento do ICN, apresentou em 1864 no International Horticultural Congress, Bruxelas, sugestões para a nomenclatura de plantas cultivadas (JIRÁSEK, 1961). Entretanto, as plantas cultiváveis foram superficialmente mencionadas nas Lois de la Nomenclature Botanique publicadas por De Candolle em 1867 (HETTERSCHEID; VAN DEN BERG; BRANDENBURG, 1996).

De Candolle sugeriu que, para as variedades hortícolas (tipos) deveria deve ser obrigatório usar exclusivamente os nomes não latinos, "fancy", isto é, os chamados nomes vernáculos ou da moda derivados de uma linguagem moderna - agora também nomes comerciais, nomes agradáveis com uma forma comercialmente atraente - de modo que os nomes das plantas cultivadas diferem acentuadamente dos

²⁹ "O elemento ao qual o nome de um táxon está permanentemente vinculado" (TURLAND *et al.*, 2018, *online*, tradução do proponente).

nomes científicos latinos de espécies e variedades incluídas na botânica sistemática, evitando assim erros, confusões e indistinções (JIRÁSEK, 1961, p. 34, tradução do proponente).

Em 1866, Alefeld propôs regras mais específicas para a classificação de plantas cultivadas na obra *Landwirtschaftliche Flora*, que seria a base para a primeira edição do ICNCP em 1952 (HETTERSCHEID; VAN DEN BERG; BRANDENBURG, 1996). Diferente do que foi proposto por Candolle (1867), as regras de Alefeld propuseram uma designação consistente para as plantas cultiváveis com descrições formais em Latim (HETTERSCHEID; VAN DEN BERG; BRANDENBURG, 1996).

Harlan & De Wet (1971) promoveram uma classificação 'racional' de plantas cultivadas usando uma perspectiva genética / biosistemática. Sua base é a espécie biológica, e o conceito de cruzabilidade é decisivo em sua divisão do pool genético em primário, secundário e terciário, dependendo do grau de cruzabilidade entre plantas cultivadas individuais (HARLAN; DE WET, 1971 citado por HETTERSCHEID; VAN DEN BERG; BRANDENBURG, 1996, p. 129, tradução do proponente).

Antes disso, em 1923 foi fundada a International Commission for Horticultural Congresses, que serviu como base para o surgimento da International Society for Horticultural Science (ISHS) em 1959, organização responsável por publicar as atualizações do código em uma publicação denominada *Acta Horticulturae* (HETTERSCHEID; VAN DEN BERG; BRANDENBURG, 1996). Atualmente, o código encontra-se na nona edição, publicada em 2016.

As disciplinas de agricultura, silvicultura e horticultura requerem um sistema preciso, estável e simples de nomenclatura de plantas que pode ser aplicado internacionalmente. Este sistema deve lidar com ambos os termos que designam as várias categorias de grupos de plantas taxonomicamente distintos (taxa, plural; táxon, singular) no cultivo e formação de nomes de tais grupos (BRICKELL *et al.*, 2009, p. 1, tradução do proponente).

Desta forma, o objetivo do código é fornecer “um método estável de nomear taxa de plantas cultiváveis³⁰, evitando e rejeitando o uso de nomes que podem causar erro ou ambiguidade” (BRICKELL *et al.*, 2009, p. 1, tradução do proponente).

³⁰ “Essas são deliberadamente plantas selecionadas que podem ter surgido por hibridização acidental ou intencional no cultivo, por seleção de estoques de cultivo existentes ou de variações dentro de populações selvagens que são mantidas como entidades reconhecíveis unicamente pela propagação contínua. Plantas deste tipo são referidas algumas vezes como cultigens” (BRICKELL *et al.* 2009, p. 1, tradução do proponente).

2.8 Considerações sobre os esquemas de codificação

De maneira geral, os códigos internacionais de nomenclatura apresentam os princípios para se nomear os táxons e as regras de nomenclatura (o formato em que devem ser escritos). Para se entender a complexidade desses códigos, é necessário estudo mais aprofundado. Entretanto, como não faz parte do escopo desta pesquisa descrever espécies, mas fazer uso de descrições prontas, decidiu-se não explorar de maneira profunda os princípios de cada código. A Figura 31 demonstra um exemplo de nomes científicos formatados conforme os códigos de nomenclatura.

Figura 31 – Exemplo de aplicação de um código internacional de nomenclatura

Nome científico aceito:	<i>Zea mays</i> L. (Nome aceito)		
Sinônimos:	<i>Mays americana</i> Baumg., nom. superfl. (sinônimo)		
	<i>Mays zea</i> Gaertn., nom. superfl. (sinônimo)		
	<i>Mayzea cerealis</i> Raf., nom. superfl. (sinônimo)		
	<i>Thalysia mays</i> (L.) Kuntze (sinônimo)		
	<i>Zea segetalis</i> Salisb., nom. superfl. (sinônimo)		
Táxon infraespecífico:	<i>Zea mays</i> subsp. <i>huehuetenangensis</i> (Iltis & Doebley) Doebley		
	<i>Zea mays</i> subsp. <i>mays</i>		
	<i>Zea mays</i> subsp. <i>parviglumis</i> Iltis & Doebley		
Nomes comuns:	-		
Classificação taxonômica:	Reino	Plantae	CoL
	Filo	Tracheophyta	CoL
	Classe	Liliopsida	CoL
	Ordem	Poales	CoL
	Família	Poaceae	WCSP
	Gênero	<i>Zea</i>	WCSP

Fonte: Govaerts (2019)

No topo da Figura 31 observa-se o nome científico aceito para uma espécie de milho de acordo com as regras do ICN, que é o nome correto a ser citado. Logo abaixo, estão relacionados os nomes que foram sinonimizados após se descobrir que representavam a mesma espécie ou por outros motivos. Os três códigos de nomenclatura das Seções 2.6.1, 2.6.2 e 2.6.3 apresentam os princípios para determinação do nome válido e para sinonimizarão.

3 METODOLOGIA

Considera-se que a metodologia desta pesquisa é, quanto aos procedimentos, de natureza aplicada (GIL, 1994; GERHARDT; SILVEIRA, 2009), uma vez que “objetiva gerar conhecimentos para aplicação prática dirigidos à solução de problemas específicos” (SILVA; MENEZES, 2001, p. 20). Quanto à abordagem do problema, esta investigação caracteriza-se como qualitativa, devido à maneira intuitiva de analisar os dados, com o suporte de especialista do domínio (GIL, 1994). Do ponto de vista de seus objetivos, a pesquisa é exploratória e descritiva, uma vez que envolveu o estudo exaustivo das temáticas tratadas (GIL, 1994).

O universo desta pesquisa são os metadados para a representação de informações sobre a agrobiodiversidade e o objeto estudado é o conjunto de princípios que fundamentam a construção de uma extensão de metadados sobre essa temática, para o padrão de metadados DwC.

Os procedimentos metodológicos foram subdivididos em duas etapas, conforme as Seções 3.1 e 3.2. As etapas foram delimitadas para facilitar a compreensão do leitor, mas a execução dos procedimentos não aconteceu de forma linear. A pesquisa aplicada exige, em diversos momentos, que experimentos sejam refeitos e que novas fontes de literatura sejam buscadas de forma pontual e direcionada, o que não implica em revisão sistemática da literatura. À vista disso, a pesquisa bibliográfica e os outros meios adotados para encontrar a literatura demandada encontram-se no Apêndice C.

3.1 Primeira etapa: exploração dos insumos metodológicos e terminológicos

A primeira etapa desta pesquisa teve como objetivo a imersão no campo temático da biologia com foco na representação da informação. A análise que foi organizada em cinco subetapas:

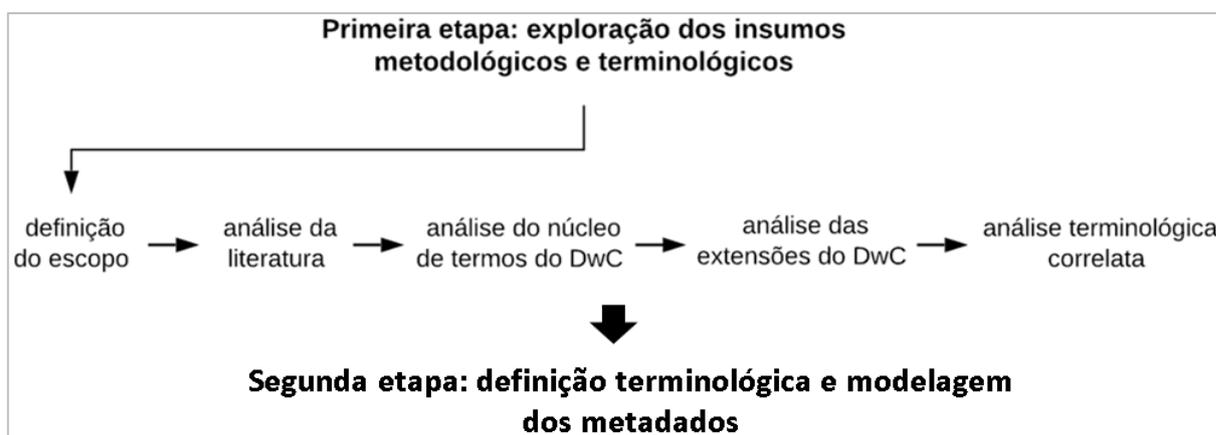
- a) definição do escopo de representação de dados da agrobiodiversidade, utilizando como insumo o *Final Report of the Task Group on GBIF Data Fitness for Use in Agrobiodiversity* (ARNAUD *et al.*, 2016), para delimitação da amostra a ser trabalhada, conforme problema e justificativas estabelecidos na introdução, em especial na Subseção 1.1;
- b) análise da literatura para coleta de informações sobre a temática da amostra,

dentro do escopo definido na primeira subetapa, tendo em vista fazer uma sistematização e gerar uma representação conceitual da temática que será trabalhada, que servirá de base para a correlação realizada na quinta e última subetapa;

- c) análise do núcleo principal de termos do DwC;
- d) análise das extensões do padrão de metadados DwC;
- e) análise terminológica correlata das classes do modelo <onto.biodiversidade>, DwC e conceitos sobre a temática da amostra, dentro do escopo definido na primeira subetapa e com os insumos obtidos na segunda subetapa.

Essas subetapas têm como função a exploração individual dos insumos metodológicos, conforme esquematizado na Figura 32.

Figura 32 – Estrutura da metodologia



Fonte: elaborado pelo proponente (2019)

Quanto à função dentro da metodologia, os insumos utilizados no desenvolvimento da pesquisa podem ser divididos em duas categorias:

- a) insumo de análise terminológica: é formado pelo núcleo principal de termos do DwC e pelas extensões de metadados do DwC, que apresentam a estrutura a ser adotada para a criação da extensão de metadados para a agrobiodiversidade; os conjuntos de termos desses insumos foram analisados a fim de se detectar termos potenciais para a nova extensão de metadados;
- b) insumo de extração terminológica: é formado pelo modelo

<onto.biodiversidade> e pelos conceitos encontrados na literatura especializada que descrevem informações sobre a temática da amostra, em conformidade com o exposto na introdução, em especial na Subseção 1.1. A seleção do modelo <onto.biodiversidade> se justificou pelo fato de ele ter sido desenvolvido como suporte a um projeto da Embrapa, instituição referência na pesquisa agropecuária brasileira, e por não haver outros modelos semelhantes na literatura.

3.2 Segunda etapa: definição terminológica e modelagem dos metadados

Os metadados são representados por predicados, que são termos e possuem significados em um dado contexto. Só é possível entender a função dos metadados por meio de definições e termos claros e bem delineados e, em alguns casos, exemplos de aplicação.

Esta segunda etapa consiste em apresentar princípios que permitam definir, de maneira clara e objetiva, a definição dos metadados, aplicando as recomendações da ISO/IEC 11179-4 (2004) para construção de metadados, apresentadas na Seção 2.4, e também a sintática e semântica recomendadas no DwC, conforme a Seção 2.6, para a definição dos metadados, conforme será apresentado e discutido nos resultados.

A essa atividade de definir os metadados, atribui-se o nome 'definição terminológica funcional', uma vez que esta atividade consiste em estabelecer a função dos metadados, apontando suas regras de aplicação, e não apenas o significado do termo que representa o predicado.

4 RESULTADOS E ANÁLISES

Nesta seção estão relacionados os resultados da aplicação dos procedimentos em cada etapa da metodologia, com as respectivas análises. As subseções correspondem à mesma divisão da metodologia.

4.1 Primeira etapa: exploração dos insumos metodológicos e terminológicos

Para tomar conhecimento sobre a biodiversidade e sobre as interações ecológicas, o proponente deste trabalho realizou algumas experiências de imersão no campo científico: visitas ao Laboratório de Aracnologia, do Instituto de Ciências Biológicas da UFMG, para observação dos procedimentos de descrição e inventário de espécimes; matrícula na disciplina Tópicos em Zoologia II, do Programa de Pós-Graduação em Zoologia da UFMG (PPGZoo-UFMG); elaboração da disciplina Representação da Informação sobre Biodiversidade, ministrada para os cursos de Biblioteconomia e Ciências Biológicas da UFMG; participação no Biodiversity Next, em Leiden, na Holanda, conferência conjunta da GBIF, Distributed System of Scientific Collections (DiSSCo), National Resource for Digitized Information about Vouchered Natural History Collections (iDigBio), Consortium of European Taxonomic Facilities (CETAF), Biodiversity Information Standards (TDWG) e e-Science and Technology European Infrastructure for Biodiversity and Ecosystem Research (LifeWatch ERIC).

Esta primeira etapa foi subdividida em cinco subetapas, conforme apresentado a seguir.

4.1.1 Definição do escopo de representação de dados da agrobiodiversidade

Para definir o escopo de representação de dados da agrobiodiversidade, foi utilizado como insumo o *Final Report of the Task Group on GBIF Data Fitness for Use in Agrobiodiversity* (ARNAUD *et al.*, 2016), que estabelece as três principais categorias de dados produzidos no contexto da agrobiodiversidade, apontando como representar alguns desses dados por meio de metadados já desenvolvidos pela GBIF. Destacam-se como dados importantes para aplicação em sistemas de cultivo dados sobre interações ecológicas, aspectos agrônômicos e práticas agrônômicas (ARNAUD *et al.*, 2016).

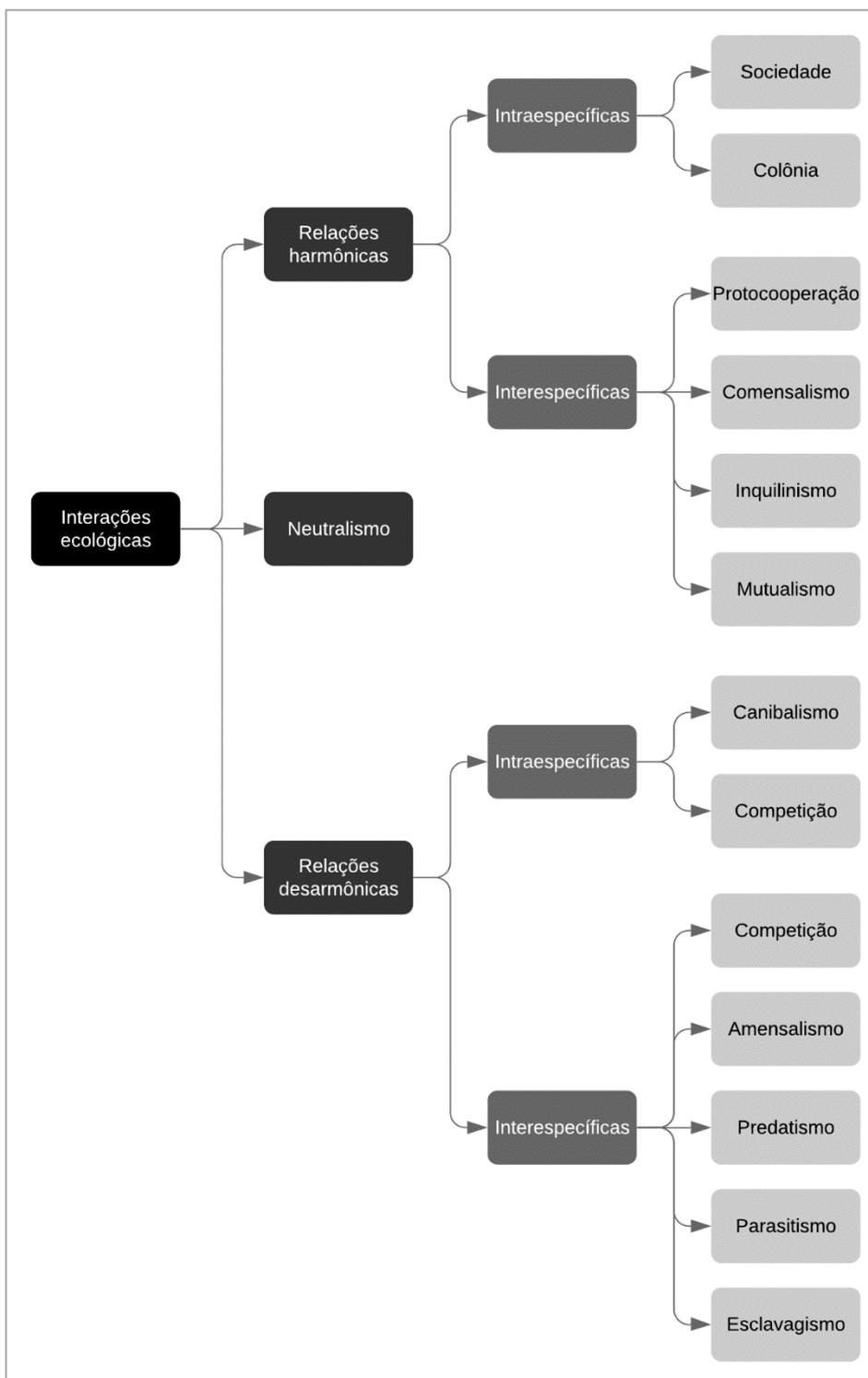
Decidiu-se adotar como amostra desta pesquisa os conceitos da agrobiodiversidade

que compõem o nicho das interações ecológicas, uma vez que elas não estão representadas nos metadados do DwC e nem em suas extensões já publicadas. A justificativa para tal escolha deve-se ao fato de as interações ecológicas estarem mais relacionadas à biodiversidade do que os demais temas da agrobiodiversidade apresentados por Arnaud *et al.* (2016), que são aspectos agronômicos e práticas agronômicas. Dentre as interações ecológicas estudadas, foram extraídos os conceitos para criação de metadados da classe parasitismo, que foi adotada como exemplo de aplicação dos princípios para criação de metadados, visto que não seria possível criar metadados para todas as classes de interações ecológicas no prazo do mestrado.

4.1.2 Análise sistemática da literatura sobre interações ecológicas

A Seção 2.2.1 dos fundamentos teórico-metodológicos é a base para a execução desta etapa dos procedimentos. A própria estruturação da referida seção teve como meta facilitar a identificação dos tipos de interações ecológicas discutidos na literatura especializada. Esta etapa da metodologia foi denominada análise sistemática da literatura porque teve como objetivo sistematizar e gerar uma representação conceitual das interações ecológicas encontradas na literatura. As interações ecológicas foram representadas em uma estrutura classificatória conforme o tipo de afetação (interação harmônica ou desarmônica) e de acordo com os organismos envolvidos (interações interespecíficas ou intraespecíficas). O resultado desta etapa pode ser conferido na Figura 33.

Figura 33 – Classificação das interações ecológicas



Fonte: elaborado pelo proponente (2019)

Nota: os relacionamentos entre os conceitos do modelo foram estabelecidos a partir de fontes consultadas na literatura (CASSINI, 2005; ODUM; BARRET, 2006; PINTO-COELHO, 2007; STEIN, 2018).

Há grande discussão no campo da Ecologia sobre a classificação das interações ecológicas (BRONSTEIN, 2015). Seria necessário um aprofundamento teórico no

tema para explicitar as relações semânticas e os papéis de cada organismo nas interações ecológicas, pois a função que cada organismo exerce na interação é o que vincula um organismo ao outro. Por exemplo, um predador se alimenta de uma presa na interação do tipo predatismo. Logo, no registro de metadados do predador, deveria haver um campo com o nome “predadorDe” onde pudesse ser incluído o(s) nome(s) e/ou identificador(es) do(s) organismo(s) que servem como alimento para o predador. Já no registro de metadados da presa, deveria haver um campo “éPredadoPor” para indicar o(s) predador(es) do organismo representado no registro de metadados. De tal maneira, “predadorDe” e “éPredadoPor” seriam propriedades da classe Predatismo. Entretanto, como será discutido adiante, o escopo da pesquisa não incluiu a definição detalhada de tais relacionamentos, exigindo uma análise mais rigorosa da literatura, o que não seria possível dentro do tempo de um mestrado. Desta forma, os resultados desta pesquisa têm como base os termos apresentados na Figura 33 que, seguindo a lógica do DwC, são classes de metadados.

4.1.3 Análise do núcleo principal de termos do Darwin Core

A análise desta subetapa consistiu em verificar as 12 classes do núcleo básico de termos do DwC, para localizar metadados em potencial para o campo das interações ecológicas na agrobiodiversidade.

Os metadados do núcleo principal do DwC foram recuperados do repositório do DwC, publicado por meio da plataforma Zenodo, sob responsabilidade do repositório publicado pelo DwCMG (2014), que foi atualizado pela última vez em 19 de outubro de 2018. A publicação oficial do DwC, e que deve ser citada, é o artigo de Wieczorek *et al.* (2012), conforme descrito no site do padrão de metadados. Entretanto, os termos analisados nesta etapa da metodologia encontram-se atualizados no repositório citado como DwCMG (2014). À vista disso, as duas referências mencionadas foram constantemente citadas conjuntamente nos resultados desta etapa da metodologia.

Os metadados do DwC são representados por termos (que também podem ser denominados propriedades ou classes). Esses termos são organizados em instâncias ou classes, portanto, a análise foi orientada pela estrutura classificatória do próprio padrão de metadados. Para facilitar o entendimento, as definições dos metadados foram traduzidas e adaptadas, de acordo com a ordem de exibição das classes no

website do DwC e organizadas do Quadro 8 ao Quadro 20.

Executou-se uma síntese do conteúdo temático das classes de metadados para determinar suas respectivas abrangências temáticas, visto que o propósito desta pesquisa não foi discutir individualmente os metadados DwC, mas detectar a existência ou inexistência de metadados úteis para a representação das interações ecológicas.

As categorias de termos do núcleo principal do DwC estão dispostas nos Quadro 8 a Quadro 20. Para cada elemento desses quadros, são associados o nome do elemento em formato *lower CamelCase*³¹ e uma definição. Os termos foram mantidos em inglês, a língua original do padrão DwC, uma vez que a tradução dos termos dos metadados deve ser evitada, pois pode gerar problemas de interoperabilidade caso os metadados não sejam aplicados corretamente, como apontado em estudo realizado por Soares, Hamanaka e Maculan (2018). Não é necessário, para efeito de comparação, estabelecer a definição clara de cada elemento do DwC neste primeiro momento, porém, uma breve definição é apresentada em cada quadro para facilitar o entendimento do conceito representado pelo elemento. Seria necessário demonstrar exemplos reais de aplicação juntamente com as definições para facilitar a compreensão dos termos, mas o excesso de informações poderia desviar a atenção do objetivo real desta etapa, que é levantar os termos dos objetos do insumo metodológico para serem correlatados na Seção 4.1.5. Quando necessário, foram incluídas informações complementares nas etapas seguintes sobre os elementos do DwC.

O Quadro 8 arrola metadados de registro documental que descrevem os dados administrativos do recurso representado. Cumprem o papel de descrever o tipo de documento que o recurso representa, utilizando para isso metadados importados do DC, que possuem o código *dcterms* no Quadro 8, e metadados próprios do DwC, que possuem o código *dwc*. De acordo com o Biodiversity Information Standards (TDWG,

³¹ O formato CamelCase “descreve uma palavra composta que usa letras maiúsculas para delimitar as partes da palavra. O nome refere-se às letras maiúsculas internas, que lembram as corcundas nas costas de um camelo. Por exemplo, ComputerHope, FedEx e WordPerfect são exemplos do CamelCase. Na programação de computadores, o CamelCase é frequentemente usado como uma convenção de nomenclatura para nomear variáveis, matrizes e outros elementos.” (COMPUTER HOPE, 2017, *online*). No caso do DwC, é aplicado o format lower CamelCase, em que a letra inicial do primeiro termo é minúscula.

2015b), a combinação de metadados do DC e DwC é utilizada na descrição do registro para possibilitar a compreensão da natureza do registro documental e a natureza da ocorrência que aquele recurso representa. Por exemplo, a fotografia de um pássaro feita por meio de uma armadilha fotográfica³² seria enquadrada no elemento *dcterms:type* como *stillimage*, que pode ser traduzida como imagem fixa, que não está em movimento (uma fotografia, que é diferente de um vídeo) (TDWG, 2015b). Já o elemento *dwc:basisOfRecord* apresentaria o valor *MachineObservation* (Observação de máquina, em português), para demonstrar o tipo de ocorrência de espécie que aquela fotografia representa, uma vez que a armadilha fotográfica é um tipo de dispositivo automático utilizado no registro de ocorrências e é considerado observação de máquina (TDWG, 2015b). Em resumo, a classe *Record-level* agrega metadados que descrevem o tipo documental e classificam a ocorrência.

Quadro 8 – Classe *Record-level*

(Continua)

ELEMENTO	DEFINIÇÃO
<i>dcterms:type</i>	A natureza ou classe do recurso.
<i>dcterms:modified</i>	A data mais recente em que o recurso foi alterado. Para DwC, a prática recomendada é a utilização de um esquema de codificação, como a ISO 8601:2004.
<i>dcterms:language</i>	O idioma do recurso. A prática recomendada é a utilização de um vocabulário controlado, como ISO 639.
<i>dcterms:license</i>	Um documento legal que concede permissão oficial para fazer algo com o recurso.
<i>dcterms:rightsHolder</i>	Uma pessoa ou organização que possui direitos sobre o recurso.
<i>dcterms:accessRights</i>	Informações sobre quem pode acessar o recurso ou uma indicação de seu status de segurança. Direitos de acesso podem incluir informações sobre acesso ou restrições baseadas em privacidade, segurança ou outras políticas.
<i>dcterms:bibliographicCitation</i>	Uma referência bibliográfica para o recurso como uma declaração indicando como esse registro deve ser citado (atribuído) quando usado. A prática recomendada é incluir suficientes detalhes bibliográficos para identificar o recurso da maneira menos ambígua possível.
<i>dcterms:references</i>	Um recurso relacionado que é referenciado, citado ou apontado de outra forma pelo recurso descrito.
<i>dwc:institutionID</i>	Um identificador para a instituição que possui a custódia do(s) objeto(s) ou das informações mencionadas no registro.

³² Dispositivo fotográfico automático montado por pesquisadores de campo em florestas ou outros tipos de ecossistemas para fotografar os seres vivos em seus ambientes naturais.

Quadro 8 – Classe Record-level

(Conclusão)

ELEMENTO	DEFINIÇÃO
dwc:collectionID	Um identificador para a coleção ou conjunto de dados a partir do qual o registro é derivado. Para amostras físicas, a prática recomendada é usar o identificador de um cadastro de coleções, como o The Global Registry of Biorepositories (GRBio) (http://grbio.org/).
dwc:datasetID	Um identificador para o conjunto de dados. Pode ser um identificador exclusivo global ou um identificador específico para uma coleção ou instituição.
dwc:institutionCode	O nome (ou sigla) em uso pela instituição que possui a custódia do(s) objeto(s) ou das informações mencionadas no registro.
dwc:collectionCode	O nome, acrônimo, coden ou inicialismo que identifica a coleção ou o conjunto de dados do qual o registro foi derivado.
dwc:datasetName	O nome que identifica o conjunto de dados do qual o registro foi derivado.
dwc:ownerInstitutionCode	O nome (ou sigla) em uso pela instituição que possui a propriedade do(s) objeto(s) ou das informações mencionadas no registro.
dwc:basisOfRecord	Termo obrigatório para registros de ocorrências. A natureza específica do registro de dados. A prática recomendada é a utilização de um vocabulário controlado, tais como o Darwin Core Type Vocabulary.
dwc:informationWithheld	Informações adicionais que existem, mas que não foram compartilhadas no registro fornecido.
dwc:dataGeneralizations	Ações feitas para tornar menos específicos ou completos, do que em sua forma original, os dados fornecidos. É sugerido que os dados alternativos de maior completitude possam estar disponíveis a pedido.
dwc:dynamicProperties	Uma lista (contínua e com os valores separados por ";") de medições, fatos, características ou afirmações adicionais sobre o registro. Pretende fornecer um mecanismo para estruturar conteúdos, tais como pares de propriedade-valor.

Fonte: DwCMG (2014) e Wieczorek *et al.* (2012)

Nota: definições traduzidas e adaptadas pelo SiBBR (2018, *online*), contendo também pequenas alterações do proponente; os termos do quadro estão de acordo com a última atualização do Darwin Core quick reference guide, que aconteceu em outubro de 2018. As traduções feitas pelo SiBBR não se encontravam mais disponíveis na data do último acesso, em 10 set. 2019.

Uma ocorrência é descrita como “a existência de um organismo em um lugar e tempo particulares” (WIECZOREK *et al.*, 2009a, *online*, tradução do proponente). O Quadro 9 compila os metadados que tem como função possibilitar a identificação da ocorrência no registro de metadados.

Quadro 9 – Classe *Occurrence*

(Continua)

ELEMENTO	DEFINIÇÃO
occurrenceID	Termo obrigatório para registros de ocorrências. Um identificador para a ocorrência. Na ausência de um identificador único, global e persistente, pode se construir um a partir de uma combinação de identificadores no registro que pode tornar o occurrenceID único globalmente. Na ausência de um identificador único, global e persistente, por exemplo, use a forma: urn:catalog:[institutionCode]:[collectionCode]:[CatalogNumber]
catalogNumber	Um identificador (de preferência único) para o registro dentro do conjunto de dados ou coleção.
recordNumber	Um identificador dado à ocorrência, quando foi registrada. Muitas vezes serve como um elo entre as notas de campo e um registro de ocorrência, tais como o número de coletor das amostras.
recordedBy	Uma lista (contínua e com os valores separados por ";") de nomes de pessoas, grupos ou organizações responsáveis pelo registro da ocorrência original. O coletor ou observador primário, especialmente aquele que dá um identificador pessoal (recordNumber), deve ser o primeiro da lista.
individualCount	O número de indivíduos representados e presente no momento da ocorrência.
organismQuantity	Um número ou valor de enumeração para a quantidade de organismos. Um dwc: organismQuantity deve ter um dwc: organismQuantityType correspondente.
organismQuantityType	O tipo de sistema de quantificação usado para a quantidade de organismos. Um dwc: organismQuantityType deve ter um dwc: organismQuantity correspondente.
Sex	O sexo do(s) indivíduo(s) biológico(s) representado(s) na ocorrência. A prática recomendada é a utilização de um vocabulário controlado.
lifeStage	A idade ou fase da vida do(s) indivíduo(s) biológico(s) quando a ocorrência foi registrada. A prática recomendada é a utilização de um vocabulário controlado.
reproductiveCondition	A condição reprodutiva do(s) indivíduo(s) biológico(s) representado(s) na ocorrência. A prática recomendada é a utilização de um vocabulário controlado.
behavior	A descrição do comportamento mostrado pelo espécime quando a ocorrência foi registrada. A prática recomendada é a utilização de um vocabulário controlado.
establishmentMeans	O processo pelo qual o(s) indivíduo(s) biológico(s) representado(s) na ocorrência ficaram estabelecidos no local. A prática recomendada é a utilização de um vocabulário controlado.

Quadro 9 – Classe *Occurrence*

(Conclusão)

ELEMENTO	DEFINIÇÃO
occurrenceStatus	Uma declaração sobre a presença ou ausência de um táxon em um local. A prática recomendada é a utilização de um vocabulário controlado.
preparations	Uma lista (contínua e com os valores separados por ";") de preparações e métodos de preservação dos espécimes.
disposition	O estado atual de um exemplar com respeito à coleção identificada em collectionCode ou collectionID. A prática recomendada é a utilização de um vocabulário controlado.
associatedMedia	Uma lista (contínua e com os valores separados por ";") de identificadores (publicação, identificador único global, URI) da mídia associada com a ocorrência.
associatedReferences	Uma lista (contínua e com os valores separados por ";") de identificadores (publicação, referência bibliográfica, identificador único e global, URI) da literatura associada à ocorrência.
associatedSequences	Uma lista (contínua e com os valores separados por ";") de identificadores (publicação, referência bibliográfica, identificador único global, URI) de informações da sequência genética associada com a ocorrência.
associatedTaxa	Uma lista (contínua e com os valores separados por ";") de identificadores ou nomes de táxons e suas associações com a ocorrência.
otherCatalogNumbers	Uma lista (concatenada e separada) de números de catálogo completos anteriores ou alternativos ou outros identificadores usados pelo homem para a mesma Ocorrência, seja no conjunto de dados ou na coleção atual ou em qualquer outro conjunto de dados. A melhor prática recomendada é separar os valores em uma lista com espaço na barra vertical ().
occurrenceRemarks	Comentários ou notas sobre a ocorrência.

Fonte: DwCMG (2014) e Wieczorek *et al.* (2012)

Nota: definições traduzidas e adaptadas pelo SiBBR (2018, online), contendo também pequenas alterações do proponente; os termos do quadro estão de acordo com a última atualização do Darwin Core quick reference guide, que aconteceu em outubro de 2018. As traduções feitas pelo SiBBR não se encontravam mais disponíveis na data do último acesso (em 10 set. 2019).

Os metadados desta classe cumprem os seguintes papéis:

- a) identificar o registro como um documento e facilitar sua organização em coleções, fornecendo números de identificação (como *occurrenceID*, *catalogNumber* e *recordNumber*);
- b) indicar os atores envolvidos no processo de registro da ocorrência (vide *recordedBy* no Quadro 9);

- c) apresentar os aspectos relacionados ao(s) organismo(s) que foram objeto de análise na ocorrência, no exato momento em que foram observados ou coletados (esses aspectos são representados por metadados como *individualCount*, *behavior* e *occurrenceStatus*);
- d) apontar as técnicas e procedimentos utilizados no registro da ocorrência (como *establishmentMeans*, *preparations* e *disposition*);
- e) fornecer informações adicionais associadas à ocorrência (tais como *associatedReferences*, *otherCatalogNumbers* e *occurrenceRemarks*).

A classe <dwc:Organism> mostrada no Quadro 10 cumpre a função de identificar o organismo presente na ocorrência. De acordo com Wieczorek *et al.* (2009a, *online*), essa classe facilita o ligamento de uma ou mais instâncias da classe <dwc:Identification> com uma ou mais instâncias da classe <dwc:Occurrence>. A classe <dwc:Identification> permite identificar de forma detalhada os organismos (<dwc:Organism>) presentes na ocorrência (<dwc:Occurrence>), logo a classe <dwc:Organism> faz a ligação entre as classes <dwc:Occurrence> e <dwc:identification>. Além disso, os metadados desta classe permitem identificar o que está sendo tratado como organismo na ocorrência (vide <dwc:organismScope> no Quadro 10 a seguir): pode-se identificar itens que recebem nomes científicos, como vírus, híbridos e líquenes (ou seja, organismos individuais); e itens que são registrados em conjuntos na ocorrência, como pacotes contendo vários itens biológicos, clones ou colônias (agregações).

Quadro 10 – Classe *Organism*

ELEMENTO	DEFINIÇÃO
organismID	Um identificador para a classe <i>Organism</i> (em oposição a um registro digital específico do organismo). Pode ser um identificador globalmente exclusivo ou um identificador específico para o conjunto de dados.
organismName	Um nome ou rótulo textual atribuído a uma instância <i>Organism</i> .
organismScope	Uma descrição do tipo de instância do organismo. Pode ser usado para indicar se a classe <i>Organism</i> representa um organismo discreto ou se representa um tipo específico de agregação.
associatedOccurrences	Uma lista (concatenada e separada) de identificadores de outros registros de Ocorrência e suas associações a essa Ocorrência.
associatedOrganisms	Uma lista (concatenada e separada) de identificadores de outros organismos e suas associações a esse organismo.
previousIdentifications	Uma lista (concatenada e separada) de atribuições anteriores de nomes ao organismo. A melhor prática recomendada é separar os valores em uma lista com espaço na barra vertical ().
organismRemarks	Comentários ou notas sobre a classe <i>Organism</i> .

Fonte: DwCMG (2014) e Wieczorek *et al.* (2012)

Nota: definições traduzidas e adaptadas pelo SiBBR (2018, *online*), contendo também pequenas alterações do proponente. As traduções feitas pelo SiBBR não se encontravam mais disponíveis na data do último acesso (em 10 set. 2019).

A classe *MaterialSample*, do Quadro 11, representa informações pertencentes a uma amostragem executada em um evento (quando há amostragem). “Em coleções biológicas, o material de amostragem é tipicamente coletado e preservado ou destrutivamente processado” (TDWG, 2015a, *online*, tradução do proponente). O material de amostragem pode ser composto por organismos inteiros (como besouros, sapos, bromélias) ou por partes de organismos (como folhas, órgãos internos, sangue); a coleta de partes de organismos pode ser considerada subamostragem e pode destruir o organismo. Essa classe pode representar também amostras de materiais não biológicos, como solo, por exemplo. Não está claro porque a classe *MaterialSample* não aparece junto às classes *PreservedSpecimen*, *FossilSpecimen* e *LivingSpecimen* (vide Quadro 20), já que todas essas classes dizem respeito aos itens que podem ser analisados em um evento (vide Quadro 12). Outra peculiaridade que envolve *MaterialSample* é que ela pode ser utilizada como valor no campo *basisOfRecord* (vide Quadro 8) para descrever a natureza do registro. Esta classe

possui apenas um elemento (*materialSampleID*); os outros metadados relacionados ao processo de amostragem estão presentes na classe evento no Quadro 12.

Quadro 11 – Classe *MaterialSample*

ELEMENTO	DEFINIÇÃO
materialSampleID	Um identificador para a classe <i>MaterialSample</i> (em contraposição a um registo digital específico do material da amostra). Na ausência de um identificador único, global e persistente, construir um a partir de uma combinação de identificadores no registro que pode tornar o materialSampleID único globalmente.

Fonte: DwCMG (2014) e Wieczorek *et al.* (2012)

Nota: definições traduzidas e adaptadas pelo SiBBR (2018, *online*), contendo também pequenas alterações do proponente. As traduções feitas pelo SiBBR não se encontravam mais disponíveis na data do último acesso (em 10 set. 2019).

Um evento, representado pelos metadados do Quadro 12, é uma “ação que acontece em determinado lugar durante algum tempo” (DwCMG, 2014, *online*, tradução do proponente). Essa ação pode ser, por exemplo, fotografar um pássaro num local específico (geralmente demarcado por coordenadas geográficas), em uma determinada floresta. A classe *Event* possui metadados que descrevem o período em que ocorreu o evento, ficando a cargo da descrição detalhada do local a classe *Location*, do Quadro 13. Inclui também metadados para descrever os dados da amostragem, o que liga a classe *Event*, do Quadro 12, à classe *MaterialSample*, do Quadro 11. Além disso, o evento e a ocorrência, do Quadro 9, estão relacionados semanticamente, e esse relacionamento é intrínseco, o que significa que o funcionalismo dessas classes depende desse relacionamento. Explica-se tal relação devido à ocorrência ser a existência de um organismo em determinado local e tempo e o evento ser o encontro do ator com esse organismo em determinado local e tempo. Esse ator pode ser um ser humano, uma armadilha fotográfica ou outro dispositivo que permita gravar o encontro. Observe alguns exemplos apresentados do DwCMG (2014, *online*, tradução do proponente):

- a) uma matilha de lobos na margem do lago Kluane em 1988;
- b) um vírus em uma folha de planta no Jardim Botânico de Nova York às 15:29 em 23/10/2014;
- c) um fungo no Central Park no verão de 1929.

Esses exemplos demonstram a existência de um organismo ou conjunto de indivíduos em determinado local, em um período específico. Pode-se dizer que aquele organismo

“ocorre” naquela região, daí o uso do termo ocorrência. Observe outros exemplos apresentados pelo DwCMG (2014, *online*, tradução do proponente):

- a) um processo de coleta de amostras;
- b) uma captura de imagem da armadilha da câmera;
- c) rede de arrasto marinha;

Os exemplos descrevem ações realizadas com o intuito de registrar a presença de determinado(s) organismo(s). Portanto, de maneira resumida, a classe evento do Quadro 12 é composta por metadados para descrever a ação ou ações executadas para registrar a ocorrência de determinado organismo. É possível notar que o DwC compreende uma rede de ligações semânticas entre as classes e que elas dificilmente poderiam ser aplicadas de maneira individual.

Quadro 12 – Classe *Event*

ELEMENTO	DEFINIÇÃO
eventID	Um identificador para o conjunto de informações associadas a um evento (algo que ocorre em um lugar e tempo). Pode ser um identificador único e global ou um identificador específico dentro do conjunto de dados.
parentEventID	Um identificador para o evento mais amplo que agrupa esse e potencialmente outros eventos.
fieldNumber	Um identificador dado ao evento no campo. Muitas vezes serve como um link entre as notas de campo e o evento.
eventDate	A data ou intervalo durante o qual ocorreu um evento. Para ocorrências, esta é a data quando o evento foi registrado. Não é adequado para um tempo em um contexto geológico. A prática recomendada é a utilização de um esquema de codificação, como a ISO 8601:2004 (E).
eventTime	A hora ou intervalo durante o qual ocorreu um evento. A prática recomendada é a utilização de um esquema de codificação, como a ISO 8601:2004 (E).
startDayOfYear	O primeiro dia ordinal do ano em que ocorreu o evento (1 para 1 de Janeiro, 365 para 31 de Dezembro, exceto em um ano bissexto, que no caso é 366).
endDayOfYear	O último dia ordinal do ano em que ocorreu o evento (1 para 1 de Janeiro, 365 para 31 de dezembro, exceto em um ano bissexto, que no caso é 366).
year	O ano de quatro dígitos em que ocorreu o evento, de acordo com o calendário da Era Comum.
month	O mês, em número inteiro, em que ocorreu o evento.
day	O dia, em números inteiros, do mês em que ocorreu o evento.
verbatimEventDate	A representação textual e original da data e as informações de tempo para um evento.
habitat	A categoria ou descrição do habitat em que ocorreu o evento.
samplingProtocol	O nome, referência ou descrição do método ou protocolo usado durante um Evento.
sampleSizeValue	Um valor numérico para uma medição do tamanho (duração, duração, área ou volume) de uma amostra em um evento de amostragem. Um sampleSizeValue deve ter um sampleSizeUnit correspondente.
sampleSizeUnit	A unidade de medida do tamanho (duração, duração, área ou volume) de uma amostra em um evento de amostragem. Um sampleSizeUnit deve ter um sampleSizeValue correspondente, por exemplo, 5 para sampleSizeValue com medidor para sampleSizeUnit.
samplingEffort	A quantidade de esforço gasto durante um evento.
fieldNotes	Pode ser: a) um indicador da existência de, b) uma referência a (publicação, URI) ou c) o texto das anotações feitas no campo sobre o Evento.
eventRemarks	Comentários ou notas sobre o evento.

Fonte: DwCMG (2014) e Wieczorek *et al.* (2012)

Nota: definições traduzidas e adaptadas pelo SiBBR (2018, *online*), contendo também pequenas alterações do proponente. As traduções feitas pelo SiBBR não se encontravam mais disponíveis na data do último acesso (em 10 set. 2019).

A classe de localização exibida no Quadro 13 permite descrever dados do local da ocorrência, do mais genérico ao mais específico. É recomendável sempre fornecer os dados mais precisos quanto possível, com as coordenadas geográficas completas (metadados como *decimalLatitude*, *decimalLongitude* e *geodeticDatum*). Em discussão com especialistas em registro de ocorrências, tomou-se nota que, na maioria dos casos, o pesquisador que registra a ocorrência não registra as coordenadas exatas do local onde o evento ocorreu, fornecendo muitas vezes coordenadas incertas no momento de escrever o registro, que geralmente é feito muito tempo após o evento e fora do local da ocorrência. Para esses casos, o DwC possui metadados para descrição mais genérica, como *higherGeography*, *waterBody*, *country* e *municipality*, conforme o Quadro 13.

Quadro 13 – Classe *Location*

(Continua)

ELEMENTO	DEFINIÇÃO
locationID	Um identificador para o conjunto de informações do local (dados associados à localização). Pode ser um identificador único e global ou um identificador específico dentro do conjunto de dados.
higherGeographyID	Um identificador da região geográfica dentro da qual ocorreu o local. A prática recomendada é a utilização de um identificador persistente de um vocabulário controlado, como o Banco de Nomes Geográficos do Brasil (http://www.ngb.ibge.gov.br/default.aspx?pagina=nomesgeograficos).
higherGeography	Uma lista (contínua e com os valores separados por ";") de nomes geográficos menos específicos do que as informações capturadas no termo <i>locality</i> .
continent	O nome do continente em que a localização ocorre. A prática recomendada é a utilização de um vocabulário controlado, como o TGN.
waterBody	O nome do corpo de água em que ocorre o local. A prática recomendada é a utilização de um vocabulário controlado, como o TGN.
islandGroup	O nome do grupo de ilhas em que ocorre o local. A prática recomendada é a utilização de um vocabulário controlado, como o TGN.
island	O nome da ilha ou perto da qual a localização ocorre. A prática recomendada é a utilização de um vocabulário controlado, como o TGN.
country	O nome do país ou maior unidade administrativa em que o local ocorre. A prática recomendada é a utilização de um vocabulário controlado, como o TGN http://www.getty.edu/research/tools/vocabularies/tgn/ ou ISO 3166 http://es.wikipedia.org/wiki/ISO_3166-1

Quadro 13 – Classe *Location*

(Continuação)

ELEMENTO	DEFINIÇÃO
countryCode	O código padrão para o país em que o local ocorre. A prática recomendada é usar o ISO 3166 http://es.wikipedia.org/wiki/ISO_3166-1
stateProvince	O nome da região administrativa menor do que país (estado), em que o local ocorre.
county	O nome completo e não abreviado de uma região administrativa menor que o stateProvince (município), em que ocorre o local.
municipality	O nome completo e não abreviado da próxima região administrativa menor do que county (Distrito ou cidade), em que ocorre o local. Não use este termo para incluir um nome de um lugar próximo que não contém o local real.
locality	A descrição específica do local. Informação Geográfica de menor especificidade pode ser fornecida em outros elementos geográficos (higherGeography, continent, country, stateProvince, county, municipality, waterBody, island, islandGroup). Este termo pode conter informações modificadas a partir da original para corrigir erros percebidos ou padronizar a descrição.
verbatimLocality	A descrição textual e original do lugar.
minimumElevationInMeters	O limite inferior do intervalo de elevação (altitude, geralmente acima do nível do mar), em metros.
maximumElevationInMeters	O limite superior do intervalo de elevação (altitude, geralmente acima do nível do mar), em metros.
verbatimElevation	A descrição original da elevação (altitude, geralmente acima do nível do mar) do local.
minimumDepthInMeters	A menor profundidade de um intervalo de profundidades abaixo da superfície do local, em metros.
maximumDepthInMeters	A maior profundidade de um intervalo de profundidades abaixo da superfície do local, em metros.
verbatimDepth	A descrição original da profundidade abaixo da superfície do local.
minimumDistanceAboveSurfaceInMeters	A menor distância de um intervalo de distância a partir de uma superfície de referência em direção vertical, em metros. Use valores positivos para locais acima da superfície, valores negativos para locais abaixo da superfície. Se as medidas de profundidade forem dadas, a superfície de referência é o local indicado pela profundidade, caso contrário, a superfície de referência estará na elevação indicada para o local.
maximumDistanceAboveSurfaceInMeters	A maior distância de um intervalo de distância a partir de uma superfície de referência em direção vertical, em metros. Use valores positivos para locais acima da superfície, valores negativos para locais abaixo da superfície. Se as medidas de profundidade forem dadas, a superfície de referência é o local indicado pela profundidade, caso contrário, a superfície de referência estará na elevação indicada para o local.

Quadro 13 – Classe *Location*

(Continuação)

ELEMENTO	DEFINIÇÃO
locationAccordingTo	As informações sobre a fonte de informação do local específico. Pode ser uma publicação (dicionário geográfico), instituição ou grupo de indivíduos.
locationRemarks	Comentários ou notas sobre o local.
decimalLatitude	A latitude geográfica (em graus decimais, utilizando o sistema de referência espacial dado em geodeticDatum) do centro geográfico de uma local. Os valores positivos são ao norte do Equador, os valores negativos são ao sul. Os valores corretos ficam entre -90 e 90.
decimalLongitude	A longitude geográfica (em graus decimais, utilizando o sistema de referência espacial dada em geodeticDatum) do centro geográfico de uma local. Os valores positivos estão a leste do meridiano de Greenwich, valores negativos são a oeste dela. Os valores legais encontram-se entre -180 e 180.
geodeticDatum	O elipsóide, datum geodésico, ou sistema de referência espacial (SRS) em que as coordenadas geográficas dadas decimalLatitude e decimalLongitude estão baseadas. A prática recomendada é usar o código EPSG como um vocabulário controlado para fornecer um SRS, se conhecido. Caso contrário, use um vocabulário controlado para o nome ou código do datum geodésico, se conhecido. Caso contrário, use um vocabulário controlado para o nome ou código do elipsoide, se conhecido. Se nenhum desses é conhecido, use o valor "desconhecido".
coordinateUncertaintyInMeters	A distância horizontal (em metros) a partir de uma latitude e longitude decimais dadas descrevendo o círculo menor que contém a totalidade do local. Deixe o valor em branco se a incerteza é desconhecida, não pode ser estimada, ou não se aplica (porque não existem coordenadas). Zero não é um valor válido para este termo.
coordinatePrecision	A representação decimal da precisão das coordenadas dadas em decimalLatitude e decimalLongitude.
pointRadiusSpatialFit	A relação entre a área do point-radius (decimalLatitude, decimalLongitude, coordinateUncertaintyInMeters) para a área da representação espacial verdadeira (original, ou mais específica) do local. Os valores permitidos são 0, maior do que ou igual a 1, ou indefinido. Um valor de 1 é uma correspondência exata ou 100% de sobreposição. Um valor de 0 deve ser utilizado se a point-radius não contém completamente a representação inicial. O pointRadiusSpatialFit é indefinido (e deve ser deixado em branco) se a representação original é um ponto e a georreferência dada não é o mesmo ponto. Se tanto a georreferência original quanto a georreferência dada são o mesmo ponto, o pointRadiusSpatialFit é 1.
verbatimCoordinates	As coordenadas espaciais originais e textuais do local. O elipsoide das coordenadas, geodeticDatum ou SRS para estas coordenadas devem ser incluídos em verbatimSRS e o sistema de coordenadas deve ser incluído em verbatimCoordinateSystem.

Quadro 13 – Classe *Location*

(Continuação)

ELEMENTO	DEFINIÇÃO
verbatimLatitude	A latitude textual e original do Local. A coordenada elipsoide, geodeticDatum ou SRS completo para estas coordenadas devem ser armazenados em verbatimSRS e o sistema de coordenadas deve ser armazenado em verbatimCoordinateSystem.
verbatimLongitude	A longitude textual e original do Local. A coordenada elipsoide, geodeticDatum ou SRS completo para estas coordenadas devem ser armazenados em verbatimSRS e o sistema de coordenadas deve ser armazenado em verbatimCoordinateSystem.
verbatimCoordinateSystem	O sistema de coordenadas espaciais para o verbatimLatitude e verbatimLongitude ou verbatimCoordinates do local. A prática recomendada é a utilização de um vocabulário controlado.
verbatimSRS	O elipsoide, datum geodésico, ou SRS para as coordenadas dadas em verbatimLatitude e verbatimLongitude, ou verbatimCoordinates no qual estão baseadas. A prática recomendada é usar o código EPSG como um vocabulário controlado para fornecer um SRS, se conhecido. Caso contrário, use um vocabulário controlado para o nome ou código do datum geodésico, se conhecido. Caso contrário, use um vocabulário controlado para o nome ou código do elipsoide, se conhecido. Se nenhum deles é conhecido, use o valor "desconhecido".
footprintWKT	Uma <i>Well-Known Text</i> ³³ (WKT) da forma (pegada, geometria) que define o local. Um local pode ter uma representação de raio de ponto (consulte decimalLatitude) e uma representação de pegada, e elas podem diferir uma da outra.
footprintSRS	Uma representação em WKT do SRS para o footprintWKT do local. Não use este termo para descrever o SRS do decimalLatitude e decimalLongitude, mesmo que seja o mesmo que para o footprintWKT, em vez use o geodeticDatum.
footprintSpatialFit	A relação entre a área de footprint (footprintWKT) e a área da representação espacial verdadeira (original, ou mais específica) do Local. Os valores permitidos são 0, maior do que ou igual a 1, ou indefinido. Um valor de 1 é uma correspondência exata ou 100% de sobreposição. Um valor de 0 deve ser utilizado se a footprint não contém completamente a representação inicial. O footprintSpatialFit é indefinido (e deve ser deixado em branco) se a representação original é um ponto e a georreferência dada não é o mesmo ponto. Se tanto a georreferência original quanto a georreferência dada são o mesmo ponto, o footprintSpatialFit é 1.

³³ Padrão [...] que é usado para representar dados espaciais em formato textual” (MICROSOFT, 2018, *online*, tradução do proponente).

Quadro 13 – Classe *Location*

(Conclusão)

ELEMENTO	DEFINIÇÃO
georeferencedBy	Uma lista (contínua e com os valores separados por ";") de nomes de pessoas, grupos ou organizações que determinaram o georreferenciamento (representação espacial) para o local.
georeferencedDate	A data em que o local foi georreferenciado. A prática recomendada é a utilização de um esquema de codificação, como a ISO 8601:2004 (E).
georeferenceProtocol	A descrição ou referência aos métodos usados para determinar a descrição espacial, coordenadas e incertezas.
georeferenceSources	Uma lista de mapas, dicionários geográficos ou outros recursos usados para georreferenciar o local (concatenados e separados), descrito com a suficiente especificidade para permitir que qualquer pessoa, no futuro, possa usar os mesmos recursos.
georeferenceVerificationStatus	Uma descrição categórica da medida em que o georreferenciamento foi verificado para representar da melhor maneira possível a descrição espacial. A prática recomendada é a utilização de um vocabulário controlado.
georeferenceRemarks	Notas ou comentários sobre a determinação da descrição espacial, explicando suposições feitas em adição ou oposição daqueles formalizados no método referido no georeferenceProtocol.

Fonte: DwCMG (2014) e Wieczorek *et al.* (2012)

Nota: definições traduzidas e adaptadas pelo SiBBR (2018, *online*), contendo também pequenas alterações do proponente.

A classe de contexto geológico no Quadro 14 aplica-se majoritariamente a recursos do tipo *FossilSpecimen*, no Quadro 20, e *MaterialSample*, no Quadro 11, considerando-se que rochas, amostras de minerais e solo podem ser considerados tipos de *MaterialSample*. Esses recursos são descritos pelo registro de metadados e categorizados pelo elemento *dwc:basisOfRecord* (vide Quadro 8). O objetivo desta classe é fornecer a descrição geológica detalhada do contexto (ambiente) onde o item foi coletado, portanto pode também ser utilizada como um qualificador da região ou local onde determinado espécime foi coletado.

Quadro 14 – Classe *GeologicalContext*

ELEMENTO	DEFINIÇÃO
geologicalContextID	Um identificador para o conjunto de informações associadas ao GeologicalContext (o local dentro de um contexto geológico, como a estratigrafia). Pode ser um identificador único e global ou um identificador específico dentro do conjunto de dados.
earliestEonOrLowestEonothem	O nome completo do eon geocronológico mais antigo possível ou eonothem cronoestratigráfico inferior ou nome informal ("Pré-cambriano") atribuível ao horizonte estratigráfico do qual o item catalogado foi coletado.
latestEonOrHighestEonothem	O nome completo do eon geocronológico mais recente possível ou eonothem cronoestratigráfico superior ou nome informal ("Pré-cambriano") atribuível ao horizonte estratigráfico do qual o item catalogado foi coletado.
earliestEraOrLowestEratthem	O nome completo da era geocronológica mais antiga possível ou eratema cronoestratigráfico inferior atribuível ao horizonte estratigráfico do qual o item catalogado foi coletado.
latestEraOrHighestEratthem	O nome completo da era geocronológica mais recente possível ou eratema cronoestratigráfico superior atribuível ao horizonte estratigráfico do qual o item catalogado foi coletado.
earliestPeriodOrLowestSystem	O nome completo do período geocronológico mais antigo possível ou sistema cronoestratigráfico inferior atribuível ao horizonte estratigráfico do qual o item catalogado foi coletado.
latestPeriodOrHighestSystem	O nome completo do período geocronológico mais recente possível ou sistema cronoestratigráfico superior atribuível ao horizonte estratigráfico do qual o item catalogado foi coletado.
earliestEpochOrLowestSeries	O nome completo da época geocronológica mais antiga possível ou séries cronoestratigráfico inferior atribuível ao horizonte estratigráfico do qual o item catalogado foi coletado.
latestEpochOrHighestSeries	O nome completo da época geocronológica mais recente possível ou séries cronoestratigráfico superior atribuível ao horizonte estratigráfico do qual o item catalogado foi coletado.
earliestAgeOrLowestStage	O nome completo da idade geocronológica mais antiga possível ou andar cronoestratigráfico inferior atribuível ao horizonte estratigráfico do qual o item catalogado foi coletado.
latestAgeOrHighestStage	O nome completo da idade geocronológica mais recente possível ou andar cronoestratigráfico superior atribuível ao horizonte estratigráfico do qual o item catalogado foi coletado.
lowestBiostratigraphicZone	O nome completo da zona bioestratigráfica inferior do horizonte estratigráfico do qual o item catalogado foi coletado.
highestBiostratigraphicZone	O nome completo da zona bioestratigráfica superior do horizonte estratigráfico do qual o item catalogado foi coletado.
lithostratigraphicTerms	A combinação de todos os nomes litoestratigráficos para a rocha da qual o item catalogado foi coletado.
group	O nome completo do grupo litoestratigráfica do qual o item catalogado foi coletado.
formation	O nome completo da formação litoestratigráfica do qual o item catalogado foi coletado.
member	O nome completo do membro litoestratigráfica do qual o item catalogado foi coletado.
bed	O nome completo da Camada litoestratigráfica da qual o item catalogado foi coletado.

Fonte: DwCMG (2014) e Wieczorek *et al.* (2012)

Nota: definições traduzidas e adaptadas pelo SiBBR (2018, *online*), contendo também pequenas alterações do proponente.

A identificação do item coletado é feita por meio dos metadados da classe de identificação, conforme descrito no Quadro 15. Esses metadados descrevem: a) os responsáveis pela identificação do espécime; b) a identificação do espécime em si, fornecendo dados sobre a classificação do espécime da ocorrência como representante de determinado táxon (para táxon, ver Quadro 16).

Quadro 15 – Classe *Identification*

ELEMENTO	DEFINIÇÃO
identificationID	Um identificador para a Identificação (o bloco de informação associada com a atribuição de um nome científico). Pode ser um identificador único e global ou um identificador específico dentro do conjunto de dados.
identificationQualifier	Uma breve frase ou um termo padrão ("cf.", "Aff.") para expressar as dúvidas do determinador sobre a identificação.
typeStatus	Uma lista (contínua e com os valores separados por ";") de tipos nomenclaturais (status do tipo, nome científico tipificado, publicação) aplicada ao espécime.
identifiedBy	Uma lista (contínua e com os valores separados por ";") de nomes de pessoas, grupos ou organizações que atribuíram o táxon ao espécime.
dateIdentified	A data em que o espécime foi identificado como representante de um táxon. A prática recomendada é a utilização de um vocabulário controlado, como o ISO 8601:2004 (E).
identificationReferences	Uma lista (contínua e com os valores separados por ";") de referências (de publicação, identificador único e global, URI) usado para a identificação.
identificationVerificationStatus	Um indicador categórico da medida em que a identificação taxonômica foi verificada como correta. A prática recomendada é a utilização de um vocabulário controlado, como o utilizado em HISPI/ABCD.
identificationRemarks	Comentários ou anotações sobre a identificação.

Fonte: DwCMG (2014) e Wieczorek *et al.* (2012)

Nota: definições traduzidas e adaptadas pelo SiBBR (2018, *online*), contendo também pequenas alterações do proponente.

O táxon, que é o conceito da classe do Quadro 16, é descrito por DwCMG (2014) como um grupo de organismos considerados homogêneos pelos taxonomistas, ou seja, seres vivos que são classificados em um mesmo táxon por possuírem as mesmas características evolutivas e certo grau de parentesco entre si. Logo, a classe de identificação do Quadro 15 é complementar à classe táxon do Quadro 16, sendo que a classe táxon padroniza a maneira de se descrever um organismo, que é a classe

do Quadro 10. Essa classe engloba metadados para a descrição taxonômica completa do organismo e recomenda o uso dos Códigos Internacionais de Nomenclatura³⁴, que são esquemas de codificação, para a padronização dos nomes científicos e indicação de responsabilidade ao registro. Em suma, é a classe táxon que classifica o espécime como representante de uma espécie. Por meio dos metadados desta classe torna-se possível identificar a espécie descrita no registro de metadados em uma coleção. Além da nomenclatura científica, a classe táxon também agrega o elemento *vernacularName*, que possibilita a exibição dos nomes populares ou comuns atribuídos à espécie da qual o espécime faz parte.

Quadro 16 – Classe *Taxon*

(Continua)

ELEMENTO	DEFINIÇÃO
taxonID	Termo obrigatório para listas de espécies. Um identificador para o conjunto de informação do táxon (dados associados com a classe Taxon). Pode ser um identificador único e global ou um identificador específico dentro do conjunto de dados.
scientificNameID	Um identificador para os detalhes nomenclaturais (não taxonômicos) do nome científico.
acceptedNameUsageID	Um identificador para o nome em uso (significado documentado do nome de acordo com uma fonte) do nome atual válido (zoologia) ou aceite (botânica) do táxon.
parentNameUsageID	Um identificador para o nome em uso (significado documentado do nome de acordo com uma fonte) do táxon parental diretamente mais próximo de maior nível hierárquico (em uma classificação) do elemento mais específico do scientificName.
originalNameUsageID	Um identificador para o nome usado (significado documentado do nome de acordo com uma fonte), em que o elemento terminal do scientificName foi originalmente criado sob as regras do nomenclaturalCode associado.
nameAccordingToID	Um identificador para a fonte em que está definido ou implícito a circunscrição do conceito do táxon específico. Veja nameAccordingTo.
namePublishedInID	Um identificador para a publicação em que o scientificName foi originalmente publicado sob as regras do nomenclaturalCode associado.
taxonConceptID	Um identificador para o conceito taxonômico a que o registro se refere - não para os detalhes de nomenclatura de um táxon.

³⁴ Veja os Códigos Internacionais de Nomenclatura na Seção 2.6 deste trabalho.

Quadro 16 – Classe *Taxon*

(Continuação)

ELEMENTO	DEFINIÇÃO
scientificName	O nome científico completo, com autoria e informações da data se conhecidas. Ao formar parte de uma identificação, este deve ser o nome em nível mais baixo da classificação taxonômica possível. Este termo não deve conter qualificações de identificação, que devem antes ser fornecidos no termo IdentificationQualifier.
acceptedNameUsage	O nome completo, com autoria e informação de data se conhecida, do nome atual válido (zoologia) ou aceitado (botânico) do táxon.
parentNameUsage	O nome completo, com autoria e data, se conhecida, do táxon diretamente mais próxima de nível hierárquico maior (em uma classificação) do elemento mais específico do scientificName.
originalNameUsage	O nome do táxon, com autoria e data, se conhecida, da maneira que apareceu originalmente quando foi publicado pela primeira vez sob as regras do nomenclaturalCode associado. O basônimo (botânica) ou basônimo (bacteriologia) do scientificName ou o homônimo sênior para os nomes substituídos.
nameAccordingTo	A referência para a fonte em que está definida ou implícita a circunscrição do conceito do táxon específico. Tradicionalmente representado por a expressão do Latim "sensu" ou "Sec" (a partir de secundum, que significa "de acordo com"). Para táxons que resultam de identificações, devem ser dadas as referências para as chaves, monografias, especialistas e outras fontes.
namePublishedIn	Uma referência para a publicação em que o scientificName foi originalmente publicado sob as regras do nomenclaturalCode associado.
namePublishedInYear	Uma lista de nomes de táxons (contínua e com os valores separados por ";") terminando no nível imediatamente superior ao nome do táxon referenciado no registro do táxon. A prática recomendada é ordenar a lista começando com o nível mais alto.
higherClassification	Uma lista de nomes de táxons (contínua e com os valores separados por ";") terminando no nível imediatamente superior ao nome do táxon referenciado no registro do táxon. A prática recomendada é ordenar a lista começando com o nível mais alto.
kingdom	O nome científico completo do reino em que o táxon é classificado.
phylum	O nome científico completo do filo em que o táxon é classificado.
class	O nome científico completo da classe em que o táxon é classificado.
order	O nome científico completo da ordem em que o táxon é classificado.

Quadro 16 – Classe *Taxon*

(Conclusão)

ELEMENTO	DEFINIÇÃO
family	O nome científico completo da família em que o táxon é classificado.
genus	O nome científico completo do gênero em que o táxon é classificado.
subgenus	O nome científico completo do subgênero em que o táxon é classificado. Os valores devem incluir o gênero para evitar confusão com o homônimo. Exemplo: Puma (Puma)
specificEpithet	O nome do epíteto específico do scientificName.
infraspecificEpithet	O nome do epíteto infraespecífico de menor categoria ou hierarquia do scientificName, excluindo qualquer designação de categoria taxonômica.
taxonRank	O nível taxonômico do nome mais específico em scientificName. A prática recomendada é a utilização de um vocabulário controlado.
verbatimTaxonRank	O nível taxonômico do nome mais específico em scientificName como aparece no registro original.
scientificNameAuthorship	Informações de autoria para o scientificName formatado de acordo com as convenções do nomenclaturalCode aplicável.
vernacularName	Um nome comum ou vernáculo.
nomenclaturalCode	O código de nomenclatura (ou códigos, no caso de um nome ambiregna) sob a qual o scientificName é construído. A prática recomendada é a utilização de um vocabulário controlado. Ex.: ICN, ICZN ₂ , ICNCP, BioCode.
taxonomicStatus	A condição de uso do scientificName como um rótulo para um táxon. Requer opinião taxonômica para definir o escopo de um táxon. Regras de prioridade são usadas para definir a condição taxonômica da nomenclatura contida nesse âmbito, combinados com a opinião de especialistas. Deve estar associada a uma referência taxonômica específica que define o conceito. A prática recomendada é a utilização de um vocabulário controlado.
nomenclaturalStatus	O estado do nome e da sua conformidade com as normas pertinentes da nomenclatura em relação com a publicação original. Baseia-se essencialmente em um algoritmo de acordo com as regras de negócio do código. Ele não exige nenhuma opinião taxonômica.
taxonRemarks	Comentários ou notas sobre o táxon ou nome.

Fonte: DwCMG (2014) e Wieczorek *et al.* (2012)

Nota: definições traduzidas e adaptadas pelo SiBBR (2018, *online*), contendo também pequenas alterações do proponente.

O DwCMG (2014) estipula que a classe do Quadro 17 representa medidas ou fatos sobre medidas de algum recurso. A definição de recurso adotada segue o esquema

RDF³⁵. Os recursos que se relacionam com a classe *MeasurementOrFact*

podem ser considerados registros identificáveis ou instâncias de classes e podem incluir, mas não precisam se limitar, a *dwc:Occurrence*, *dwc:Organism*, *dwc:MaterialSample*, *dwc:Event*, *dwc:Location*, *dwc:GeologicalContext*, *dwc:Identification*, ou *dwc:Taxon* (DwCMG, 2014, *online*, tradução do proponente).

Em outras palavras, a classe de medidas pode ser associada a outras classes do DwC para fornecer informações de medida sobre essas classes. Por exemplo, se combinada com a classe organismo do Quadro 10, pode ser utilizada para fornecer o tamanho de um organismo, ou de parte deste organismo; combinada com a classe evento, pode fornecer o tempo de duração total de um evento; combinada a classe de localização do Quadro 13 permite descrever medidas de espaço; dentre outros exemplos.

Quadro 17 – Classe *MeasurementOrFact*

ELEMENTO	DEFINIÇÃO
measurementID	Um identificador para o Medições ou fatos (informações relativas às medidas, fatos, características ou afirmações). Pode ser um identificador único e global ou um identificador específico dentro do conjunto de dados.
measurementType	A natureza da medição, fato, característica ou afirmação. A prática recomendada é a utilização de um vocabulário controlado.
measurementValue	O valor da medição, fato, característica ou afirmação.
measurementAccuracy	A descrição do erro potencial associado com o measurementValue.
measurementUnit	As unidades associadas com o measurementValue. Prática recomendada é a utilização das unidades do SI.
measurementDeterminedBy	Uma lista (contínua e com os valores separados por ";") de nomes de pessoas, grupos ou organizações que determinaram o valor de Medições ou fatos.
measurementDeterminedDate	A data em que a Medições ou fatos foi feita. A prática recomendada é a utilização de um esquema de codificação, como a ISO 8601:2004 (E).
measurementMethod	Uma descrição ou referência para (publicação, URI) o método ou protocolo utilizado para determinar a medida, fato, característica ou afirmação.
measurementRemarks	Comentários ou notas que acompanham o Medições ou fatos.

Fonte: DwCMG (2014) e Wieczorek *et al.* (2012)

Nota: definições traduzidas e adaptadas pelo SiBBR (2018, *online*), contendo também pequenas alterações do proponente.

³⁵ Para relembrar a definição e a lógica do RDF, veja a Seção 2.3.3 deste trabalho.

O Quadro 18 enumera metadados que formalizam as relações entre recursos (sujeito e objeto) no registro de metadados. Por meio dos metadados desta classe, é possível relacionar dois ou mais registros cujos recursos representados compartilham certa afinidade entre si, especificando quem determinou (*relationshipAccordingTo*), quando (*relationshipEstablishedDate*) e qual a natureza dessa afinidade (*relationshipOfResource*). Esta classe é útil para vincular itens dispostos em coleções biológicas distintas, como por exemplo um fóssil de dinossauro em museu no Brasil que é “irmão” de outro fóssil que faz parte da coleção de um museu na Holanda; ou para determinar relações entre organismos que possuem registros de metadados organizados numa mesma coleção.

Quadro 18 – Classe *ResourceRelationship*

ELEMENTO	DEFINIÇÃO
resourceRelationshipID	Um identificador de uma instância de relação entre um recurso (o sujeito) e outro (<i>relatedResource</i> , o objeto).
resourceID	Um identificador para o recurso que é o tema da relação.
relatedResourceID	Um identificador para um recurso relacionado (o objeto, ao invés do tema da relação).
relationshipOfResource	A relação do recurso identificado em <i>relatedResourceID</i> ao sujeito (opcionalmente identificado pelo <i>resourceID</i>). A prática recomendada é a utilização de um vocabulário controlado.
relationshipAccordingTo	A fonte (pessoa, organização, publicação de referência), que estabelece o relacionamento entre os dois recursos.
relationshipEstablishedDate	A data e hora em que foi estabelecida a relação entre os dois recursos. A prática recomendada é usar um esquema de codificação, tais como ISO 8601: 2004 (E).
relationshipRemarks	Comentários ou anotações sobre a relação entre os dois recursos.

Fonte: DwCMG (2014) e Wieczorek *et al.* (2012)

Nota: definições traduzidas e adaptadas pelo SiBBR (2018, *online*), contendo também pequenas alterações do proponente.

O Quadro 19 elenca todos os termos do DwC que devem ter como valor um IRI³⁶ (DwCMG, 2014), ou seja, todos os campos que recomendam esquemas de codificação. Além dos IRIs, essa classe sugere o uso de vocabulários controlados

³⁶ É “[...] um complemento para URIs. Um IRI é uma sequência de caracteres do Conjunto Universal de Caracteres (Unicode/ISO10646). Há um mapeamento de IRIs para URIs, o que significa que os IRIs podem ser usados em vez dos URIs, quando apropriado, para identificar recursos” (DÜRST, 2011, *online*, tradução do proponente).

como esquemas de codificação para padronizar certos valores que necessitam de descrição textual (que seria considerado um valor literal).

Quadro 19 – Classe *UseWithIRI*

(Continua)

ELEMENTO	DEFINIÇÃO
inDescribedPlace	Use para vincular uma instância de <i>dcterms: Location</i> sujeita ao recurso descrito hierarquicamente e padronizado de nível mais baixo. Uma "propriedade de conveniência" que substitui os termos de valor literal do Darwin Core relacionados aos locais.
identifiedBy	Uma pessoa, grupo ou organização que atribuiu o <i>Taxon</i> ao sujeito. Os termos no <i>namespace</i> <i>dwciri</i> devem ser usados no RDF com objetos não literais.
recordedBy	Uma pessoa, grupo ou organização responsável por registrar a Ocorrência original. Os termos no <i>namespace</i> <i>dwciri</i> devem ser usados no RDF com objetos não literais.
toTaxon	Use para vincular uma classe <i>dwc:Identification</i> sujeita a uma entidade taxonômica, como um táxon, conceito de táxon ou uso de nome de táxon. Uma "propriedade de conveniência" que substitui os termos de valor literal do Darwin Core relacionados a entidades taxonômicas.
inCollection	Use para vincular qualquer recurso de assunto que faça parte de uma coleção à coleção que contém o recurso. A melhor prática recomendada é usar um IRI de um registro controlado. Uma "propriedade de conveniência" que substitui termos de valor literal relacionados a coleções e instituições.
georeferencedBy	Uma pessoa, grupo ou organização que determinou a georreferência (representação espacial) do Local. Os termos no <i>namespace</i> <i>dwciri</i> devem ser usados no RDF com objetos não literais.
behavior	Uma descrição do comportamento mostrado pelo sujeito quando a ocorrência foi registrada. A melhor prática recomendada é usar um vocabulário controlado. Os termos no <i>namespace</i> <i>dwciri</i> devem ser usados no RDF com objetos não literais.
dataGeneralizations	Ações tomadas para tornar os dados compartilhados menos específicos ou completos do que em sua forma original. Sugere que dados alternativos de qualidade superior possam estar disponíveis mediante solicitação. Os termos no <i>namespace</i> <i>dwciri</i> devem ser usados no RDF com objetos não literais.
disposition	O estado atual de uma amostra em relação à coleção identificada em <i>collectionCode</i> ou <i>collectionID</i> . A melhor prática recomendada é usar um vocabulário controlado. Os termos no <i>namespace</i> <i>dwciri</i> devem ser usados no RDF com objetos não literais.

Quadro 19 – Classe *UseWithIRI*

(Continuação)

ELEMENTO	DEFINIÇÃO
earliestGeochronological Era	Use para vincular uma instância <i>dwc:GeologicalContext</i> a períodos cronoestratigráficos no nível mais baixo possível em uma hierarquia padronizada. Use essa propriedade para apontar para o período geológico possível mais antigo a partir do qual o item catalogado foi coletado. A melhor prática recomendada é usar um IRI de um vocabulário controlado. Uma "propriedade de conveniência" que substitui os termos de valor literal do Darwin Core relacionados ao contexto geológico.
establishmentMeans	O processo pelo qual o (s) indivíduo (s) biológico (s) representado (s) na Ocorrência foram estabelecidos no local. A melhor prática recomendada é usar um vocabulário controlado. Os termos no <i>namespace</i> <i>dwciri</i> devem ser usados no RDF com objetos não literais.
fieldNotes	Pode ser: a) um indicador da existência, b) uma referência a (publicação, URI) ou c) o texto das anotações feitas no campo sobre o Evento. O assunto é uma instância <i>dwc:Event</i> e o objeto é um recurso (possivelmente identificado pelo IRI) que são as notas de campo
fieldNumber	Um identificador fornecido ao evento no campo. Frequentemente serve como um link entre as anotações de campo e o Evento. O assunto é um recurso (possivelmente identificado pelo IRI) que são as notas de campo e o objeto é uma instância <i>dwc:Event</i> .
footprintSRS	Uma representação de WKT do SRS para o <i>footprintWKT</i> do local. Não use esse termo para descrever o SRS de <i>decimalLatitude</i> e <i>decimalLongitude</i> , mesmo que seja o mesmo do <i>footprintWKT</i> - use o <i>geodeticDatum</i> . Os termos no <i>namespace</i> <i>dwciri</i> devem ser usados no RDF com objetos não literais.
footprintWKT	Uma representação WKT da forma (pegada, geometria) que define o local. Um local pode ter uma representação de raio de ponto (consulte <i>decimalLatitude</i>) e uma representação de pegada, e elas podem diferir uma da outra. Os termos no <i>namespace</i> <i>dwciri</i> devem ser usados no RDF com objetos não literais.
fromLithostratigraphicUnit	Use para vincular uma instância <i>dwc:GeologicalContext</i> a uma unidade litoestratigráfica identificada pelo IRI no nível mais baixo possível em uma hierarquia. A melhor prática recomendada é usar um IRI de um vocabulário controlado. Uma "propriedade de conveniência" que substitui os termos de valor literal do Darwin Core relacionados ao contexto geológico.
geodeticDatum	O elipsoide, dado geodésico ou SRS no qual as coordenadas geográficas são fornecidas em <i>decimalLatitude</i> e <i>decimalLongitude</i> como base. A melhor prática recomendada é usar um IRI para o código EPSG do SRS, se conhecido. Caso contrário, use um IRI ou vocabulário controlado para o nome ou código do dado geodésico, se conhecido. Caso contrário, use um IRI ou vocabulário controlado para o nome ou código do elipsoide, se conhecido. Se nada disso for conhecido, use o valor <i>unknown</i> .

Quadro 19 – Classe *UseWithIRI*

(Continuação)

ELEMENTO	DEFINIÇÃO
georeferenceProtocol	Uma descrição ou referência aos métodos usados para determinar a pegada espacial, coordenadas e incertezas. Os termos no <i>namespace</i> dwciri devem ser usados no RDF com objetos não literais.
georeferenceSources	Um mapa, gazeta ou outro recurso usado para georreferenciar o local. Os termos no <i>namespace</i> dwciri devem ser usados no RDF com objetos não literais.
georeferenceVerificationStatus	Uma descrição categórica da extensão em que a georreferência foi verificada para representar a melhor descrição espacial possível. A melhor prática recomendada é usar um vocabulário controlado. Os termos no <i>namespace</i> dwciri devem ser usados no RDF com objetos não literais.
habitat	Uma categoria ou descrição do habitat em que o Evento ocorreu. Os termos no <i>namespace</i> dwciri devem ser usados no RDF com objetos não literais.
identificationQualifier	Um valor controlado para expressar as dúvidas do determinante sobre a identificação. Os termos no <i>namespace</i> dwciri devem ser usados no RDF com objetos não literais.
identificationVerificationStatus	Um indicador categórico da extensão em que a identificação taxonômica foi verificada como correta. Os termos no <i>namespace</i> dwciri devem ser usados no RDF com objetos não literais. A melhor prática recomendada é usar um vocabulário controlado, como o usado no HISPID e ABCD.
inDataset	Use para vincular um registro de conjunto de dados de assunto ao conjunto de dados que o contém. Um nome literal de sequência de caracteres do conjunto de dados pode ser fornecido usando o termo <i>dwc:datasetName</i> .
informationWithheld	Informações adicionais que existem, mas que não foram compartilhadas no registro fornecido. Os termos no <i>namespace</i> dwciri devem ser usados no RDF com objetos não literais.
latestGeochronologicalEra	Use para vincular uma instância <i>dwc:GeologicalContext</i> a períodos cronoestratigráficos no nível mais baixo possível em uma hierarquia padronizada. Use esta propriedade para apontar para o período geológico possível mais recente do qual o item catalogado foi coletado. A melhor prática recomendada é usar um IRI de um vocabulário controlado. Uma "propriedade de conveniência" que substitui os termos de valor literal do Darwin Core relacionados ao contexto geológico.
lifeStage	A classe etária ou estágio de vida do(s) indivíduo(s) biológico(s) quando a Ocorrência foi registrada. A melhor prática recomendada é usar um vocabulário controlado. Os termos no <i>namespace</i> dwciri devem ser usados no RDF com objetos não literais.
locationAccordingTo	Informações sobre a fonte dessas informações de localização. Pode ser uma publicação (gazetteer), instituição ou equipe de indivíduos. Os termos no <i>namespace</i> dwciri devem ser usados no RDF com objetos não literais.
measurementDeterminedBy	Uma pessoa, grupo ou organização que determinou o valor do MeasurementOrFact. Os termos no espaço de nomes dwciri devem ser usados no RDF com objetos não literais.
measurementMethod	O método ou protocolo usado para determinar a medida, fato, característica ou afirmação. Os termos no <i>namespace</i> dwciri devem ser usados no RDF com objetos não literais.

Quadro 19 – Classe *UseWithIRI*

(Conclusão)

ELEMENTO	DEFINIÇÃO
measurementType	A natureza da medida, fato, característica ou afirmação. A melhor prática recomendada é usar um vocabulário controlado. Os termos no <i>namespace</i> <i>dwciri</i> devem ser usados no RDF com objetos não literais.
measurementUnit	As unidades associadas com o <i>measurementValue</i> .
occurrenceStatus	A melhor prática recomendada é utilizar o International System of Units (SI). Uma declaração sobre a presença ou ausência de um táxon em um local. A melhor prática recomendada é usar um vocabulário controlado. Os termos no <i>namespace</i> <i>dwciri</i> devem ser usados no RDF com objetos não literais.
organismQuantityType	O tipo de sistema de quantificação usado para a quantidade de organismos. Um <i>dwc:organismQuantityType</i> deve ter um <i>dwc:organismQuantity</i> correspondente.
preparations	Um método de preparação ou preservação de uma amostra. Os termos no <i>namespace</i> <i>dwciri</i> devem ser usados no RDF com objetos não literais.
recordNumber	Um identificador fornecido à Ocorrência quando foi gravado. Frequentemente serve como um link entre as anotações de campo e um registro de Ocorrência, como o número de um coletor de amostras. O assunto é um <i>dwc:Occurrence</i> e o objeto é um recurso (possivelmente identificado pelo IRI) que são as notas de campo.
reproductiveCondition	A condição reprodutiva do(s) indivíduo(s) biológico(s) representado(s) na Ocorrência. A melhor prática recomendada é usar um vocabulário controlado. Os termos no <i>namespace</i> <i>dwciri</i> devem ser usados no RDF com objetos não literais.
sampleSizeUnit	A unidade de medida do tamanho (duração, duração, área ou volume) de uma amostra em um evento de amostragem. Um <i>sampleSizeUnit</i> deve ter um <i>sampleSizeValue</i> correspondente. A melhor prática recomendada é usar um vocabulário controlado, como a Ontology of Units of Measure http://www.wurvoc.org/vocabularies/om-1.8/ de unidades SI, unidades derivadas ou outras unidades não SI aceitas para uso no SI.
samplingProtocol	O método ou protocolo usado durante um evento. Os termos no <i>namespace</i> <i>dwciri</i> devem ser usados no RDF com objetos não literais.
sex	O sexo do(s) indivíduo(s) biológico(s) representado(s) na Ocorrência. A melhor prática recomendada é usar um vocabulário controlado. Os termos no <i>namespace</i> <i>dwciri</i> devem ser usados no RDF com objetos não literais.
typeStatus	Um tipo nomenclatural (status do tipo, nome científico tipificado, publicação) aplicado ao assunto. Os termos no <i>namespace</i> <i>dwciri</i> devem ser usados no RDF com objetos não literais.
verbatimCoordinateSystem	O sistema de coordenadas espaciais para <i>verbatimLatitude</i> e <i>verbatimLongitude</i> ou <i>verbatimCoordinates</i> do local. A melhor prática recomendada é usar um vocabulário controlado. Os termos no <i>namespace</i> <i>dwciri</i> devem ser usados no RDF com valores não literais.
verbatimSRS	O elipsoide, o dado geodésico ou o SRS no qual as coordenadas fornecidas em <i>verbatimLatitude</i> e <i>verbatimLongitude</i> , ou <i>verbatimCoordinates</i> são baseadas. A melhor prática recomendada é usar um IRI para o código EPSG do SRS, se conhecido. Caso contrário, use um IRI ou vocabulário controlado para o nome ou código do dado geodésico, se conhecido. Caso contrário, use um IRI ou vocabulário controlado para o nome ou código do elipsoide, se conhecido. Se nada disso for conhecido, use o valor <i>unknown</i> .

Fonte: DwCMG (2014) e Wieczorek *et al.* (2012)

Nota: definições traduzidas e adaptadas pelo proponente.

Os termos do Quadro 20 são considerados classes no vocabulário de termos do DwC,

mas podem ser consideradas instâncias se aplicados como valor para descrever o tipo de registro no elemento *basisOfRecord* (vide Quadro 8). O DwC foi concebido para facilitar o compartilhamento de dados de espécimes de coleções (WIECZOREK *et al.*, 2012), logo, as classes do Quadro 20 devem ser suficientes para determinar a base do registro, na maioria dos casos.

Quadro 20 – Evidência de ocorrência

ELEMENTO	DEFINIÇÃO
LivingSpecimen	Um espécime que está vivo.
PreservedSpecimen	Um espécime que foi preservado.
FossilSpecimen	Um espécime preservado que é um fóssil.
HumanObservation	Um resultado de um processo de observação humana.
MachineObservation	Um resultado de um processo de observação de máquina.

Fonte: DwCMG (2014) e Wieczorek *et al.* (2012)

Nota: definições traduzidas pelo proponente.

Ao propor os ligamentos entre termos de diferentes classes e elucidar as melhores maneiras de trabalhar esses ligamentos, o DwC cria uma estrutura de dados ligados que atribui mais semântica ao registro de metadados, por permitir que diferentes categorias de dados da ocorrência do organismo sejam registradas de maneira relacional e em contexto. Todas as classes de metadados apresentadas nesta seção estão relacionadas. Os termos de nível de registro (vide Quadro 8) descrevem todo o registro de metadados e são utilizados para organizar o registro em uma coleção ou base de dados; esse registro pode representar uma amostra de material (vide Quadro 11), um espécime vivo, um espécime fóssil, um espécime conservado ou pode consistir apenas uma observação, que pode ser humana ou de máquina (vide Quadro 20); esse registro de metadados representa uma ocorrência (vide Quadro 9), que é a existência de um organismo (vide Quadro 10) em determinado tempo e local; um organismo (vide Quadro 10) é estudado na ocorrência e, para que isso aconteça, muitas vezes é necessário coletar amostras (vide Quadro 11) deste organismo; o ato de coletar as amostras é denominado evento (vide Quadro 12) e acontece em um determinado intervalo de tempo, em certo lugar (vide Quadro 13); esse lugar representa uma região que possui características geológicas (vide Quadro 14); o espécime coletado deve ser identificado (vide Quadro 15) de maneira que fique claro

a qual táxon (vide Quadro 16) aquele espécime pertence – logo, a identificação é o elo de ligação entre a ocorrência e a espécie (táxon) representada pelo espécime; é possível fornecer informações adicionais de medidas (vide Quadro 17) de alguns recursos representados no DwC; por fim, outros recursos podem ser relacionados ao recurso descrito por meio de relações formais (vide Quadro 18). Há também a recomendação de que sejam utilizadores identificadores universais (vide Quadro 19) como valores em diversos termos do DwC como forma de estabelecer linked data de maneira estável. A representação das ligações entre as classes do DwC pode ser visualizada na Figura 28, na Seção 2.6, e representa de maneira clara a lógica do RDF.

Não há na estrutura de representação do núcleo do DwC metadados para a representação de interações ecológicas, como foi possível observar após a análise do núcleo básico de termos. Entretanto, como exposto na Seção 2.6.1, o DwC possui extensões de metadados, e foi necessário verificar se essas extensões continham metadados para a representação das interações ecológicas, o que será exposto na Seção 4.1.4.

4.1.4 Análise das extensões do padrão de metadados Darwin Core

As extensões do DwC também foram analisadas para detectar termos existentes nas extensões do DwC que pudessem ser incorporados a uma extensão de metadados no escopo temático das interações ecológicas na agrobiodiversidade, aplicadas ao controle de pragas e doenças, ciclagem de nutrientes no solo, polinização dentre outras funções.

A estrutura semântica das extensões de metadados é muito semelhante à estrutura do núcleo principal de termos do DwC. Como descrito na Seção 2.6.1, cada extensão é considerada uma classe de metadados, e cada classe possui diversos elementos (propriedades) de áreas temáticas não cobertas pelos metadados do núcleo principal do DwC. As extensões são um complemento aos metadados existentes no DwC.

Para proporcionar a análise das classes de metadados das extensões e facilitar a compreensão do escopo temático de cada classe, foi feita a tradução para o português das definições das classes de metadados disponíveis no repositório da GBIF (2013). Em alguns casos, as definições fornecidas pela GBIF (2013) não foram o suficiente

para a compreensão total do conteúdo temático das extensões, exigindo pesquisas adicionais para o entendimento de determinados termos e expressões e, também, a análise dos elementos que compõem as classes de metadados, cujas definições ajudaram a clarear o entendimento da classe temática. A análise das classes foi guiada pela busca de termos do campo das interações ecológicas, levantados por meio de análise da literatura que teve como resultado a Seção 4.1.2.

De maneira bem semelhante ao que foi discutido na Seção 4.1.3, a análise das classes das extensões resultou em uma síntese de conteúdo, representada pelo Quadro 22 e Quadro 23.

As extensões analisadas nesta etapa dos resultados foram elencadas no Apêndice A e no Apêndice B, que contêm uma descrição básica de todas as extensões criadas para o DwC. As extensões estáveis foram inseridas no Apêndice B e as extensões em desenvolvimento no Apêndice B, ambas representadas em quadros com a estrutura do Quadro 21.

Quadro 21 – Estrutura do Apêndice A e do Apêndice B

TÍTULO	ATRIBUTOS	DEFINIÇÃO
GGBN Amplification Extension	Definição	Suporte para Amplificações de DNA como uma extensão para a classe de dados de amostragem <i>MaterialSample</i> no DwC. Destina-se a ser uma relação do tipo um-para-muitos (1-n) para a classe <i>MaterialSample</i> .
	Disponível em	http://terms.tdwg.org/wiki/GGBN_Amplification_Vocabulary
	Propriedades	34
	Nome da classe	Amplification
	Name space	http://data.ggbn.org/schemas/ggbn/terms/
	RowType	http://data.ggbn.org/schemas/ggbn/terms/Amplification

Fonte: extrato do Apêndice A

A estrutura do Quadro 21 evidencia o modelo de triplas (ver Figura 16) que é o cerne dos esquemas de metadados: o título, que designa o nome da extensão, representa o *sujeito* (ou recurso representado); os descritores representam os *predicados* (propriedades) sobre o sujeito e a definição apresenta os *objetos* (valores) para os predicados.

A análise do Apêndice A gerou uma síntese do escopo temático de cada extensão para facilitar a detecção da presença do tema interações ecológicas nas classes de metadados do Quadro 22.

Quadro 22 – Síntese do conteúdo das extensões estáveis do DwC

(Continua)

NOME	DEFINIÇÃO
Global Genome Biodiversity Network (GGBN) Amplification Extension	Metadados para a descrição de amplificação de DNA ³⁷ ; é um complemento à classe <i>MaterialSample</i> .
GGBN DNA Cloning Extension	Metadados para descrição de Clonagem de DNA como uma extensão para a classe de dados de amostragem <i>MaterialSample</i> no DwC.
GGBN Gel Image Extension	Suporte para propriedades de imagem em gel como uma extensão para a classe de dados de amostragem <i>MaterialSample</i> no DwC.
GGBN Loan Extension	Suporte para propriedades de empréstimo de espécimes, tecidos ou amostras de DNA entre instituições como uma extensão para a classe de dados de amostragem <i>MaterialSample</i> no DwC
GGBN Material Sample Extension	Suporte para propriedades de amostra de material (que pode ser um tecido, DNA ou RNA) como uma extensão para a classe de dados de amostragem <i>MaterialSample</i> no DwC.
GGBN Permit Extension	Suporte para todos os tipos de permissões legais relacionadas a aspectos de aquisição de amostras, empréstimo e uso como uma extensão para a classe de dados de amostragem <i>MaterialSample</i> no DwC.
GGBN Preparation Extension	Suporte para propriedades de preparação do espécime, da amostra de tecido ou extração de DNA como uma extensão para uma classe do tipo <i>MaterialSample</i> .
GGBN Preservation Extension	Suporte para todos os tipos de preservação de amostra em uma coleção física como uma extensão para a classe de dados de amostragem <i>MaterialSample</i> no DwC.
EOL Media Extension 1.0	Esta extensão importa metadados do Audubon Core, Dublin Core e outros para coletar informação sobre texto e multimídia. Ele foi projetado para conter todos os metadados que precisam ser indexados pela Encyclopedia of Life (EOL), mas essa extensão é, com sorte, geral o suficiente para ser útil a todos os provedores e consumidores de texto e mídia.
EOL References Extension 1.0	Essa extensão importa metadados da BIBO (http://bibliontology.com/), Dublin Core e outros para coletar informações sobre citações e referências bibliográficas.
Germplasm accession (v20140515)	O Germplasm Vocabulary fornece um conjunto de termos (complementando os termos do Darwin Core) para descrever os acessos do banco de genes. Esses termos são mantidos pela comunidade temática de Recursos Genéticos de Plantas para Alimentação e Agricultura (PGRFA).
Trait measurement score (v20140515)	Medições de traços ou os chamados dados de Caracterização e Avaliação para recursos genéticos de PGRFA.
Trait descriptor (v20140515)	Metadados que descrevem métodos e protocolos seguidos ao se fazer medições de características ou os chamados dados de Caracterização e Avaliação) para recursos genéticos de PGRFA.
Trait measurement trial (v20140515)	Ensaio de medição em campo ou estufa para coletar características ou os chamados dados de Caracterização e Avaliação para recursos genéticos de PGRFA.

³⁷ A amplificação de DNA consiste em um “Aumento seletivo no número de cópias de um gene codificado por uma proteína específica sem um aumento proporcional nos outros genes” (AMPLIFICAÇÃO..., 2019, *online*).

Quadro 22 – Síntese do conteúdo das extensões estáveis do DwC
(Conclusão)

NOME	DEFINIÇÃO
Taxon Description	Extensão de metadados para a classe Taxon do DwC. Permite atribuir descrições simples de táxons baseados em texto. Não é adequado para descrições com chaves estruturadas, mas é útil para criar páginas de espécies.
Species Distribution	Distribuição geográfica de um táxon.
Alternative Identifiers	Identificadores alternativos para metadados de Táxons, e bases de dados do DwC.
Simple Multimedia	Extensão simples para troca de metadados sobre recursos de multimídia, em particular links para arquivos de imagem, vídeo e áudio.
Literature References	Bibliografia, ou seja, lista de referências bibliográficas
GBIF Relevé	Suporte para medições de levantamento de área de vegetação auxiliares àquelas relatadas usando a classe Event do DwC com extensão de Ocorrência. Observa-se que todas as medidas de cobertura estão em porcentagem.
Species Profile	Um perfil básico de espécies com características além da descrição textual que são cobertas pela extensão Taxon Description.
Types and Specimen	Uma extensão para espécimes e tipos, incluindo espécimes tipo, espécie tipo e gêneros tipo e espécimes simples não relacionados a tipos.
Vernacular Names	Extensão para a classe <i>Taxon</i> que lista nomes vernáculos para um táxon científico.
Extended Measurement Or Facts	Suporte para medições genéricas ou fatos, versão estendida da classe MeasurementOrFacts do DwC ligada a ocorrências.
Audubon Media Description	O Audubon Core é um conjunto de vocabulários criados para representar metadados de recursos multimídia e coleções de biodiversidade.
Darwin Core Identification History	Suporte para múltiplas identificações/determinações de ocorrências de espécies como espécimes.
ChronometricAge	Extensão para a classe Occurrence para capturar informações sobre a idade cronológica.
ChronometricDate	Extensão para a classe Occurrence para capturar informações de datação cronométrica.

Fonte: adaptado do Apêndice A

Dentre as extensões de metadados indicadas no Quadro 22 não há indicação de metadados sobre as interações ecológicas. Entretanto, observa-se que há metadados para representar outra dimensão da agrobiodiversidade que é denominada traços agrônômicos (*agronomic traits*, em inglês). Relacionam-se aos traços agrônômicos às classes que contêm metadados que representam PGRFA no Quadro 22.

A análise do Apêndice B gerou uma síntese do escopo temático de cada extensão do DwC em desenvolvimento, para facilitar a detecção da presença do tema interações ecológicas nas classes de metadados do Quadro 23.

Quadro 23 – Síntese do conteúdo das extensões em desenvolvimento do DwC
(Continua)

NOME	DEFINIÇÃO
MlxS Sample	Termos baseados em amostragem do MlxS ³⁸ estendendo a classe Occurrence do DwC.
Plinian Distribution Extension	Definição geográfica das espécies.
Plinian Endemicity Extension	Um marcador que indica que um organismo vive exclusivamente em um território específico e não pode ser encontrado em nenhum outro lugar. Lista de áreas estruturadas e categorizadas pelo grau de segurança.
Plinian Legislation Extension	Lei proposta a nível nacional ou regional ou grupo de leis.
Plinian ManagementAndConservation Extension	Gestão: ações direcionadas à conservação ou restauração de espécies. Conservação: intervenções realizadas para preservar espécies.
Plinian Core Simple Extension	O Plinian Core é um conjunto de conceitos que define os atributos básicos necessários para integrar e recuperar as informações sobre espécies de organismos requeridas por usuários especializados em biodiversidade, bem como usuários de outras áreas.
Plinian Synonym Extension	Nomes diferentes para este táxon. Esse conceito é um campo de espaço reservado.
Plinian ThreatStatus Extension	Informações sobre o estado de conservação do táxon.
Plinian Uses Extension	Formas em que as espécies são utilizadas pelas pessoas. Incluindo Folclore. Também aplicado a animais e fungos.
CoL Name Relations	Suporte para relações nomenclaturais de nomes a serem usadas em um arquivo DwC da classe Taxon, especificando uma relação direcionada do nome do core para um nome relacionado.
Invasive Species Distribution	Extensão da classe Taxon do DwC para trocar dados de distribuições de áreas de espécies invasoras. Baseia-se na extensão de distribuição do GBIF, mas acrescenta invasividade e usa outros vocabulários.
Invasive Species Pathways	Extensão da classe Taxon do DwC para trocar dados sobre rotas de espécies invasoras.
Taxon Concept Relationship	Uma extensão do DwC-A para especificar as relações entre conceitos de táxon, com base no elemento 'TaxonRelationshipAssertion' do TDWG Taxon Concept Transfer XML Schema. Os tipos de relacionamento incluem tipos de teoria de conjuntos (included_in etc.) e relacionamentos pais-filhos taxonômicos.
Chromosomes Count	Informação sobre o número de cromossomos de uma espécie.

³⁸ Minimum Information about any (x) Sequence (MlxS).

Quadro 23 – Síntese do conteúdo das extensões em desenvolvimento do DwC
(Conclusão)

NOME	DEFINIÇÃO
Ellenbergs Ecological Indicators for European Plants	O nome abreviado de Ellenberg, Pointer Values, para os "Indicadores Ecológicos de Plantas na Europa Central" é um método de classificação para plantas da Europa Central, descrito em detalhes por Heinz Ellenberg em meados dos anos 70, de acordo com seu "comportamento" ecológico e propriedades botânicas. Os <i>pointer values</i> de acordo com Ellenberg são derivados de observações e experiências ecológicas e botânicas e são agora amplamente confirmados por análises de locais e investigações ecofisiológicas ou parâmetros confirmados para espécies de plantas individuais.
Web Links	Links de páginas da Web relacionadas. Os links podem ser classificados com palavras-chave (dc:subject) e devem ser usados para páginas da web tenham interfaces amigáveis.
Extension for mapping National Biodiversity Network (NBN) eXchange Format (NXF) files to DwC-A	Extensão NBN Darwin Core do Reino Unido para mapear arquivos NXF para DwC-A. Esta extensão é, provavelmente, aplicável apenas para a Rede Nacional de Biodiversidade do Reino Unido e seus fornecedores de dados, e contém algumas peculiaridades específicas do Reino Unido.
Ocean Biogeographic Information System (OBIS) Marine BioGeography (MBG) Common Terms Extension	Suporte para propriedades de OBIS MBG como uma extensão para a classe Occurrence do DwC.
OBIS MBG Fish Abundance Extension	Suporte para propriedades de OBIS MBG Fish Abundance como uma extensão para a classe Occurrence do DwC.
OBIS MBG Visual Line Transect Survey Extension	Suporte para as propriedades Visual Line Transect Survey (VLTS) do MBG OBIS como uma extensão para a classe Occurrence do DwC.
Audubon Media Description	Esses termos são metadados dependentes de representação, referentes a representações digitais específicas de um recurso.
Global Invasive Species Information Network (GISIN) Species Status	Não apresenta uma descrição, apenas o status "em desenvolvimento". Ao analisar os metadados que compõem esta classe, entretanto, é possível observar que tem por funcionalidade representar a distribuição de espécies.

Fonte: adaptado do Apêndice B

Assim como as extensões estáveis do DwC, as extensões em desenvolvimento também não apresentam metadados para o campo temático das interações

ecológicas, e, diferentemente daquelas, não apresentam outros conceitos dentro do domínio da agrobiodiversidade. Entretanto, os metadados do núcleo principal de termos e das extensões cobrem boa parte da representação de dados das interações, uma vez que só é possível registrar as interações a partir dos registros de ocorrência. As interações ecológicas representam, na lógica RDF, ligações entre registros de metadados de espécimes que são membros de táxons (espécies) distintos. Portanto, é necessário que haja uma descrição detalhada de uma espécie para que seja possível vinculá-la a outra espécie. Desta forma, conclui-se que o DwC é uma estrutura semântica rica e que pode suportar a representação de informações sobre as interações ecológicas, desde que sejam criadas “ligações”. Essas ligações devem ser descritas de forma clara para que seja possível a identificação do tipo de interação que há entre as duas espécies cujos registros de metadados estão ligados.

A seguir, na Seção 4.1.5 apresentam-se as análises dos termos do modelo <onto.biodiversidade> que teve em vista verificar se a cobertura temática do modelo proposto pela Embrapa é compatível com a estrutura de representação do DwC e se esse modelo comporta termos que possam representar a agrobiodiversidade, com foco nas interações ecológicas.

4.1.5 *Análise terminológica correlata das classes do modelo <onto.biodiversidade>, DwC e conceitos de interações ecológicas*

Estudou-se a estrutura do modelo <onto.biodiversidade>, desenvolvido na Embrapa (ANEXOS A a I), em correlação aos campos do DwC. O <onto.biodiversidade> é um modelo geral de representação da informação, que inclui conceitos de ocorrência de espécies, espécimes, coleção, georreferenciamento e outros.

Foi feita uma correlação entre os termos presentes no padrão DwC e os termos do modelo <onto.biodiversidade> para detectar quais conceitos do modelo já estão representados no DwC e quais não estão.

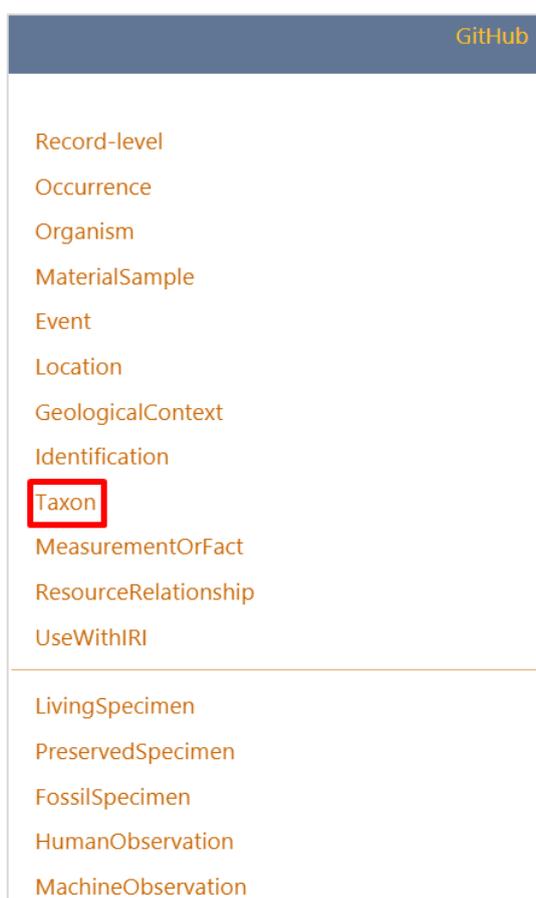
Para analisar as partes do modelo dos Anexos A a I, foram utilizadas as próprias classes desses modelos como categorias gerais de análise, de forma que foi possível identificar o escopo temático de cada classe do modelo <onto.biodiversidade>, aplicando o mesmo procedimento realizado nas etapas anteriores da metodologia, cujos resultados podem ser verificados nas Seções 4.1.2 e 4.1.3.

Posteriormente, foram analisadas as classes do <onto.biodiversidade> que não possuem equivalentes no DwC para verificar se esses conceitos podem representar as interações ecológicas levantadas por meio da análise da literatura, visto que as classes que são comuns entre o <onto.biodiversidade> e o DwC não possuem metadados sobre as interações ecológicas.

A busca por termos equivalentes entre o <onto.biodiversidade> e o DwC foi executada da seguinte maneira:

- a) selecionou-se um termo do modelo <onto.biodiversidade>, como Taxa, que representa uma classe, como pode ser visualizado no Anexo E;
- b) no Darwin Core Quick Reference Guide (WIECZOREK *et al.*, 2009a), que possui a lista de termos do vocabulário do DwC, buscou-se o termo equivalente ao selecionado no modelo <onto.biodiversidade>, primeiramente nas classes gerais do DwC, listadas na Figura 34;

Figura 34 – Classes gerais do DwC



Fonte: Wieczorek *et al.* (2009a)

O resultado da busca foi a localização do termo *taxon* como equivalente para o termo Taxa. Caso o termo selecionado não tivesse sido localizado nas classes do núcleo principal do DwC, a busca deveria continuar pelos termos mais específicos (elementos) subordinados às classes do vocabulário;

- c) o termo equivalente foi buscado nos termos subordinados a cada classe, denominados *property*, conforme o exemplo da Figura 35;

Figura 35 – Termos subordinados às classes do DwC



Fonte: Wiczorek *et al.* (2009a)

- d) identificado o termo equivalente em potencial, foi verificada a sua definição para garantir que são totalmente equivalentes. Na Figura 35 o termo *kingdom* em destaque é um exemplo de termo equivalente em potencial ao termo reino do Anexo E; ao clicar sobre *kingdom*, conforme a Figura 35, o site exibe a ficha terminológica da Figura 36;

Figura 36 – Ficha terminológica do termo *kingdom*

kingdom		Property
Identifier	http://rs.tdwg.org/dwc/terms/kingdom	
Definition	The full scientific name of the kingdom in which the taxon is classified.	
Comments		
Examples	Animalia , Archaea , Bacteria , Chromista , Fungi , Plantae , Protozoa , Viruses	

Fonte: Wieczorek *et al.* (2009a)

A análise dos atributos *Definition*, *Comments* e *Examples* da Figura 36 permite confirmar se o termo equivalente em potencial é um termo equivalente verdadeiro.

Em casos onde o termo não foi localizado no núcleo básico de termos do DwC, a busca continuou pelas extensões. As extensões do DwC apresentadas nos Apêndices A e B estão disponíveis no repositório da GBIF denominado Darwin Core Archive Validator³⁹. Cada extensão possui uma ficha de apresentação, exemplificada na Figura 37.

Figura 37 – Apresentação das extensões do DwC

GGBN Amplification Extension



Support for DNA Amplifications as an extension for Material Sample core relation to the Material Sample core.
See also http://terms.tdwg.org/wiki/GGBN_Amplification_Vocabulary

Properties 34

Name Amplification

Namespace <http://data.ggbn.org/schemas/ggbn/terms/>

RowType <http://data.ggbn.org/schemas/ggbn/terms/Amplification>

Keywords

Fonte: GBIF (2019a)

Ao clicar sobre o nome da extensão, como indicado pela seta vermelha na Figura 37, a navegação é automaticamente redirecionada para uma página contendo os termos (*property*) da extensão com suas respectivas definições, como é possível verificar na Figura 38.

³⁹ Disponível em: <https://tools.gbif.org/dwca-validator/extensions.do>. Acesso em: 12 mar. 2019.

Figura 38 – Termos da extensão

Amplification

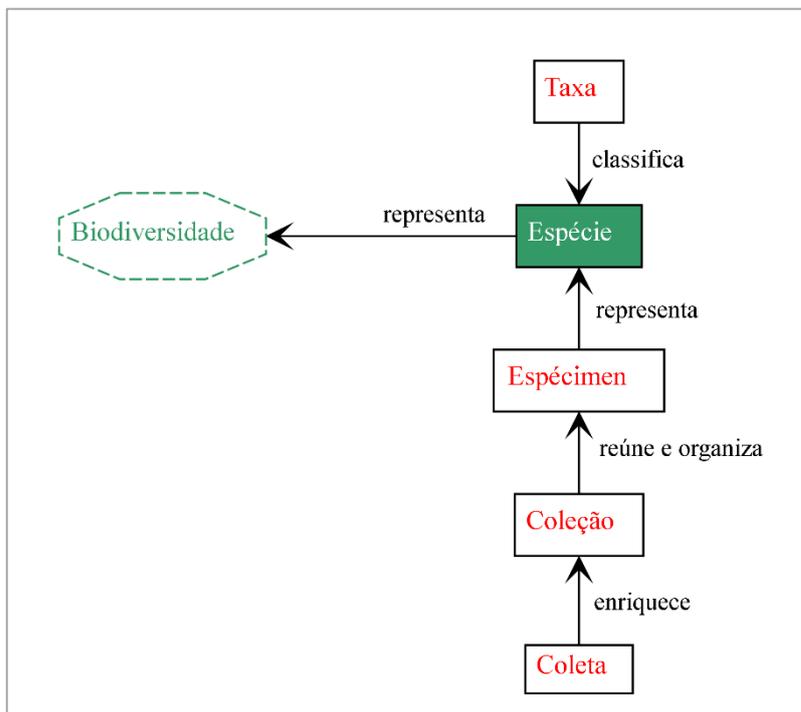
amplificationDate	<p>Date when the amplification was carried out See also http://terms.tdwg.org/wiki/ggbn:amplificationDate</p> <p><i>Examples:</i> 2008-10-17, May 2007</p> <p>Qualified Name http://data.ggbn.org/schemas/ggbn/terms/amplificationDate Namespace http://data.ggbn.org/schemas/ggbn/terms/ Group Amplification Data Type string Required false</p>
amplificationStaff	<p>Person or Institution who performed the amplification See also http://terms.tdwg.org/wiki/ggbn:amplificationStaff</p> <p><i>Examples:</i> Jon Doe, NMNH</p> <p>Qualified Name http://data.ggbn.org/schemas/ggbn/terms/amplificationStaff Namespace http://data.ggbn.org/schemas/ggbn/terms/ Group Amplification Data Type string Required false</p>
amplificationSuccess	<p>true/false or yes/no whether the amplification was successful in general See also http://terms.tdwg.org/wiki/ggbn:amplificationSuccess</p> <p><i>Examples:</i> yes, no, true, false</p> <p>Qualified Name http://data.ggbn.org/schemas/ggbn/terms/amplificationSuccess Namespace http://data.ggbn.org/schemas/ggbn/terms/ Group Amplification Data Type string Required false</p>

Fonte: GBIF (2019b)

A análise da definição do termo e dos *Examples* da Figura 38 permite confirmar se o termo equivalente em potencial é um termo totalmente equivalente.

Na primeira parte da análise desta subetapa, o modelo <onto.biodiversidade> foi explorado em partes, que estão relacionadas nos Anexos A a I. As classes mais gerais do modelo elaborado por Pierozzi Junior (2012) estão representadas na Figura 39.

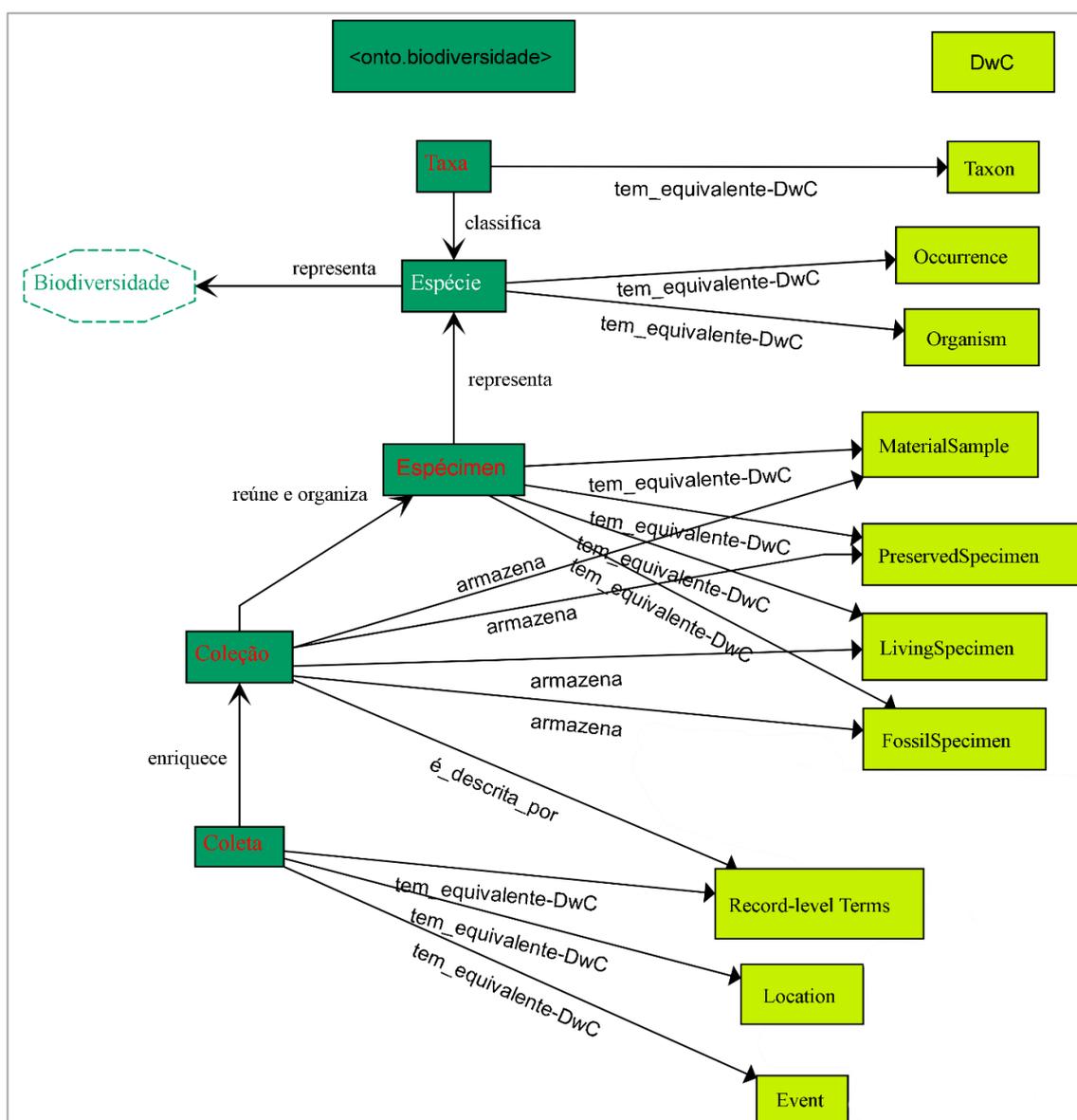
Figura 39 – Classes gerais do modelo <onto.biodiversidade>



Fonte: Pierozzi Junior (2012)

Os termos de modelo da Figura 39, que representam classes, foram analisados de maneira correlacionada ao vocabulário do DwC. A análise permitiu constatar que não há equivalentes para todos os termos propostos como classe no modelo <onto.biodiversidade>, conforme é possível observar na Figura 40. Entretanto, foi possível estabelecer relações entre os conceitos do modelo da Figura 39 com os termos do DwC, mesmo que não haja equivalentes. O modelo da Figura 40 apresenta relacionamentos entre os conceitos como um esquema de entidades e relacionamentos.

Figura 40 – Classes modelo <onto.biodiversidade> e equivalentes no DwC



Fonte: elaborado pelo proponente (2019)

Na Figura 40 considerou-se que o termo *taxa*, no plural, possui como equivalente o termo *taxon* no DwC, no singular. O termo espécie não é adotado como uma classe ou instância do DwC, mas é representado por meio das ligações semânticas entre as classes do DwC. A classe *Organism* (vide Quadro 10), ligada à classe *Taxon* (vide Quadro 16) por meio da classe *Identification* (vide Quadro 15) no DwC representa o conceito de espécie⁴⁰, pois permite identificar o táxon do(s) organismo(s) descrito(s)

⁴⁰ Não confundir conceito de espécie tratado aqui, no contexto da representação da informação, com o conceito biológico de espécie, em torno do qual há grande discussão teórica no campo da Biologia e que não faz parte do escopo desta pesquisa.

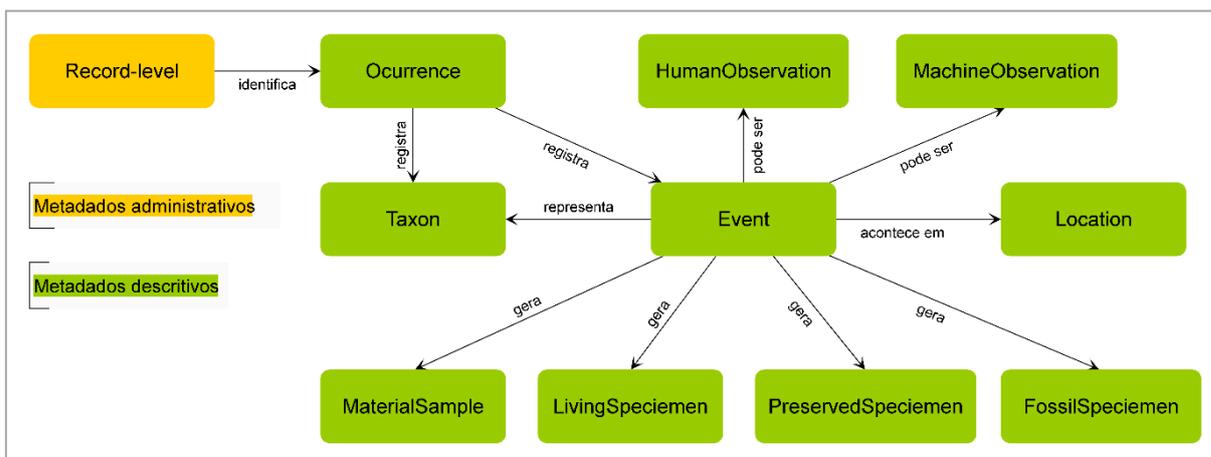
no registro de metadados.

O termo *espécimen* tem como equivalentes os elementos *MaterialSample*, *PreservedSpecimen*, *LivingSpecimen*, *FossilSpecimen*, que representam espécimes e amostras de materiais nos registros do DwC. Neste ponto, o modelo <onto.biodiversidade> é mais lógico ao propor uma classe que agregue todos os tipos de espécimes, o que não acontece no DwC. Nesse caso, o termo espécime do <onto.biodiversidade> se enquadraria no campo *basisOfRecord* (vide Quadro 8) do DwC.

As classes Coleção e Coleta, na Figura 39, não possuem classes equivalentes diretas no DwC, como é possível visualizar na Figura 40. Os conceitos de coleção e coleta do <onto.biodiversidade> estão representados por conceitos de outras classes no DwC. A coleção que armazena o item descrito no registro de ocorrência pode ser identificada com os termos de nível de registro importados do DC, ou seja, os Record-level Terms do Quadro 8. A classe Coleta não possui equivalente direto, mas faz parte de um evento que envolve amostragem. Neste caso, a coleta não poderia ser considerada uma classe, e sim um valor pouco descritivo para a classe *MaterialSample* do Quadro 11 e a ação de coletar itens biológicos seria descrita na classe Evento do Quadro 12, por meio dos metadados que descrevem as técnicas de amostragem (*samplingProtocol*), tamanho da amostra (*sampleSizeValue*), dentre outros.

A Figura 41 é um esquema de classes do DwC que equivale à estrutura conceitual desenvolvida no modelo <onto.biodiversidade>, na Figura 39. O modelo geral de registro de metadados DwC de representação de ocorrência de espécies com interesse para a agropecuária no contexto da Embrapa pode, então, ser representado conforme a Figura 41.

Figura 41 – Classes do modelo de representação para a Empresa



Fonte: elaborado pelo proponente (2019)

As classes da Figura 41 cobrem a representação de dados da biodiversidade em geral, não necessariamente no contexto da agrobiodiversidade. Desta maneira, foi necessário observar termos mais específicos do modelo <onto.biodiversidade> que pudessem refinar a representação dos dados da agrobiodiversidade. No Anexo B, os termos ‘importância econômica’, ‘importância socioeconômica’, ‘importância no sistema natural’ e ‘uso potencial/real’ apontam as relações de uso das espécies pelos seres humanos, o que se encaixa no contexto da agrobiodiversidade. A importância da espécie para o sistema natural representa a função do organismo no ecossistema em uma perspectiva ecológica, de modo que as ações desenvolvidas por aquele organismo são fundamentais para manutenção da estabilidade do ecossistema. Isso se aplica também aos sistemas agroecológicos, que dependem de diversos organismos para a manutenção de serviços como fixação de nitrogênio, polinização, combate de pragas, dentre outros, conforme apontado no referencial teórico-metodológico. As funções ecológicas (que aparecem como uma instância no Anexo D) envolvem as ligações entre os organismos no ecossistema e podem ser representadas por meio de metadados aplicando-se a lógica de relacionamentos do RDF. Entretanto, não existem metadados específicos no DwC ou nas suas extensões para este tipo de representação. O modelo <onto.biodiversidade> não inclui refinamentos para esses termos, apesar de indicá-los. Chegou-se a essa conclusão após a análise da literatura especializada, realizada durante o desenvolvimento da investigação.

Portanto, o pressuposto inicial desta pesquisa de que os modelos

<onto.biodiversidade> apresentavam termos que poderiam ser metadados em potencial para a representação de dados sobre as interações ecológicas na agrobiodiversidade, apresentou resultado negativo, não tendo sido comprovada. A partir deste resultado, o escopo conceitual e temático adotado para as interações ecológicas foi o modelo conceitual elaborado a partir da análise da literatura especializada, apresentado na Figura 33.

Ressalta-se que a compreensão do campo temático da extensão de metadados foi fundamental para se definir o escopo de representação. Desta forma, um dos princípios fundamentais para a criação de uma extensão de metadados é a definição do escopo de representação temática, que deve ser feito antes da etapa de modelagem. A análise do campo temático permitiu detectar que os conceitos das interações ecológicas não estavam representados nos instrumentos terminológicos analisados, o que serve como justificativa para dar continuidade com o estudo nesse campo.

4.2 Segunda etapa: definição terminológica e modelagem dos metadados

Para elaborar as definições para os metadados, adotou-se a definição terminológica funcional, exemplificada pelo Quadro 24.

Quadro 24 – Definição terminológica funcional

higherClassification		Property
Identifier	http://rs.tdwg.org/dwc/terms/higherClassification	
Definition	A list (concatenated and separated) of taxa names terminating at the rank immediately superior to the taxon referenced in the taxon record.	
Comments	Recommended best practice is to separate the values in a list with space vertical bar space (), with terms in order from the highest taxonomic rank to the lowest.	
Examples	Plantae Tracheophyta Magnoliopsida Ranunculales Ranunculaceae Ranunculus , Animalia , Animalia Chordata Vertebrata Mammalia Theria Eutheria Rodentia Hystricognatha Hystricognathi Ctenomyidae Ctenomyini Ctenomys	

Fonte: Wieczorek *et al.* (2009a)

Observa-se no campo *Definition* do Quadro 24 que a definição é o que se espera que a *label higherClassification* tenha como valor atribuído, e os campos *Comments* e *Examples* complementam as orientações de formatação do valor.

Além disso, os termos adotados como termos de metadados devem seguir um esquema de codificação de vocabulário controlado. No campo da agrobiodiversidade, recomenda-se o tesouro Agricultural Information Management Standards (AGROVOC⁴¹), que foi desenvolvido pela FAO.

A extensão de metadados a ser criada com base nos princípios propostos nesta pesquisa deverá combinar as três formas de se estender um padrão de metadados:

- a) criando metadados, utilizando-se termos que não estão presentes no núcleo principal do padrão de metadados;
- b) utilizando-se qualificadores para metadados existentes para que possam representar valores específicos;
- c) esquemas de codificação, como os Códigos Internacionais de Nomenclatura na Seção 2.7 desta dissertação.

Para atender a essas três formas de extensão aplicadas ao escopo desta pesquisa, alguns princípios básicos devem ser seguidos:

- a) estudo aprofundado da temática que se deseja representar por meio de metadados;
- b) incluir termos que não estão representados, previamente, no *core* de metadados do DwC, para que valores mais refinados possam ser representados. Os novos metadados deverão possuir nomes, *labels*, definições e funções diferentes dos metadados preexistentes no *core* do DwC. A definição deve estar bem clara para que não haja dúvida quanto aos valores que podem ser atribuídos aos elementos, eliminando ambiguidade semântica. A análise dos metadados em potencial será feita conforme os princípios do DCAM simplificado. Em outras palavras, deve-se verificar se os conceitos levantados

⁴¹ Disponível em: <http://aims.fao.org/standards/agrovoc/concept-scheme>. Acesso em: 30 set. 2019.

- como potenciais metadados podem formar triplas (sujeito-predicado-objeto);
- c) criar qualificadores para especificar os valores a partir de metadados pré-existentes no *core* no padrão de metadados. O qualificador, que designa um valor específico, deve ser inserido junto ao termo já existente no padrão de metadados. A forma gráfica de representar os qualificadores obedecerá às regras do DwC: não há elementos gráficos separando os termos que compõem o elemento com o qualificador, apenas as iniciais em maiúsculo de cada termo, exceto o primeiro termo. Por exemplo, o qualificador *amplificationDate*, da GGBN Amplification Extension, tem a inicial do segundo termo, *Date*, maiúscula, mas sem ponto final separando os elementos, o que é conhecido como formato *lower CamelCase*). Os qualificadores devem possuir, em suas descrições, os mesmos atributos dos metadados do *core* principal, ou seja, *term name*, *label*, *definition* e outros atributos definidos pelo padrão de metadados. Também aqui a definição dos qualificadores deverá estar bem clara para que não haja dúvida quanto aos valores que podem ser atribuídos aos elementos, eliminando ambiguidade semântica;
 - d) usar esquemas de codificação, que não envolve a criação de novos elementos, uma vez que esses esquemas são instrumentos independentes e não são elementos do DwC: fornecem diretrizes para formatação dos valores a serem atribuídos aos elementos, conforme exemplo de Pomerantz (2015) descrito na Seção 2.3.1. O elemento *date* do DC exemplifica essa aplicação: “Comment: Date may be used to express temporal information at any level of granularity. Recommended best practice is to use an encoding scheme, such as the W3C-DTF profile of ISO 8601” (DCMI USAGE BOARD, 2012, *online*).

Cada extensão presente nos Apêndices A e B está representada no site da GBIF por um registro de metadados, conforme ilustra a Figura 42.

Figura 42 – Registro de uma extensão autorizada

GGBN Amplification Extension	Support for DNA Amplifications as an extension for Material Sample correlation to the Material Sample core. See also http://terms.tdwg.org/wiki/GGBN_Amplification_Vocabulary
Properties	34
Name	Amplification
Namespace	http://data.ggbn.org/schemas/ggbn/terms/
RowType	http://data.ggbn.org/schemas/ggbn/terms/Amplification
Keywords	

Fonte: GBIF (2019b)

Nota-se que o registro da Figura 42 descreve a extensão como um elemento de metadados e seus atributos:

- a) *extension name*: nome que designa a extensão com o criador;
- b) *definition*: definição do escopo de representação dos metadados da extensão;
- c) *see also*: indica uma fonte de informação com discussões, vocabulário ou o histórico de desenvolvimento da extensão;
- d) *properties*: número de propriedades da extensão;
- e) *name*: nome da classe que extensão representa a ser utilizado em um registro de metadados;
- f) *namespace*: os termos do DwC devem ser identificados com um URI. Essas URIs estão agrupadas em coleções denominadas Darwin Core namespaces (DARWIN CORE TASK GROUP, 2009a);
- g) *rowType*: URI atribuído ao termo para identificar a classe de dados representada (DARWIN CORE TASK GROUP, 2009b);
- h) *keywords*: termos que descrevem as categorias de assunto da extensão.

Os atributos *see also*, *namespace* e *rowType* são campos de controle definidos pelo TDWG.

Cada um dos metadados da extensão possui uma ficha terminológica que define seus atributos, conforme a Figura 43.

Figura 43 – Ficha terminológica de metadados das extensões

amplificationDate	Date when the amplification was carried out See also http://terms.tdwg.org/wiki/ggbn:amplificationDate <i>Examples:</i> 2008-10-17, May 2007
	Qualified Name http://data.ggbn.org/schemas/ggbn/terms/amplificationDate Namespace http://data.ggbn.org/schemas/ggbn/terms/ Group Amplification Data Type string Required false

Fonte: GBIF (2019b)

Cada elemento deve ser definido por um grupo de atributos, observados na Figura 43:

- a) *termName*: nome do elemento que pode ser utilizado como nome de campo em um registro;
- b) *definition*: definição do termo;
- c) *see also*: indica uma fonte de informação com discussões, vocabulário ou o histórico do termo;
- d) *qualified name*: URI do termo;
- e) *examples*: exemplos de valores que podem ser atribuídos ao metadado;
- f) *namespace*: os termos do DwC devem ser identificados com um URI. Essas URIs estão agrupadas em coleções denominadas Darwin Core namespaces (DARWIN CORE TASK GROUP, 2009a);
- g) *group*: a classe que arrola o metadado;
- h) *data type*: o tipo de valor (dado) que pode ser inserido no campo;
- i) *required*: uso obrigatório.

Para se criar uma extensão de metadados para as interações ecológicas que seja compatível com o DwC, esses atributos devem ser seguidos para padronizar e formalizar a descrição de cada elemento.

A título de demonstração do que se espera alcançar após a definição terminológica dos metadados e a completa modelagem dos metadados, enuncia-se a seguinte situação: deseja-se descrever as interações entre espécies utilizadas no controle biológico.

A premissa básica do controle biológico é controlar as pragas agrícolas e os insetos transmissores de doenças a partir do uso de seus inimigos naturais, que podem ser outros insetos benéficos, predadores, parasitoides, e microrganismos, como fungos, vírus e bactérias (EMBRAPA, 2019, p. 3).

O controle biológico depende do tipo de interação ecológica que há entre os organismos. As interações ecológicas sempre acontecem entre dois ou mais indivíduos diferentes, da mesma espécie ou não. É possível descrever essas relações de acordo com o princípio de triplas para representar o tipo de interação entre um indivíduo e outro.

Para demonstrar essa relação, o controle biológico de percevejos que atacam a soja foi utilizado como exemplo.

Os percevejos das espécies *Euschistus heros*, *Piezodorus guildinii* e *Nezara viridula* são pragas que atacam a soja. De acordo com Pacheco e Corrêa-Ferreira (2000), espécies parasitoides de ovos são utilizadas em diversos países para controlar as populações de percevejos que atacam a soja. No estudo conduzido pelos autores, que são pesquisadores da Embrapa, foi utilizada a espécie de microvespa *Telenomus podisi* Ashmead.

Os trabalhos foram realizados em duas lavouras de soja, na safra 1996/97, (2500 m² cada), localizadas na Fazenda Experimental da Embrapa Soja, no Distrito da Warta, semeadas com as cultivares BR-48 e BR-37 (Campos 1 e 2, respectivamente), em Londrina, PR, onde nenhuma aplicação de inseticida foi realizada. (PACHECO; CORRÊA-FERREIRA, 2000, p. 297)

A partir do estudo de caso dos autores, infere-se que há uma interação ecológica do tipo parasitismo entre a microvespa e o percevejo-da-soja. Essa relação poderia ser representada em um modelo de triplas, seguindo os princípios do DCAM, conforme a Figura 44.

Figura 44 – Relação de parasitismo no modelo de triplas



Fonte: elaborado pelo proponente (2019)

O predicado *parasite of*, que descreve a relação entre a espécie de vespa *Telenomus*

podisi (sujeito) e o percevejo-da-soja *Euschistus heros* (objeto) na Figura 44, pode ser formalmente representado como um elemento DwC conforme a ficha do Quadro 26⁴².

A classe *Parasitism* do Quadro 25 possui duas propriedades associadas, sendo elas descritas no Quadro 26 e Quadro 27.

Quadro 25 – Classe *Parasitism*

Parasitism		Class
Definition	Negative ecological interaction in which an organism, called a parasite, develops at the expense of another organism, called a host, harming it.	
Comments		
Examples	Tapeworms in the human gut; chaparral dodder in Orange tree.	

Fonte: elaborado pelo proponente (2019)

A definição terminológica (*definition*) foi baseada em Sorci e Garnier (2008). A label *Parasitism* foi importada do AGROVOC (PARASITISM, 2011).

As classes cumprem a função de contextualizar o papel do organismo na interação ecológica, pois um mesmo papel pode representar o favorecimento ou desfavorecimento de um organismo, dependendo do tipo de interação ocorrida. Por exemplo, o predicado *hostOf*, que no Quadro 27 permite atribuir os nomes de organismos parasitários, poderia ser utilizado para descrever a relação mutualística (em que ambos os envolvidos são beneficiados) entre o ser humano e as bactérias benéficas que habitam o intestino. O predicado *hostOf* seria atribuído ao registro de metadados do ser humano e receberia como valores os nomes dessas bactérias benéficas.

⁴² As definições dos Quadro 26, Quadro 27 e Quadro 25 foram escritas em inglês visando a interoperabilidade com o padrão de metadados DwC. É recomendado que os nomes dos termos sempre sejam aplicados em inglês, que é a língua oficial do DwC. No Apêndice D foram inseridas as respectivas traduções dos termos apresentados nos três quadros.

Quadro 26 – Elemento *parasiteOf*

parasiteOf		Property
Definition	A living being that lives at the expense of another living being, harming it.	
Comments	Recommended best practice is to use the full scientific name of the parasitized species, with authorship and date if possible, in accordance with International Nomenclature Codes.	
Examples	<i>Euschistus heros</i> Fabricius, 1794	

Fonte: elaborado pelo proponente (2019)

A definição terminológica (*definition*) do Quadro 26 foi baseada em Sorci e Garnier (2008). A label *parasite* foi importada do AGROVOC (PARASITES, 2011) que apresenta o termo *parasites*, entretanto, o termo foi adotado no singular para atender ao requisito da ISO/IEC 11179-4 (2004).

O atributo *comments* foi definido conforme os modelos de metadados e recomendações de uso dos Códigos Internacionais de Nomenclatura apresentados por Wieczorek *et al.* (2009a).

A definição atribuída à propriedade no Quadro 26 aplica os requisitos básicos da ISO/IEC 11179-4 (2004) para a construção de metadados:

- a) o termo usado para representar a propriedade foi indicado no singular;
- b) a definição declara o que o conceito é, de fato, e não apenas o que não é;
- c) a sentença de definição é descritiva, ou seja, esclarece o significado do termo usado como elemento;
- d) não contém abreviações de difícil entendimento;
- e) não incorpora definições de outros metadados ou conceitos subjacentes.

Estes requisitos determinam a semântica do elemento. Entretanto, encontrou-se limitações no uso dos requisitos e recomendações da ISO/IEC 11179-4 (2004). Dentre os requisitos, não incorporar definições de outros metadados ou conceitos subjacentes é inviável tratando-se de metadados para áreas temáticas específicas. No Quadro 26, por exemplo, observa-se o uso do termo *living being*, que significa ser

vivo em português. Esse termo representa outro conceito das ciências biológicas que não foi definido na ficha terminológica do elemento *parasiteOf*, porém, é indispensável utilizá-lo na definição do elemento *parasiteOf* para apresentar seu significado.

Quanto às recomendações de melhores práticas da ISO/IEC 11179-4 (2004), detectou-se como limitação: os elementos de metadados da classe *parasitism* não podem ser compreendidos sozinhos, pois dependem das relações semânticas com outros elementos da classe para serem compreendidos e aplicados. Tal fato é contrário à recomendação “ser capaz de permanecer sozinho [e ser compreendido sem demais explicações” (ISO; IEC, 2004, p. 3).

Já o comentário atribuído ao elemento indica que é possível inserir dados em formato textual no campo, como *Euschistus heros* (Fabricius, 1794), ou identificadores únicos, como o número de DOI no exemplo do Quadro 26. Desta forma, reproduz a codificação sintática do elemento, de acordo com o que é recomendado pelo DwC (BASKAUF *et al.*, 2015): usar URIs quando possível, de maneira a facilitar o compartilhamento de dados aplicando-se o RDF.

Aplicando-se a lógica do RDF é possível estruturar um registro de metadados de uma ocorrência da espécie *Telenomus podisi* em formato XML, acrescentando-se o elemento *parasiteOf* do Quadro 26.

Observe o registro no Esquema 3.

Esquema 3 – Registro XML de metadados DwC com a propriedade *parasiteOf*

```

1
2 <dwr:DarwinRecordSet>
3   xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
4   xsi:schemaLocation="http://rs.tdwg.org/dwc/dwcrecord/ http://rs.tdwg.org/dwc/xsd/tdwg_dwc_classes.xsd"
5   xmlns:dcterms="http://purl.org/dc/terms/"
6   xmlns:dwc="http://rs.tdwg.org/dwc/terms/"
7   xmlns:dwr="http://rs.tdwg.org/dwc/dwcrecord/">
8   <dcterms:Location>
9     <dwc:locationID>xsi:nil="true"/>
10    <dwc:country>Bolivia</dwc:country>
11    <dwc:countryCode>BO</dwc:countryCode>
12    <dwc:stateProvince>Santa Cruz</dwc:stateProvince>
13    <dwc:locality>Santa Cruz Dept., Bolivia</dwc:locality>
14  </dcterms:Location>
15  <dwc:Occurrence>
16    <dcterms:type>PhysicalObject</dcterms:type>
17    <dcterms:modified>2015-10-29T09:49:22Z</dcterms:modified>
18    <dwc:institutionCode>C.A. Triplehorn Insect Collection, Ohio State University, Columbus, OH (OSUC)</dwc:institutionCode>
19    <dwc:collectionCode>Insects</dwc:collectionCode>
20    <dwc:occurrenceID>urn:lsid:biosci.ohio-state.edu:osuc_occurrences:OSUC__72632</dwc:occurrenceID>
21    <dwc:basisOfRecord>PreservedSpecimen</dwc:basisOfRecord>
22    <dwc:locationID>xsi:nil="true"/>
23  </dwc:Occurrence>
24  <dwc:Identification>
25    <dwc:identificationID>urn:lsid:biosci.o...ences:OSUC__72632</dwc:identificationID>
26    <dwc:identifiedBy>Johnson, N. F. (Norman F.)</dwc:identifiedBy>
27    <dwc:dateIdentified>2003</dwc:dateIdentified>
28    <dwc:identificationQualifier>xsi:nil="true"/>
29    <dwc:occurrenceID>urn:lsid:biosci.o...ences:OSUC__72632</dwc:occurrenceID>
30    <dwc:taxonID>xsi:nil="true"/>
31  </dwc:Identification>
32  <dwc:Taxon>
33    <dwc:taxonID>xsi:nil="true"/>
34    <dwc:scientificName>Ctenomys</dwc:scientificName>
35    <dwc:taxonRank>genus</dwc:taxonRank>
36    <dwc:nomenclaturalCode>ICZN</dwc:nomenclaturalCode>
37    <dwc:genus>Telenomus</dwc:genus>
38  </dwc:Taxon>

```

```

39 <dwc:Identification>
40   <dwc:identificationID>xsi:nil="true"/>
41   <dwc:identifiedBy>Johnson, N. F. (Norman F.)</dwc:identifiedBy>
42   <dwc:dateIdentified>2001-09-14</dwc:dateIdentified>
43   <dwc:occurrenceID>urn:lsid:biosci.ohio-state.edu:osuc_occurrences:OSUC__72632</dwc:occurrenceID>
44   <dwc:taxonID>xsi:nil="true"/>
45 </dwc:Identification>
46 <dwc:Taxon>
47   <dwc:taxonID>xsi:nil="true"/>
48   <dwc:parentNameUsageID>xsi:nil="true"/>
49   <dwc:scientificName>Telenomus podisi</dwc:scientificName>
50   <dwc:scientificNameAuthorship>Ashmead, 1893</dwc:scientificNameAuthorship>
51   <dwc:taxonRank>species</dwc:taxonRank>
52   <dwc:nomenclaturalCode>ICZN</dwc:nomenclaturalCode>
53   <dwc:higherClassification>Animalia; Arthropoda; Insecta; Hymenoptera; Platygastroidea; Scelionidae; Telenomus</dwc:higherClassification>
54   <dwc:kingdom>Animalia</dwc:kingdom>
55   <dwc:phylum>Arthropoda</dwc:phylum>
56   <dwc:class>Insecta</dwc:class>
57   <dwc:order>Hymenoptera</dwc:order>
58   <dwc:family>Scelionidae</dwc:family>
59   <dwc:genus>Telenomus</dwc:genus>
60   <dwc:specificEpithet>podisi</dwc:specificEpithet>
61 </dwc:Taxon>
62 <dwc:ResourceRelationship>
63   <dwc:resourceRelationshipID>xsi:nil="true"/>
64   <dwc:resourceID>https://www.gbif.org/species/1401314</dwc:resourceID>
65   <dwc:relatedResourceID>https://www.gbif.org/species/7632236</dwc:relatedResourceID>
66   <dwc:relationshipOfResource>ecological interaction</dwc:relationshipOfResource>
67   <dwcab:parasitism>
68     <dwcab:parasiteOf>Euschistus heros Fabricius, 1794</dwcab:parasiteOf>
69   </dwcab:parasitism>
70 </dwc:ResourceRelationship>
71 </dwr:DarwinRecordSet>

```

72 Fonte: elaborado pelo proponente (2019)

73 Nota: apresenta o modelo de registro de metadados do Darwin Core Task Group (2009e), dados da ocorrência extraídos da GBIF (JOHNSON; CORA, 2019)
74 e caso de parasitismo como controle biológico descrito por Pacheco e Corrêa-Ferreira (2000).

Para diferenciar os metadados da extensão, que foi nomeada como Darwin Core Agrobiodiversity Extension (DwCAB), dos metadados originários do DwC, foi adotado o código <dwcab> para os metadados da extensão, <dwc> para os metadados do núcleo do DwC e <dwr> para os metadados do Simple DwC.

Para intercalar a propriedade `parasiteOf` no registro de metadados do Esquema 3, utilizou-se a propriedade <dwc:relationshipOfResource> (linha 66), que faz parte o núcleo de termos do DwC, para definir o tipo de recurso relacionado como `ecological interaction`, que é o valor desse campo. O campo <dwc:resourceRelationshipID> (linha 63), não possui valor associado, por isso foi inserido o código `xsi:nil="true"/>` que, de acordo com o Darwin Core Task Group (2009e), representa o valor nulo. Não é recomendado deixar campos vazios num registro de metadados, e, como alternativa, o Darwin Core Task Group (2009e) recomenda excluir o campo do registro de metadados ou usar o código `xsi:nil="true"/>`. Decidiu-se manter o campo e usar o valor nulo para demonstrar sua importância para o registro: é por meio do <dwc:resourceRelationshipID> que se conecta um registro de metadados ao outro. Porém, não existe um ID para esse relacionamento, visto que ele foi criado como demonstração para ser apresentado nesta pesquisa. Em uma situação real de representação, deve-se criar todos os IDs para permitir que os dados sejam ligados.

O par propriedade-valor presente na linha 64 do Esquema 3 representa o identificador único para a espécie que é o sujeito do registro de metadados, ou seja, a espécie *Telenomus podisi* (Ashmead, 1893). Por sua vez, o par propriedade-valor da linha 65 representa o identificador único para a espécie que é o objeto da relação, ou seja, *Euschistus heros* (Fabricius, 1794). Ambos os identificadores foram importados do repositório GBIF.org e encontram-se disponíveis também nas respectivas referências na bibliografia deste trabalho.

O valor *ecological interaction* foi proposto como tipo de recurso e descreve uma relação do tipo <dwcab:parasitism> (linha 67), que possui como propriedade <dwcab:parasiteOf>, que recebeu o valor *Euschistus heros* Fabricius, 1794, que é o nome⁴³ da espécie parasitada pela vespa descrita no registro de metadados. O par propriedade-valor dessa relação é apresentado como <dwcab:parasiteOf>Euschistus

⁴³ O formato de citação do nome segue o código ICZN₂.

heros Fabricius, 1794<dwcab:parasiteOf> (vide linha 68 do Esquema 3).

Ainda é preciso aprofundar a discussão quanto à melhor forma de intercalar os metadados das interações ecológicas no registro de metadados baseado no DwC. Apresentou-se como uma alternativa ao modelo apresentado no Esquema 3, a descrição do organismo que é o objeto na relação por meio da classe táxon, conforme o exemplo do Esquema 4.

Esquema 4 – Alternativa de descrição da classe *Parasitism*

```
<dwcab:parasitism>
  <dwcab:parasiteOf>
    <dwc:Taxon>
      <dwc:scientificName>Euschistus heros</dwc:scientificName>
      <dwc:scientificNameAuthorship>Fabricius, 1794</dwc:scientificNameAuthorship>
      <dwc:taxonRank>species</dwc:taxonRank>
      <dwc:nomenclaturalCode>ICZN</dwc:nomenclaturalCode>
    <dwcab:parasiteOf>
  </dwcab:parasitism>
```

Fonte: elaborado pelo proponente (2019)

No Esquema 4 a classe *Parasitism* utiliza os metadados classe *Taxon* do DwC para descrever a espécie *Euschistus heros* (Fabricius, 1794), que é o objeto na relação, parasitada pela vespa que é o sujeito do registro de metadados do Esquema 3.

Observa-se que o Esquema 4 cumpre com todos requisitos do Quadro 26:

- a) “recommended best practice is to use the full scientific name of the parasitized species”; este requisito é atendido pelos pares propriedade-valor <dwc:scientificName>Euschistus heros</dwc:scientificName> e <dwc:scientificNameAuthorship>Fabricius, 1794</dwc:scientificNameAuthorship>;
- b) “in accordance with International Nomenclature Codes”; este requisito é atendido pelo par propriedade-valor <dwc:nomenclaturalCode>ICZN</dwc:nomenclaturalCode>, que aponta qual código de nomenclatura foi aplicado.

Levando-se em consideração que o objeto do exemplo da Figura 44, ou seja, a espécie *Euschistus heros* (Fabricius, 1794), também pode ser sujeito de um registro de metadados caso a tripla seja invertida, foi criado outro predicado, com o percevejo-da-soja como sujeito, conforme o Quadro 27.

Quadro 27 – Elemento *hostOf*

hostOf		Property
Definition	A living being that provides shelter for another living being.	
Comments	Recommended best practice is to use the full scientific name of the parasitic, commensal, or mutualistic living being that the organism serves as a host for, with authorship and date if possible, in accordance with International Nomenclature Codes.	
Examples	<i>Telenomus podisi</i> Ashmead, 1893	

Fonte: elaborado pelo proponente (2019)

A definição terminológica (*definition*) do Quadro 27 foi baseada em Sorci e Garnier (2008). A *label host* foi importada do AGROVOC (HOSTS, 2011) que apresenta o termo *hosts* no plural, entretanto o termo foi adotado no singular para atender ao requisito da ISO/IEC 11179-4 (2004).

O atributo *comments* foi definido conforme os modelos de metadados e recomendações de uso dos Códigos Internacionais de Nomenclatura apresentados por Wieczorek *et al.* (2009a).

O exemplo do Quadro 27 segue as mesmas regras de especificação sintática e semântica aplicadas no exemplo do Quadro 26. Porém, a natureza deste elemento é mais complexa: há hospedeiros em três tipos de interações ecológicas: parasitismo, comensalismo e mutualismo. Uma mesma espécie pode estar envolvida em uma ou mais interações ao mesmo tempo, portanto é possível que ela assuma o papel de hospedeiro mais de uma vez, simultaneamente, em interações diferentes. É necessário desambiguar o significado do papel de hospedeiro nos diferentes casos.

Uma possível solução para esse problema é a organização dos metadados em classes, conforme foi proposto. O elemento *hostOf*, do Quadro 27, subordinado à classe *Parasitism* do Quadro 25, em um registro de metadados perde a ambiguidade, pois torna-se evidente que se trata de um hospedeiro que abriga um parasito.

Ainda é necessário discutir qual seria a melhor maneira de relacionar a extensão de metadados com o núcleo de termos do DwC. Há dois níveis de representação possíveis: o da espécie e o do espécime. O nível da espécie é mais subjetivo, abstrato; ao passo que o nível do espécime é mais realista e concreto, uma vez que ele é um

objeto físico, passível de manipulação e que demonstrou comportamentos e característica quando foi coletado ou observado em uma ocorrência. Em outras palavras, o espécime é um organismo que pode ser observado, portanto é possível dizer com precisão quais são as características daquele organismo. Por esse motivo, considera-se que a melhor maneira de se apresentar as interações ecológicas é descrevê-las a partir de dados de observação, ou seja, no nível do espécime.

Com base nesse conhecimento, o Esquema 3 relaciona a interação ecológica parasitismo para vincular duas espécies representadas por registros de metadados de seus respectivos espécimes. O espécime que é o sujeito do registro do Esquema 3 foi observado em Santa Cruz, Bolívia, e relacionado ao espécime de um percevejo-da-soja. No Esquema 3 foram demonstrados dados de interação hipotéticos, uma vez que não foram recuperadas informações reais de observação desses espécimes em interação.

Para elucidar esses problemas epistemológicos, para a criação efetiva de uma extensão para a temática, é preciso participar da comunidade de prática que estuda a agrobiodiversidade, tendo em vista discutir a viabilidade de aplicação desses metadados.

4.3 Sistematização dos princípios para construção de uma extensão ao DwC

O principal resultado desta pesquisa foi a sistematização do conjunto de princípios que orientam a construção de uma extensão para o padrão de metadados DwC. Tomou-se o cuidado de classificar os princípios como recomendações ou requisitos.

Etimologicamente, o termo “recomendação” tem origem no latim, com o sentido de norma. Entende-se por recomendações os princípios que compreendem as melhores práticas de criação de metadados, construídas de maneira conjunta e compartilhada, cuja aplicação não é obrigatória. As recomendações assumem o papel de conselho, estabelecendo como base ou uma advertência sobre a importância de se aplicar tal recomendação para que o objetivo seja alcançado de maneira eficiente. Dentro do conjunto de princípios, as recomendações foram elaboradas em conformidade com a literatura especializada estudada (ISO; IEC, 2004; NISO, 2004; NISO, 2007; POMERANTZ, 2015; POWELL *et al.*, 2007; RILEY, 2017; ZENG, 2015; ZENG; QIN, 2008).

Etimologicamente, a palavra “requisito” tem origem do latim, com o sentido de procurado ou buscado.

Já os princípios classificados como requisitos se referem a uma condição necessária e essencial, a uma exigência imprescindível, sem os quais não será possível satisfazer o objetivo ou se alcançar determinado fim. São critérios, aspectos e parâmetros concretos que devem ser obrigatoriamente seguidos. A maioria dos requisitos foram importados da norma ISO/IEC 11179-4 (2004) e da documentação do DwC (DARWIN CORE TASK GROUP, 2009A; DARWIN CORE TASK GROUP, 2009b; WIECZOREK *et al.*, 2012).

Os princípios metodológicos, constituídos por recomendações e requisitos, para a construção de uma extensão de metadados para o padrão DwC, conteúdos que foram discutidos ao longo da seção de resultados, estão sintetizados no Quadro 28.

Quadro 28 – Síntese dos princípios metodológicos para criação de uma extensão de metadados para o DwC

RECOMENDAÇÕES	1	Estudar e analisar a temática sobre os dados que se pretende representar por meio de metadados, para se definir o escopo e a abordagem da representação.
	1.1	Classificar os dados em categorias de análises.
	2	Analisar as 12 classes de elementos do núcleo principal de termos do DwC, para verificar a existência de metadados para a área temática que se deseja representar.
	3	Analisar os núcleos de termos das extensões estáveis e daquelas em desenvolvimento do padrão, para verificar a existência de metadados para o escopo da área temática que se deseja representar, dentro da estrutura classificatória do próprio padrão de metadados.
	4	Buscar ontologias, taxonomias, tesouros, esquemas de metadados ou quaisquer outros modelos conceituais pré-existentes para o campo temático que se pretende representar.
	4.1	Executar a análise correlata entre os termos do modelo conceitual e do DwC, caso seja encontrado algum modelo conceitual para a área temática, classificando os termos do modelo conceitual em duas categorias: termos equivalentes e termos não-equivalentes.
	4.1.1	Tratar os termos não-equivalentes como novos metadados em potencial.
	4.1.2	Adotar os termos equivalentes como metadados DwC.
	5	Adotar princípios e recomendações fundamentados em normas e em modelos presentes na literatura sobre metadados para propor a estrutura sintática e semântica dos elementos extensão de metadados. Sugere-se:

	5.1	a norma a ISO/IEC 11179-4 (2004), que fornece recomendações com definições claras e exemplificadas de melhores práticas na criação de metadados.
	5.2	O DCAMI (POWELL <i>et al.</i> , 2007), como modelo da literatura, por representar a base conceitual dos padrões de metadados existentes.
	6	Adotar as três formas de extensão do DwC, se aplicável: 1) Incluir novos metadados, que deverão possuir nomes, <i>labels</i> , definições e funções diferentes dos metadados preexistentes no <i>core</i> do DwC, conforme os itens 6.1 a 6.2.9; 2) Incluir qualificadores (designa um valor específico, deve ser inserido junto ao termo já existente no padrão de metadados), conforme os itens 6.3 a 6.3.2; 3) usar esquemas de codificação, que fornecem diretrizes para formatação dos valores a serem atribuídos aos elementos, conforme item 6.4.
	6.1	Criar uma ficha terminológica para cada um dos metadados da extensão, definindo seus atributos.
REQUISITOS	6.2	Aplicar os requisitos da norma ISO/IEC 11179-4 (2004), descritos nos itens 6.2.1 a 6.2.6 deste quadro para a criação de novos metadados e recomendações adicionais, do item 6.2.7 ao 6.2.9.
	6.2.1	Indicar o termo que representa o elemento no singular.
	6.2.2	Definir o que o conceito é, de fato, e não apenas o que não é.
	6.2.3	Adotar uma sentença de definição descritiva, ou seja, que esclarece o significado do termo usado como elemento.
	6.2.4	Desconsiderar abreviações de difícil entendimento.
	6.2.5	Não incorporar definições de outros metadados ou conceitos subjacentes.
RECOMENDAÇÕES	6.2.6	Adotar um vocabulário controlado para atribuição dos termos utilizados para descrever os novos metadados.
	6.2.7	Buscar uma definição na literatura científica quando o vocabulário controlado não fornecer definição conceitual subjacente para o termo importado como um elemento de metadados.
	6.2.8	Redigir os termos de metadados seguindo o formato <i>lower CamelCase</i> . Por exemplo: occurrenceRemark, lifeStage, reproductiveCondition.
	6.2.9	Criar uma <i>tag</i> para identificar os metadados da extensão no formato XML (ex. 'dwcab' para Darwin Core Agrobiodiversity Extension).
	6.3	Criar qualificadores a partir de metadados pré-existentes no núcleo de metadados do DwC, acrescidos de termos que tornem a representação mais específica.
	6.3.1	Desconsiderar abreviações de difícil entendimento.
	6.3.2	Redigir os termos de metadados, seguindo o formato <i>lower CamelCase</i> . Por exemplo: organismQuantityType, parentEventID, endDayOfYear.
	6.4	Aplicar os esquemas de codificação prescritos no DwC para a formatação dos dados.

REQUISITOS	7	Adicionar a participação da comunidade de prática na construção da extensão de metadados, para que esses metadados sejam pertinentes às necessidades de representação de dados de quem vai utilizá-los na prática.
	8	Criar um registro de metadados para identificar a extensão de metadados como um todo, incluindo os seguintes atributos (8.1 a 8.7):
	8.1	<i>definition</i> : definição do escopo de representação dos metadados da extensão;
	8.2	<i>see also</i> : indica uma fonte de informação com discussões, vocabulário ou o histórico de desenvolvimento da extensão;
	8.3	<i>properties</i> : número de propriedades da extensão;
	8.4	<i>name</i> : nome da classe que extensão representa a ser utilizado em um registro de metadados;
	8.5	<i>namespace</i> : os termos do DwC devem ser identificados com um URI. Essas URIs estão agrupadas em coleções denominadas Darwin Core namespaces (DARWIN CORE TASK GROUP, 2009a);
	8.6	<i>rowType</i> : URI atribuído ao termo para identificar a classe de dados representada (DARWIN CORE TASK GROUP, 2009b);
	8.7	<i>keywords</i> : termos que descrevem as categorias de assunto da extensão.
	9	Cada elemento da extensão deve ser definido por um grupo de atributos (9.1 a 9.9):
	9.1	<i>termName</i> : nome do elemento que pode ser utilizado como nome de campo em um registro;
	9.2	<i>definition</i> : definição do termo;
	9.3	<i>see also</i> : indica uma fonte de informação com discussões, vocabulário ou o histórico do termo;
	9.4	<i>qualified name</i> : URI do termo;
	9.5	<i>examples</i> : exemplos de valores que podem ser atribuídos ao metadado;
	9.6	<i>namespace</i> : os termos do DwC devem ser identificados com um URI. Essas URIs estão agrupadas em coleções denominadas Darwin Core namespaces (DARWIN CORE TASK GROUP, 2009a);
	9.7	<i>group</i> : a classe que compreende o metadado;
	9.8	<i>data type</i> : o tipo de valor (dado) que pode ser inserido no campo;
	9.9	<i>required</i> : uso obrigatório.

Fonte: resultado de pesquisa (2019)

Acredita-se que os princípios sintetizados no Quadro 28 poderão ser aplicados para construir extensões de metadados para o padrão DwC para a representação de outras situações ou temáticas, uma vez que a sintaxe e a semântica do DwC devem ser mantidas em qualquer contexto de tratamento temático para que os metadados sejam

interoperáveis com o DwC.

Os princípios do Quadro 28 preservam o formato do DwC e apresentam os princípios, com os requisitos e recomendações, para criar metadados de acordo com as especificações desse padrão, tendo sido extraídos da documentação do próprio padrão de metadados e de normas contemporâneas vigentes.

5 CONSIDERAÇÕES FINAIS

A apresentação e discussão dos princípios que fundamentam a criação de metadados revela a natureza essencial da representação da informação, que dentre outras atividades trata de eleger conceitos para representar conteúdos em SRIs. A demanda de organização da informação na Web exige que os metadados adotados nos SRIs sejam precisos em suas definições, para facilitar o entendimento de usuários humanos, e normalizados, para que o compartilhamento automático de dados seja eficiente.

A recuperação, acesso, interoperabilidade e reuso dos dados científicos depende de metadados bem estruturados, sendo relevante estudos que tragam contribuições frente à realidade de elevada produção de dados científicos em meio digital.

O estudo dos metadados direcionado ao campo da agrobiodiversidade permitiu entender esse campo científico e as complexas relações em torno de seus conceitos, assim como as dificuldades em se alcançar certos consensos em nível epistemológico, o que é comum a todas as ciências. Essas dificuldades, que tendem a levantar questionamentos sobre a classificação das interações ecológicas, por exemplo, apresentadas nesta pesquisa, demonstram que o processo de criação de metadados deve acionar a discussão com a comunidade de prática para que se obtenha consenso na representação dos dados. Esse consenso é importante para evitar a criação de metadados com definição ambígua, que é uma das principais causas de erros na aplicação de padrões de metadados para a criação de registros.

O pressuposto inicial desta pesquisa, que buscou verificar se era possível definir o escopo de representação dos dados da agrobiodiversidade, notadamente das interações ecológicas, tendo por base o modelo conceitual <onto.biodiversidade> e, a partir dele, propor princípios para a criação de metadados para a agrobiodiversidade como uma extensão ao padrão de metadados DwC, não se confirmou. Já na aplicação da primeira etapa da metodologia, de exploração dos insumos metodológicos e terminológicos, constatou-se que o escopo dos conceitos do modelo <onto.biodiversidade> é uma representação mais geral e abstrata da temática, sendo que, originalmente, o seu escopo não inclui o campo das interações ecológicas.

Logo, o escopo temático foi definido a partir da análise da literatura especializada, o

que resultou em um modelo conceitual dos tipos de interações ecológicas, em especial para as classes que as representam. Indo além dessas classes, para a criação da extensão de metadados ao DwC é necessário ampliar a discussão em torno dos papéis dos organismos, tendo em vista encontrar os melhores metadados para a representação das interações ecológicas. Para isso, deve-se encontrar meios de realizar a desambiguação, para que o *hostOf* de uma relação parasitária não seja confundido com o *hostOf* de uma relação mutualística, por exemplo. Isso exige uma abordagem mais aprofundada da literatura, pois as publicações científicas são as fontes qualificadas para o entendimento do desenvolvimento dos variados tipos de interações ecológicas na agricultura, e são também fonte de vocabulário especializado. Elas ainda cumprem o papel de justificar a necessidade de curadoria dos dados produzidos em torno dessa temática. Também é preciso incluir a discussão com especialistas de domínio que aplicam o padrão de metadados DwC e/ou trabalham com a gestão de dados da agrobiodiversidade, pois são os usuários reais e/ou potenciais desse instrumento.

Quanto aos objetivos desta pesquisa, considera-se que o primeiro objetivo específico, que tratou de ‘conhecer o padrão DwC e suas extensões para entender como os dados sobre a biodiversidade são representados, tendo em vista estabelecer uma correlação entre esses dados e a estrutura semântica do DwC’ foi atendido. Conforme discutido nas Seções 4.1.3 e 4.1.4, houve uma análise bastante detalhada desses insumos, o que permitiu estabelecer os princípios básicos de formatação dos metadados DwC, ou seja, a sintaxe do padrão de metadados, e seus aspectos semânticos.

O segundo objetivo específico que foi ‘conhecer os princípios do DCMI e das recomendações da ISO/IEC 11179-4 (2004) para subsidiar a determinação dos princípios para criação de metadados para o DwC’ foi atendido. Foram apresentados os princípios que fundamentam a elaboração de metadados, descrevendo recomendações de melhores práticas, requisitos e procedimentos, expostos na Seção 2.4 do referencial teórico-metodológico. Além disso, demonstrou-se exemplos práticos de aplicação desses princípios nos metadados criados como protótipo na Seção 4.2 dos resultados.

O terceiro objetivo específico, que tratou de ‘examinar a literatura sobre interações

ecológicas para compreender os conceitos e sua relação com o campo da agrobiodiversidade, com vistas a verificar a possibilidade de uso do padrão DwC para a descrição de dados dessa temática' foi atendido, tendo a temática sido introduzida na Seção 2.2.1 do referencial teórico-metodológico e os conceitos das interações sistematizados na Figura 33, na Seção 4.1.2 dos resultados. Entretanto, destaca-se certa limitação do estudo realizado no tempo disponível para um mestrado, uma vez que, acredita-se, se faz necessário ampliar a discussão em torno da classificação das interações ecológicas na literatura, assim como das relações que podem ser estabelecidas entre o registro de ocorrência e coleções de espécimes⁴⁴ com os dados produzidos pela pesquisa agropecuária sobre as interações ecológicas no contexto apresentado.

Quanto ao quarto objetivo específico, que tratou de 'contribuir para as práticas agrícolas sustentáveis ao propor princípios que podem ser empregados na representação de informação sobre a agrobiodiversidade, na forma de metadados', acredita-se tê-lo atendido parcialmente. Acredita-se que é ainda necessário dar continuidade à pesquisa, como trabalho futuro, para que se construa um sistema que dê acesso aos dados da agrobiodiversidade, construído com base em metadados que sigam os princípios apresentados nesta pesquisa.

Mediante os resultados obtidos no desenvolvimento dos objetivos específicos, considera-se atendido o objetivo geral da pesquisa, que foi 'propor princípios para a criação de uma extensão de metadados para o padrão DwC com vistas ao atendimento das necessidades de representação da informação sobre interações ecológicas no contexto da agrobiodiversidade brasileira'. Espera-se que os princípios propostos nesta pesquisa possam ser aplicados a outros domínios para a criação de metadados. Em suma, estes princípios tratam de:

- a) compreensão do campo temático dos dados que se deseja representar com os metadados;
- b) análise de padrões de metadados pré-existentes e/ou modelos conceituais aceitos por uma comunidade científica para verificar a existência de metadados pré-existentes para representar o campo temático estudado, que pode resultar

⁴⁴ Que são o domínio de representação temática dos metadados do DwC.

em dois direcionamentos:

- metadados já existentes para o campo temático: neste caso, deve-se estudar o padrão de metadados, entender suas regras de aplicação, e adotá-lo como padrão para criação de registros, caso esse seja estável e já tenha sido aplicado por outras instituições. Neste caso, não é necessário criar uma extensão de metadados;
 - após busca exaustiva, não foram encontrados metadados específicos para o assunto: neste caso, deve-se adotar o padrão de metadados com maior proximidade do campo temático e propor uma extensão de metadados que atenda às necessidades mais específicas de representação não atendidas pelo núcleo de termos do padrão existente;
- c) adotar princípios e recomendações fundamentados em normas e em modelos presentes na literatura sobre metadados para propor a estrutura sintática e semântica da extensão de metadados. Recomenda-se como norma a ISO/IEC 11179-4 (2004) por fornecer recomendações com definições claras e exemplificadas de melhores práticas na criação de metadados e, como modelo da literatura, o DCAM (POWELL *et al.*, 2007), por representar a base conceitual dos padrões de metadados existentes;
- d) adicionar a participação da comunidade de prática na construção da extensão de metadados, para que esses metadados sejam pertinentes às necessidades de representação de dados de quem vai utilizá-los na prática.

Considera-se que a questão de pesquisa apresentada no problema: ‘Quais princípios são necessários adotar para a criação de uma extensão de metadados que siga as especificações sintáticas e semânticas do DwC, de maneira que possa representar os conceitos da agrobiodiversidade brasileira, especificamente das interações ecológicas?’ foi respondida. Isso porque os resultados obtidos forneceram um conjunto de princípios, diretrizes e procedimentos que podem ser adotados como base na construção de uma extensão ao DwC sobre a agrobiodiversidade brasileira. Como restrição desses resultados, destaca-se que não houve a realização de uma aplicação efetiva para a validação final desses princípios, ainda que se tenha feito uma aplicação exemplificadora, em uma situação hipotética.

Acredita-se que a pergunta desta pesquisa pode ser reinterpretada de duas maneiras, que são complementares:

- a) é possível criar metadados para a agrobiodiversidade com os mesmos princípios sintáticos e semânticos do DwC?;
- b) é possível representar as relações da agrobiodiversidade a partir do DwC?

A resposta da primeira pergunta é sim, já que demonstrou-se que é possível aplicar os mesmos princípios sintáticos e semânticos do DwC para criar os novos metadados; a segunda pergunta também tem resposta positiva, entretanto, há uma ressalva a se fazer: o núcleo de termos do DwC tem como escopo a descrição de espécimes, logo, a representação é baseada em objetos físicos. Já as interações ecológicas, acredita-se, relacionam-se com maior afinidade com a descrição da espécie, que é um conceito mais abstrato. Um espécime, que representa uma espécie, pode manifestar certa interação ou um conjunto de interações com outro indivíduo quando coletado ou pode simplesmente não estar em interação com outro indivíduo quando foi coletado; as duas situações são possíveis, mas a ausência de interações na segunda ocasião não determina que aquela espécie não manifeste interações ecológicas.

Como limitação nesta pesquisa, ressalta-se que é necessário ampliar a discussão em torno das recomendações da norma ISO/IEC 11179-4 (2004) para substanciar a etapa da definição terminológica e modelagem dos metadados, questão que se sugere seja retomada em estudo futuro.

Espera-se em pesquisa futura abordar os fundamentos da Teoria do Conceito e da Terminologia como insumos para a fase de definição dos termos de metadados. Mesmo que a definição conceitual dos metadados seja diferente da que é proposta nessas teorias, acredita-se que o aporte sobre a organização do conhecimento apresentado por elas pode fornecer embasamento para solucionar casos de ambiguidade semântica mencionados anteriormente neste trabalho.

Com base na discussão contemplada nesta pesquisa, evidenciou-se um questionamento que deve ser investigado em uma pesquisa futura, que o proponente espera levar adiante: as interações ecológicas devem ser representadas no nível da espécie, o que determina que todo registro de metadados sobre ocorrência de

indivíduos de uma espécie apresente todas as possíveis interações que aquela espécie pode desenvolver, ou as interações devem ser representadas no nível do espécime, aparecendo apenas nos registros de metadados dos espécimes os quais foi possível observar a(s) interação(ões) acontecendo em determinado lugar e tempo?

É preciso compreender, em pesquisa futura, como associar a representação das interações ecológicas com a representação em nível de espécime (indivíduo separado) e espécie (conjunto de indivíduos) no DwC. Espera-se, nessa pesquisa futura, descrever a relação entre uma extensão de metadados para a agrobiodiversidade e o núcleo básico de termos do DwC, fornecendo orientações de como os metadados devem ser organizados para que as interações ecológicas possam ser representadas em um registro DwC.

Outro ponto que se espera alcançar em pesquisa futura, é uma metodologia estável e detalhada quanto aos procedimentos para criação de uma extensão de metadados para o DwC, tendo como recorte as interações ecológicas na agrobiodiversidade, mas que possa ser aplicada a qualquer domínio relacionado ao estudo da biodiversidade.

Para tanto, destaca-se que é necessária uma contínua participação em comunidade de prática sobre a temática. Isso ficou evidente em um primeiro contato com a comunidade, que aconteceu em outubro de 2019, na conferência Biodiversity Next, na Holanda, onde um resultado parcial desta pesquisa foi apresentado e discutido entre os partícipes. Pode-se citar como exemplos de membros da comunidade a Dra. Elizabeth Arnaud, coordenadora do Task Group on Data Fitness for Use in Agrobiodiversity, da GBIF; Dr. Antonio Saraiva, coordenador do Interest Group in Biological Interactions do TDWG; Dra. Cynthia Parr, pesquisadora do United States Department of Agriculture; Dra. Jennifer Hammock, pesquisadora do Smithsonian Institution e coordenadora da Encyclopedia of Life; Dra. Debora Drucker, pesquisadora da Embrapa. Ao final do evento, a proposta de intenção estabelecida foi a de criar um grupo de discussão em torno do tema, de maneira que a futura extensão de metadados a ser criada possa ser efetivamente utilizada pela comunidade de prática ao redor do mundo.

No campo da agrobiodiversidade, considera-se que a contribuição dada por esta pesquisa foi iniciar a discussão em torno dos conceitos das interações ecológicas nos

sistemas agroecológicos, e a possibilidade de representar essas interações por meio de metadados. Trazendo um olhar do profissional bibliotecário, a função dos metadados é organizar dados e informações em SRIs e, como tal, acredita-se que um repositório de dados científicos sobre a agrobiodiversidade pode ser ferramenta estratégica para o fomento de práticas agrícolas sustentáveis, em que há alocação correta de recursos.

A contribuição de cunho social se revela uma vez que disponibilizar esses dados em um repositório pode trazer elementos para práticas sustentáveis, ao dar acesso a conhecimento sobre formas de solucionar problemas na agricultura com menor impacto nos ecossistemas. A Agenda 2030 das Nações Unidas determina como uma de suas metas a implementação de práticas agrícolas sustentáveis para o bem do planeta. Para que isso seja viável, aliado ao desenvolvimento de um SRI, é necessário conciliar a divulgação do mesmo junto à sua comunidade de uso real e potencial, assim como realizar testes de usabilidade, para garantir que o recurso criado seja de fato aproveitado por aqueles a quem se destina e que dele necessite. Esta é uma das características seminais da Ciência da Informação: explorar a relação entre o usuário e a informação.

Como contribuição científica no campo da CI, considera-se que a maior contribuição desta pesquisa é a proposta de princípios, compostos de recomendações e requisitos para a criação de uma extensão de metadados, acolhendo à lacuna verificada, cujo desenvolvimento será continuado em pesquisa futura, no mesmo escopo da agrobiodiversidade, para atender a uma demanda nacional e internacional de curadoria dos dados que estão sendo produzidos neste campo. Acredita-se que este trabalho apresenta avanços para o desenvolvimento de uma metodologia para criação de uma extensão de metadados, e os metadados são parte fundamental do campo de estudos da Representação da Informação. Não foi recuperado na literatura qualquer proposta metodológica para a criação de metadados, portanto, acredita-se que os princípios metodológicos apresentados como resultado desta pesquisa são o primeiro passo para a criação de uma metodologia para criação de esquemas de metadados.

REFERÊNCIAS

- ALENCAR, A. *et al.* *Desmatamento na Amazônia: indo além da “emergência crônica”*. Belém: IPAM-Belém, 2004. Não paginado. Disponível em: <https://ipam.org.br/bibliotecas/desmatamento-na-amazonia-indo-alem-da-emergencia-cronica/>. Acesso em: 19 set. 2019.
- ALVES, N. O. *et al.* Biomass burning in the Amazon region: aerosol source apportionment and associated health risk assessment. *Atmospheric Environment*, [Oxford, GB], v. 20, p. 277-285, Nov. 2015. DOI: <https://doi.org/10.1016/j.atmosenv.2015.08.059>. Available from: <https://www.sciencedirect.com/science/article/pii/S1352231015303046>. Access on: 20 Sept. 2019.
- AMPLIFICAÇÃO de genes. *In: DeCS: Descritores em Ciências da Saúde*. São Paulo: Biblioteca Virtual em Saúde, [2019]. Available from: <http://decs.bvs.br/cgi-bin/wxis1660.exe/decsserver/#2>. Access: 7 out. 2019.
- ARNAUD, E. *et al.* *Final Report of the Task Group on GBIF Data Fitness for Use in Agrobiodiversity*. Copenhagen: GBIF, 2016. Available from: https://assets.ctfassets.net/uo17ejk9rkwj/5R4l2VTdy8AqOWUkgyeue/54187bfa12be6bd91aaba4957dac8888/Final-report-of-the-task-group-on-GBIF-data-fitness-for-use-in-agrobiodiversity_1.pdf. Access on: 1 Oct. 2019.
- BACA, M. (ed.). *Introduction to metadata*. 3rd ed. Los Angeles: Getty P, 2016.
- BAKER, T. *et al.* *Dublin Core Application Profile Guidelines*. [Maryland]: DCMI, 2005. Available from: <https://www.dublincore.org/specifications/dublin-core/application-profile-guidelines/>. Access on: 23 Sept. 2019.
- BALVANERA, P. *et al.* Status and trends: indirect and direct drivers of change. *In: DÍAZ, S. et al.* (ed.). *Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services*. Bonn: IPBES secretariat, 2019. Cap. 2, p. Available from: https://www.ipbes.net/system/tdf/ipbes_global_assessment_chapter_2_1_drivers_unedited_31may.pdf?file=1&type=node&id=35278. Access on: 5 Nov. 2019.
- BASKAUF, S. J. *et al.* *Darwin Core RDF guide*. [San Francisco]: TDWG, 2015. Available from: <https://dwc.tdwg.org/rdf/#25-terms-in-the-dwciri-namespace-normative>. Access on: 4 Nov. 2019.
- BASKAUF, S. J.; WEBB, C. O. Darwin-SW: Darwin Core-based terms for expressing biodiversity data as RDF. *Semantic Web Journal*, [Amsterdam], v. 7, n. 6, p. 629-643, 2016. Available from: <http://www.semantic-web-journal.net/content/darwin-sw-darwin-core-based-terms-expressing-biodiversity-data-rdf-0>. Access on: 5 Nov. 2019.
- BEGON, M.; TOWNSEND, C. R.; HARPER, J. L. *Ecologia: de indivíduos a ecossistemas*. 4. ed. Tradução Adriano Sanches Melo *et al.* Porto Alegre: Artmed, 2007. 752 p.

BERNERS-LEE, T. *Linked Data*. [S. l.], 27 July 2006. Available from: <https://www.w3.org/DesignIssues/LinkedData>. Access on: 13 Nov. 2019.

BIODIVERSITY INFORMATION STANDARDS (TDWG). *Darwin Core Material Sample*. [San Francisco]: TDWG Terms Wiki, 2015a. Available from: https://terms.tdwg.org/wiki/Darwin_Core_Material_Sample. Access on: 30 Sept. 2019.

BIODIVERSITY INFORMATION STANDARDS (TDWG). *Darwin Core Record-level Terms*. [San Francisco]: TDWG Terms Wiki, 2015b. Available from: https://terms.tdwg.org/wiki/Darwin_Core_Record-level_Terms. Access on: 20 Sept. 2019.

BIOLOGICAL interaction. *In*: AGROVOC. [New York]: FAO, 2011. Available from: http://aims.fao.org/aos/agrovoc/c_49896.html. Access on: 20 Sept. 2019.

BLUME, E.; REINIGER, L. *Fundamentos da agroecologia*. Santa Maria, RS: UFSM: NTE: UAB, 2007. 87 p. Disponível em: https://repositorio.ufsm.br/bitstream/handle/1/16153/Curso_Agric-Famil-Sust_Fundam-Agroecologia.pdf?sequence=1&isAllowed=y. Acesso em: 20 set. 2019.

BOUCHER, D. H.; JAMES, S.; KEELER, K. H. The ecology of mutualism. *Annual Review of Ecology and Systematics*, [Palo Alto], v. 13, p. 315-347, Nov. 1982. DOI: <https://doi.org/10.1146/annurev.es.13.110182.001531>. Available from: <https://www.annualreviews.org/doi/10.1146/annurev.es.13.110182.001531>. Access on: 20 Sept. 2019.

BRASIL. Decreto nº 4.339, de 22 de agosto de 2002. Institui princípios e diretrizes para a implementação da Política Nacional da Biodiversidade. *Diário Oficial da União*, Brasília, DF, 23 ago. 2002. Seção 1, p. 2. Disponível em: <https://www.lexml.gov.br/urn/urn:lex:br:federal:decreto:2002-08-22;4339>. Acesso em: 8 ago. 2018.

BRASIL. Lei nº 13.123, de 20 de maio de 2015. Regulamenta o inciso II do § 1º e o § 4º do art. 225 da Constituição Federal, o Artigo 1, a alínea j do Artigo 8, a alínea c do Artigo 10, o Artigo 15 e os §§ 3º e 4º do Artigo 16 da Convenção sobre Diversidade Biológica, promulgada pelo Decreto nº 2.519, de 16 de março de 1998; dispõe sobre o acesso ao patrimônio genético, sobre a proteção e o acesso ao conhecimento tradicional associado e sobre a repartição de benefícios para conservação e uso sustentável da biodiversidade; revoga a Medida Provisória nº 2.186-16, de 23 de agosto de 2001; e dá outras providências. *Diário Oficial da União*, Brasília, DF, 21 maio 2015. Seção 1, p. 1. Disponível em: <http://www.lexml.gov.br/urn/urn:lex:br:federal:lei:2015-05-20;13123>. Acesso em: 8 ago. 2018.

BRAY, T. *et al. Extensible Markup Language (XML) 1.0*. 5th ed. [S. l.]: W3C, 2008. Available from: <https://www.w3.org/TR/REC-xml/#sec-intro>. Access on: 23 Apr. 2019.

BRICKELL, C. D. *et al* (ed.). *International Code of Nomenclature for Cultivated Plants (ICNCP or Cultivated Plant Code)*: incorporating the Rules and Recommendations for naming plants in cultivation. 8th ed. Vienna: ISHS, 2009. (Regnum Vegetabile, 151).

BRICKLEY, D.; GUHA, R. V.; MCBRIDE, B. *RDF Schema 1.1: W3C Recommendation* 25 February 2014. W3C, [S. l.], 25 Feb. 2014. Available from: <https://www.w3.org/TR/rdf-schema/#bib-RDF11-CONCEPTS>. Access on: 7 fev. 2019.

BRONSTEIN, J. L. (ed.). *Mutualism*. New York: Oxford University Press, 2015.

CAIN, M. L.; BOWMAN, W. D.; HACKER, S. D. *Ecologia*. 3. ed. Tradução Armando Molina Divan Junior *et al*. Porto Alegre: Artmed, 2018. 720 p. Revisão técnica de Fernando Joner e Paulo Luiz de Oliveira.

CANDOLLE, A. *Lois de la nomenclature Botanique*. In: INTERNATIONAL BOTANICAL CONGRESS, 1867, Paris. *Actes du Congrès international de botanique tenu a Paris en août 1867, sous les auspices de la Société botanique de France*. Paris: V. Masson et fils, 1867. Disponible à: <https://www.biodiversitylibrary.org/item/235808#page/3/mode/1up>. Consulté à: 23 nov. 2019.

CANHOS, V. P. Informática para biodiversidade: padrões, protocolos e ferramentas. *Ciência e Cultura*, São Paulo, v. 55, n. 2, p. 45-47, abr. 2003. Disponível em: http://cienciaecultura.bvs.br/scielo.php?script=sci_arttext&pid=S0009-67252003000200025&lng=en&nrm=iso. Acesso em: 9 maio 2018.

CAPECHE, C. L. Impactos das queimadas na qualidade do solo: degradação ambiental e manejo e conservação do solo e água. In: ENCONTRO CIENTÍFICO DO PARQUE ESTADUAL DOS TRÊS PICOS, 2., 2012, Cachoeiras de Macacu. *Anais [...]*. Cachoeiras de Macacu: Parque Estadual dos Três Picos, 2012. p. 17-20. Disponível em: <https://ainfo.cnptia.embrapa.br/digital/bitstream/item/82108/1/II-Encontro-Cientifico.pdf>. Acesso em: 20 set. 2019.

CASSINI, S. T. *Ecologia: conceitos fundamentais*. Vitória: PPGEA-UFES, 2005. 69 p. Disponível em: https://www.inf.ufes.br/~neyval/Gestao_ambiental/Tecnologias_Ambientais2005/Ecologia/CONC_BASICOS ECOLOGIA_V1.pdf. Acesso em: 29 abr. 2019.

COMPUTER HOPE. CamelCase. In: COMPUTER terms, dictionary, and glossary. [S. l.], 7 dez. 2017. Available from: <https://www.computerhope.com/jargon/c/camelcas.htm>. Access on: 9 May 2019.

CONVENTION ON BIOLOGICAL DIVERSITY. [Report]. Rio de Janeiro: United Nations, 1992. Available from: <https://www.cbd.int/doc/legal/cbd-en.pdf>. Access on: 19 Sept. 2019.

CONVENTION ON BIOLOGICAL DIVERSITY. *COP 5 Decisions: Agricultural biological diversity: Review of phase I of the programme of work and adoption of a multi-year work programme: Annex: Programme of work on agricultural biodiversity*. Nairobi, 2000. Available from: <https://www.cbd.int/doc/decisions/COP-05-dec-en.pdf>. Access on: 9 April 2019.

CONVENTION ON BIOLOGICAL DIVERSITY. Subsidiary Body on Scientific, Technical and Technological Advice. *Report on how to improve sustainable use of biodiversity in a landscape perspective*. Montreal, 2011. Available from: <https://www.cbd.int/doc/meetings/sbstta/sbstta-15/official/sbstta-15-13-en.pdf>. Access on: 20 Sept. 2019.

COYLE, K. *DCAM Explained*. [S. l.], 2008. Available from: http://kcoyle.net/dcam_simple.html. Access on: 25 Sept. 2019.

CRUZ, J. C.; PEREIRA FILHO, I. A.; ALBUQUERQUE FILHO, M. R. Árvore do conhecimento: milho: espaçamento e densidade. *Agência Embrapa de Informação Tecnológica*, Brasília, DF, [3 mar. 2011]. Disponível em: https://www.agencia.cnptia.embrapa.br/gestor/milho/arvore/CONTAG01_49_168200511159.html. Acesso em: 13 nov. 2019.

Cryptolaemus montrouzieri (Mulsant, 1853). In: GBIF SECRETARIAT. *GBIF Backbone Taxonomy*. Copenhagen: GBIF.org, 2019. DOI: <https://doi.org/10.15468/39omei>. Available from: <https://www.gbif.org/dataset/d7dddbf4-2cf0-4f39-9b2a-bb099caae36c>. Access on: 13 Nov. 2019.

DARWIN CORE MAINTENANCE GROUP. *Darwin Core*. [San Fransisco]: TDWG, 2014. Available from: <https://doi.org/10.5281/zenodo.592792>. Access on: 20 Sept. 2019.

DARWIN CORE TASK GROUP. *Darwin Core namespace policy*. [San Fransisco]: TDWG, 2009a. Available from: <http://rs.tdwg.org/dwc/terms/namespace/>. Access on: 9 Apr. 2019.

DARWIN CORE TASK GROUP. *Darwin Core text guide*. [San Fransisco]: TDWG, 2009b. Available from: <http://rs.tdwg.org/dwc/terms/guides/text/>. Access on: 9 Apr. 2019.

DARWIN CORE TASK GROUP. *Darwin Core*. [San Fransisco]: TDWG, [2009c]. Available from: <http://rs.tdwg.org/dwc/>. Access on: 8 Sept. 2018.

DARWIN CORE TASK GROUP. *Simple Darwin Core*. [San Fransisco]: TDWG, 2009d. Available from: <http://rs.tdwg.org/dwc/terms/simple/>. Access on: 22 Jan. 2018.

DARWIN CORE TASK GROUP. *Darwin Core XML guide*. [San Fransisco]: TDWG, 2009e. Available from: <http://rs.tdwg.org/dwc/terms/guides/xml/>. Access on: 20 Sept. 2019.

DCMI USAGE BOARD. *DCMI Metadata Terms*. [Maryland]: DCMI, 2012. Available from: <https://www.dublincore.org/specifications/dublin-core/dcmi-terms/#terms-abstract>. Access on: 23 Sept. 2019.

DESCUBRA ideias sobre ensino de biologia. Pinterest, [2019?]. Disponível em: <https://br.pinterest.com/pin/226094843771732447/>. Acesso em: 2 jan. 2019.

DIGITAL CURATION CENTRE. *Biodiversity*. Edinburgh: DCC, 2018. Available from: <http://www.dcc.ac.uk/resources/metadata-standards/disciplinary/biodiversity>. Access on: 3 Aug. 2018.

DOUBILET, D. A diver explores corals on the Great Barrier Reef off Australia, the largest living structure on Earth (photo). In: PARKER, L. Coral reefs could be gone in 30 years. *National Geographic [News]*, 26 June 2017. Available from: <https://www.nationalgeographic.com.au/australia/coral-reefs-could-be-gone-in-30-years.aspx>. Access on: 19 Sept. 2019.

DUBLIN CORE METADATA INITIATIVE. *Dublin Core Metadata Element Set, Version 1.1: Reference Description*. [Maryland], 2012. Available from: <http://dublincore.org/documents/dces/>. Access on: 19 ago. 2018.

DUBOIS, A. *The International Code of Zoological Nomenclature must be drastically improved before it is too late*. Auckland: Magnolia, 2011. 104 p. (Bionomina, 2).

DÜRST, M. *Internationalized Resource Identifiers (IRIs)*. [S. l.]: W3C Internationalization, 7 Mar. 2011. Available from: <https://www.w3.org/International/O-URL-and-ident.html>. Access on: 7 Out. 2019.

DUVAL, E. *et al.* (2002). Metadata principles and practicalities. *D-Lib Magazine*, [Reston], v. 8, n. 4, [p. 1-15], Apr. 2002. Available from: <http://www.dlib.org/dlib/april02/weibel/04weibel.html>. Access on: 23 Sept. 2019.

EARTH SUMMIT+5. 1., 1997, New York. *Special session of the General Assembly to Review and Appraise the Implementation of Agenda 21*. New York: United Nations, 1997. Available from: <https://www.un.org/esa/earthsummit/>. Access on: 19 Sept. 2019.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA. *Controle biológico*. Portal Embrapa. Versão 3.59.4. 2019. Disponível em: <https://www.embrapa.br/tema-controle-biologico>. Acesso em: 29 abr. 2019.

ESTUDO alerta para o potencial invasor do eucalipto. *Mundo Português*, Lisboa, 18 mar. 2019. Disponível em: <https://www.mundoportugues.pt/estudo-alerta-para-o-potencial-invasor-do-eucalipto/>. Acesso em: 18 set. 2019.

Euschistus heros (Fabricius, 1794). In: GBIF SECRETARIAT. *GBIF Backbone Taxonomy*. Copenhagen: GBIF.org, 2019. DOI: 10.15468/39omei. Available from: <https://www.gbif.org/species/7632236>. Access on: 4 nov. 2019.

FEGRAUS, E. H. *et al.* Maximizing the value of ecological data with structured metadata: an introduction to Ecological Metadata Language (EML) and principles for metadata creation. *Bulletin Ecological Society of America*, [Washington, DC], v. 86, n. 3, p. 158-168, July 2005. Available from: <https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/0012-9623%282005%2986%5B158%3AMTVOED%5D2.0.CO%3B2>. Access on: 25 Sept. 2019.

FERREIRA, E. *Orquídeas nas Superquadras*. 1 fotografia, color., Disponível em: <http://www.ambienciabrasilia.com.br/2015/02/>. Acesso em: 20 set. 2019.

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS. *Save and grow: a policymaker's guide to the sustainable intensification of smallholder crop production*. Rome, 2011. 102 p. Available from: <http://www.fao.org/3/a-i2215e.pdf>. Access on: 20 Sept. 2019.

GERHARDT, T. E.; SILVEIRA, D. T. *Métodos de pesquisa*. Porto Alegre: Ed. UFRGS, 2009. 120 p. (Educação à Distância).

GIL, A.C. *Métodos e técnicas de pesquisa social*. 4 ed. São Paulo: Atlas, 1994. 207 p.

GELAMBI, M. Neutralismo (relación biológica): teorías y ejemplos. *In: lifeder.com*, 2019. Disponible en: <https://www.lifeder.com/neutralismo/>. Acceso em: 25 nov. 2019.

GILLILAND, A. J. Setting the stage. *In: BACA, M. (ed.). Introduction to metadata*. 3rd ed. Los Angeles: Getty P, 2016.

GLOBAL BIODIVERSITY INFORMATION FACILITY. *Darwin Core Archive Validator*. [Copenhagen], [2019a]. Available from: <https://tools.gbif.org/dwca-validator/extensions.do>. Access on: 20 Apr. 2019.

GLOBAL BIODIVERSITY INFORMATION FACILITY. *Data quality requirements: occurrence-only datasets*. Copenhagen: GBIF Secretariat, 2018. Available from: <https://www.cbd.int/doc/legal/cbd-en.pdf>. Access on: 20 Sept. 2019.

GLOBAL BIODIVERSITY INFORMATION FACILITY. *GGBN Amplification Extension*. *In: DARWIN Core Archive Validator*. [Copenhagen], [2019b]. Available from: <https://tools.gbif.org/dwca-validator/extension.do?id=http://data.ggbn.org/schemas/ggbn/terms/Amplification>. Access on: 20 Sept. 2019.

GLOBAL BIODIVERSITY INFORMATION FACILITY. *Data standards*. Copenhagen: GBIF Secretariat, [2019c]. Available from: <https://www.gbif.org/pt/standards>. Access on: 20 Sept. 2019.

GLOBAL BIODIVERSITY INFORMATION FACILITY. *Registered Extensions*. *In: DARWIN Core Archive Validator*. [Copenhagen], 2013. Available from: <https://tools.gbif.org/dwca-validator/extensions.do>. Access on: 31 Jan. 2019.

GONÇALVES, G. S. *et al.* Controle de erva-de-passarinho em citros com herbicidas. *In: CONGRESSO BRASILEIRO DA CIÊNCIA DAS PLANTAS DANINHAS*, 28., 2012, Campo Grande, MS. *A ciência das plantas daninhas na era da biotecnologia*. Campo Grande, MS: Sociedade Brasileira da Ciência das Plantas Daninhas, 2012. p. 193-197. Disponível em: <http://www.alice.cnptia.embrapa.br/handle/doc/949311>. Acesso em: 20 set. 2019.

GOVAERTS R. (ed.). WCSP: World Checklist of Selected Plant Families (version Aug 2017). *In: ROSKOV et al. (ed.). Species 2000 & ITIS Catalogue of Life: 2019 Annual Checklist*. Leiden: Species 2000, 2019. Available from: www.catalogueoflife.org/annual-checklist/2019. Access on: 19 Sept. 2019.

GUIMARÃES, C. *Introdução a linguagens de marcação: HTML, XHTML, SGML, XML*. Campinas: Instituto de Computação da Unicamp, 2005. Disponível em: <http://www.ic.unicamp.br/~celio/inf533/docs/markup.html>. Acesso em: 20 set. 2019.

GUIMARÃES, D. T. *Dicionário técnico jurídico*. 2. ed. rev. e atual. São Paulo: Rideel, 1999.

GUIMARÃES, P. P. *et al.* Análise dos impactos ambientais de um incêndio florestal. *Agrarian Academy*, Goiânia, v. 1, n. 1, p. 38-60, abr. 2014. Disponível em: <http://www.conhecer.org.br/Agrarian%20Academy/2014a/analise%20dos%20impactos.pdf>. Acesso em: 20 set. 2019.

HARLAN, J. R.; DE WET, J. M. J. Toward a rational classification of cultivated plants. *Taxon*, [Utrecht], v. 20, n. 4, p. 509-517, Aug. 1971.

HESSE, H. Perda de biodiversidade, conservação de biodiversidade... mas o que é biodiversidade?. In: BENSUSAN, N. (org.). *Seria melhor mandar ladrilhar?: biodiversidade como, para que, por quê*. 2. ed. rev. e ampl. São Paulo: Editora Petrópolis, 2008. p. 23-25.

HETTERSCHEID, W. L. A.; VAN DEN BERG, R. G.; BRANDENBURG, W. A. An annotated history of the principles of cultivated plant classification. *Acta Botanica Neerlandica*, [Amsterdam], v. 45, n. 2, p. 123-134, June 1996.

HILLMANN, D. *Using Dublin Core*. [Maryland]: DCMI, 2005. Available from: <https://www.dublincore.org/specifications/dublin-core/usageguide/2005-11-07/>. Access on: 23 Sept. 2019.

HOLLAND, J. N.; BRONSTEIN, J. L. Mutualim. In: JØRGENSEN, S. E.; FATH, B. F. *Encyclopedia of Ecology*. Amsterdam: Elsevier, 2008. p. 2485-2491. DOI: 10.1016/B978-008045405-4.00673-X. Available from: <https://arizona.pure.elsevier.com/en/publications/mutualism>. Access on: 5 Nov. 2019.

Hoplias intermedius (Günther, 1864). In: GBIF SECRETARIAT. *GBIF Backbone Taxonomy*. Copenhagen: GBIF.org, 2019. DOI: <https://doi.org/10.15468/39omei>. Available from: <https://www.gbif.org/dataset/d7dddb4-2cf0-4f39-9b2a-bb099caae36c>. Access on: 5 Nov. 2019.

HOSTS. In: AGROVOC Multilingual Thesaurus. [S. l.]: FAO, 2011. Available from: http://aims.fao.org/aos/agrovoc/c_3673. Access on: 24 Nov. 2019.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA. Gerenciar metadados. *Ibict WIKI*, [Brasília, DF], 18 set. 2013. Disponível em: http://wiki.ibict.br/index.php/Gerenciar_metadados. Acesso em: 10 out. 2018.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA. Nome DOI. *Ibict Wiki*, [Brasília, DF], 5 maio 2015. Disponível em: http://wiki.ibict.br/index.php/Nome_Doi. Acesso em: 25 set. 2019.

INSTITUTO BRASILEIRO DO MEIO AMBIENTE E DOS RECURSOS NATURAIS. Instrução normativa nº 160, de 27 de abril de 2007. *Diário Oficial da União*, Brasília, DF, n. 82, p. 404-405, abr. 2007. Seção 1. Disponível em: <http://www.ibama.gov.br/sophia/cnia/legislacao/IBAMA/IN0160-270407.PDF>. Acesso em: 29 maio 2018.

INSTITUTO CHICO MENDES DE CONSERVAÇÃO DA BIODIVERSIDADE. Instrução Normativa nº 2, de 25 de novembro de 2015. *Diário Oficial da União*, Brasília, DF, n. 226, p. 83, nov. 2015. Seção 1. Disponível em: http://www.icmbio.gov.br/cepsul/images/stories/legislacao/Instrucao_normativa/2015/in_icmbio_2_2015_politica_dados_informa%C3%A7oes_icmbio.pdf. Acesso em: 2 jun. 2018.

INSTITUTO NACIONAL DE METEOROLOGIA (Brasil). *INMET*: nota meteorológica São Paulo capital. São Paulo, 20 ago. 2019. Disponível em: http://www.inmet.gov.br/portal/notas_tecnicas/2019/nota_7dis_sp_20190820.pdf. Acesso em: 5 nov. 2019.

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS. *Monitoramento dos focos ativos por países*. [Brasília, DF], 2019. Disponível em: http://queimadas.dgi.inpe.br/queimadas/portal-static/estatisticas_paises/. Acesso em: 19 set. 2019.

INSTITUTO SOCIEDADE, POPULAÇÃO E NATUREZA. *O bioma ameaçado*. Brasília, DF, [2018]. Disponível em: <http://www.ispn.org.br/o-cerrado/o-bioma-ameacado/>. Acesso em: 13 maio 2018.

INTERGOVERNMENTAL SCIENCE-POLICY PLATFORM ON BIODIVERSITY AND ECOSYSTEM SERVICES. *Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services*. Bonn: IPBES secretariat, 2019. 56 p. Available from: https://ipbes.net/system/tdf/inline/files/ipbes_global_assessment_report_summary_for_policymakers.pdf?file=1&type=node&id=36213. Access on: 1 Dec. 2019.

INTERNATIONAL COMMISSION ON ZOOLOGICAL NOMENCLATURE. Amendment of Articles 8, 9, 10, 21 and 78 of the International Code of Zoological Nomenclature to expand and refine methods of publication. *ZooKeys*, Sofia, v. 219, p. 1-10, Sept. 2012. DOI: 10.3897/zookeys.219.3994. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/22977348>. Access on: 13 Nov. 2019.

INTERNATIONAL COMMISSION ON ZOOLOGICAL NOMENCLATURE. *International Code of Zoological Nomenclature*. 4th ed. London: The International Trust for Zoological Nomenclature, 1999. 149 p. Adopted by the International Union of Biological Sciences.

INTERNATIONAL COMMISSION ON ZOOLOGICAL NOMENCLATURE. Proposed amendment of the International Code of Zoological Nomenclature to expand and refine methods of publication. *Zootaxa*, Auckland, v. 1908, p. 57-67, Oct. 2008. DOI: <http://dx.doi.org/10.11646/zootaxa.1908.1.5>. Available from: <https://www.biotaxa.org/Zootaxa/article/view/zootaxa.1908.1.5>. Access on: 13 Nov. 2019.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. *ISO 8601:2004*: Data elements and interchange formats: Information interchange: Representation of dates and times. [Geneva], 2004. 33 p.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. *ISO 8879:1986(E)*: Information processing: text and office systems: Standard Generalized Markup Language (SGML). [Geneva], 1986.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION; INTERNATIONAL ELECTROTECHNICAL COMMISSION. *ISO/IEC 11179-4*: information technology: metadata registries (MDR): part 4: formulation of data definitions. 2nd ed. Geneva, 2004. 9 p.

INTERNATIONAL UNION FOR CONSERVATION OF NATURE. *IUCN*: A brief history. © 2018. Available from: <https://www.iucn.org/about/iucn-brief-history>. Access on: 29 May 2018.

JAIN, K. *International Code of Zoological Nomenclature*. Biology Discussion, [2018]. Available from: <http://www.biologydiscussion.com/animals-2/international-code-of-zoological-nomenclature/32399>. Access on: 3 Aug. 2018.

JAVLIN, A. S. Part I: overview. In: JAVLIN, A. S. *CloverDX Designer*. User's Guide. [S. l.]: CloverDX, 2019. This User's Guide refers to CloverDX Designer 5.1.x release. Available from: <https://doc.cloverdx.com/documentation/UserGuide/index.jsp?topic=/com.cloveretl.proufiler.doc/docs/index.html>. Access on: 22 Feb. 2019.

JIRÁSEK, V. Evolution of the proposals of taxonomical categories for the classification of cultivated plants. *Taxon*, [Utrecht], v. 10, n. 2, p. 34-45, Feb. 1961. Available from: <http://www.jstor.org/stable/1217450>. Access on: 8 Sept. 2018.

JOHNSON, N.; CORA, J. C. A. *Triplehorn Insect Collection (OSUC)*. Columbus: Museum of Biological Diversity, The Ohio State University, [2019?]. DOI: <https://doi.org/10.15468/efb17f>. Available from: <https://www.gbif.org/occurrence/872918732>. Accessed on: 22 Nov. 2019.

KARTHICK, B.; WILLIAMS, D. M. The International Code for Nomenclature for algae, fungi and plants: a significant rewrite of the International Code of Botanical Nomenclature. *Current Science*, Bangalore, v. 102, n. 4, p. 551-553, Feb. 2012.

KNAPP, S.; POLASZEK, A.; WATSON, M. Spreading the word. *Nature*, [London], v. 446, p. 261-262, Mar. 2007. DOI: <https://doi.org/10.1038/446261a>. Available from: <https://www.nature.com/articles/446261a>. Access on: 3 Aug. 2018.

Latrodectus mactans (Fabricius, 1775). In: GBIF SECRETARIAT. *GBIF Backbone Taxonomy*. Copenhagen: GBIF.org, 2019. DOI: <https://doi.org/10.15468/39omei>. Available from: <https://www.gbif.org/dataset/d7dddbf4-2cf0-4f39-9b2a-bb099caae36c>. Access on: 5 Nov. 2019.

LOPACKA, J. *Reproduction of painting Mona Lisa by Leonardo da Vinci and light graphic effect*. Adobe Stock, c2019. Available from: https://stock.adobe.com/br/images/id/138089510?as_campaign=pixabay&as_content=api&tduid=d4fe30c2334ec775b87419baeb43491d&as_channel=affiliate&as_campaign=redirect&as_source=arvato. Access on: 28 Apr. 2019.

LOURENÇO, C. A. Representação de informação, metadados, interoperabilidade e recuperação da informação na atualidade. In: SEMINÁRIO DO GRUPO DE PESQUISA <MHTX>, 2., 2016, Belo Horizonte. *Anais [...]*. Belo Horizonte: ECI/UFMG, 2016. p. 95-101.

MAGGYDURCH. [*Favo de mel*]. 1 fotografia, color., 4752x3168 px. 27 abr. 2012. Disponível em: <https://pixabay.com/pt/photos/abelhas-rainha-favo-de-mel-1163028/>. Acesso em: 19 set. 2019.

MARQUES, M. C. M. *et al.* Mata Atlântica: o desafio de transformar um passado de devastação em um futuro de conhecimento e conservação. In: PEIXOTO, A. L.; LUZ, J. R. P.; BRITO, M. A. (org.). *Conhecendo a biodiversidade*. Brasília, DF: MCTIC: CNPq: PPBio, 2016. p. 50-68.

MAY, P. H. (coord.); VEIGA NETO, F. C.; POZO, O. V. C. *Valoração econômica da biodiversidade: estudos de caso no Brasil*. [Brasília, DF]: Ministério do Meio Ambiente, 2000. 198 p.

METADATA, structural. In: FEDERAL AGENCIES DIGITAL GUIDELINES INITIATIVE. *Glossary*. [Washington, DC: Library of Congress], 2017. Available from: <http://www.digitizationguidelines.gov/term.php?term=metadatastructural>. Access on: 12 Feb. 2019.

MICROSOFT. *Well Known Text Module*. [S. l.], 27 Feb. 2018. Available from: <https://docs.microsoft.com/en-us/bingmaps/v8-web-control/modules/well-known-text-module>. Access on: 7 Out. 2019.

PRINCÍPIO. In: *MICHAELIS: moderno dicionário da língua portuguesa*. Disponível em: <https://michaelis.uol.com.br/moderno-portugues/busca/portugues-brasileiro/princ%C3%ADpio/>. Acesso em: 2 jan. 2020.

MINELLI, A. Zoological nomenclature in the digital era. *Frontiers in Zoology*, London, v. 10, n. 4, p. 1-7, 2013. DOI: 10.1186/1742-9994-10-4. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3568006/>. Access on: 3 ago. 2018.

MINISTÉRIO DO MEIO AMBIENTE (Brasil). *Espécies ameaçadas*. Brasília, DF, 2019. Disponível em: <https://www.mma.gov.br/biodiversidade/conservacao-de-especies/fauna-ameacada.html>. Acesso em: 5 nov. 2019.

MORIN, P. J. *Community ecology*. 2nd ed. Oxford: Wiley-Blackwell, 2011. 407 p.

MOYES, C. D.; SCHULTE, P. M. *Princípios de fisiologia animal*. Tradução Adriane Belló Klein *et al.* 2. ed. Porto Alegre: Artmed, 2010. 792 p.

NAÇÕES UNIDAS BRASIL. *Objetivo 15*: proteger, recuperar e promover o uso sustentável dos ecossistemas terrestres, gerir de forma sustentável as florestas, combater a desertificação, deter e reverter a degradação da terra e deter a perda de biodiversidade. Rio de Janeiro: ONUBR, 2015. Disponível em: <https://nacoesunidas.org/pos2015/ods15/>. Acesso em: 13 nov. 2019.

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION. *Scientific consensus: Earth's climate is warming*. Pasadena: NASA's Jet Propulsion Laboratory: California Institute of Technology, 2019. Available from: <https://climate.nasa.gov/scientific-consensus/>. Access on: 17 Nov. 2019.

NATIONAL INFORMATION STANDARDS ORGANIZATION. *A framework of guidance for building good digital collections*. 3rd ed. Baltimore, 2007. 95 p.

NATIONAL INFORMATION STANDARDS ORGANIZATION. *Understanding metadata*. Bethesda: NISO Press, 2004.

NICOLSON, D. H. A History of Botanical Nomenclature. *Annals of the Missouri Botanical Garden*, St. Louis, MO, v. 78, n. 1, p. 33-56, 1991.

NOBRE, C. A. *et al.* Land-use and climate change risks in the Amazon and the need of a novel sustainable development paradigm. *Proceedings of the National Academy of Sciences of the United States of America*, [Washington, DC], v. 113, n. 39, p. 10759-10768, Sept. 2016. DOI: 10.1073/pnas.1605516113. Available from: <http://www.pnas.org/content/pnas/113/39/10759.full.pdf>. Access on: 7 May 2018.

NORMA. *In*: MICHAELIS: dicionário brasileiro da língua portuguesa. Disponível em: <https://michaelis.uol.com.br/moderno-portugues/busca/portugues-brasileiro/norma/>. Acesso em: 2 jan. 2020.

ODUM, E. P.; BARRET, G. W. *Fundamentos de ecologia*. Tradução de Pégasus Sistemas e Soluções. São Paulo: Cengage Learning, 2006. 632 p. Título original: *Fundamentals of ecology*. Tradução da 5. ed. norte americana.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. *Interaction (between species)*. *In*: GLOSSARY of statistical terms. [Paris], 2001. Available from: <https://stats.oecd.org/glossary/about.asp>. Access on: 20 Sept. 2019.

PACHECO, D. J. P.; CORRÊA-FERREIRA, B. S. Parasitismo de *Telenomus podisi* Ashmead (Hymenoptera: Scelionidae) em Populações de Percevejos Pragas da Soja. *Anais da Sociedade Entomológica do Brasil*, [Londrina], v. 29, n. 2, p. 295-302, jun. 2000. Disponível em: <http://www.scielo.br/pdf/aseb/v29n2/v29n2a11.pdf>. Acesso em: 29 abr. 2019.

Panthera leo (Linnaeus, 1758). *In*: GBIF SECRETARIAT. *GBIF Backbone Taxonomy*. Copenhagen: GBIF.org, 2019. DOI: <https://doi.org/10.15468/39omei>. Available from: <https://www.gbif.org/dataset/d7ddd4-2cf0-4f39-9b2a-bb099caae36c>. Access on: 5 Nov. 2019.

PARASITES. *In*: AGROVOC Multilingual Thesaurus. [S. l.]: FAO, 2011. Available from: http://aims.fao.org/aos/agrovoc/c_5574. Access on: 24 Nov. 2019.

PARASITISM. *In*: AGROVOC Multilingual Thesaurus. [S. l.]: FAO, 2011. Available from: http://agrovoc.uniroma2.it/agrovoc/agrovoc/en/page/c_5577. Access on: 5 Nov. 2019.

PEARCE, D.; MORAN, D. *The Economic Value of Biodiversity*. London: Earthscan, 1994. 106 p.

PEGORIM, J. Por que o céu escureceu em São Paulo? *CLIMATEMPO*, São Paulo, 19 ago. 2019. Seção Notícias. Disponível em: <https://www.climatempo.com.br/noticia/2019/08/19/por-que-o-ceu-escureceu-em-sao-paulo--7543>. Acesso em: 5 nov. 2019.

PIEROZZI JUNIOR, I. [*Modelo de dados*] <onto.biodiversidade>. Campinas, 2012. Não publicado. Produzido no software yEd Graph Editor.

PINTO-COELHO, R. M. *Fundamentos em ecologia*. Porto Alegre: Artmed, 2007. 252 p.

PIXABAY. [*Abelha*]. 1 fotografia, color., 6000 x 4000 px. 31 mar. 2019. Disponível em: <https://pixabay.com/pt/photos/abelha-mel-de-abelha-inseto-p%C3%B3len-4092987/>. Acesso em: 20 set. 2019.

PIXABAY. [*Zebra e girafa*]. 1 fotografia, color., 3939x2215 px. 2 fev. 2016. Disponível em: <https://pixabay.com/pt/photos/zebra-gnu-girafa-%C3%A1frica-nam%C3%ADbia-1170177/>. Acesso em: 25 nov. 2019.

PLANTAS Daninhas: saiba as principais plantas que todo produtor precisa se cuidar. *Tecnologia no Campo*, 12 maio 2019. Disponível em: <https://tecnologianocampo.com.br/plantas-daninhas/>. Acesso em: 1 ago. 2019.

PLINIAN CORE TASK GROUP. *Plinian Core*. [San Fransisco]: TDWG, 2018. Available from: <https://github.com/tdwg/PlinianCore>. Access on: 25 Sept. 2019.

POLLOCK, J. *Web Semântica para Leigos*. Rio de Janeiro: Alta Books, 2010. 424 p.

POMERANTZ, J. *Metadata*. Cambridge, MA: The MIT Press, 2015.

PORTAIL des Collections des Musées de France. *Joconde*. [Paris], c1999. Disponible à: http://www2.culture.gouv.fr/public/mistral/joconde_fr?ACTION=CHERCHER&FIELD_1=REF&VALUE_1=000PE025604. Consulté à: 28 abr. 2019.

POWELL, A. *et al.* *DCMI Abstract Model*. [Maryland]: Dublin Core Metadata Initiative, 2007. Available from: <http://dublincore.org/documents/abstract-model/>. Access on: 14 fev. 2019.

PRADO, J. *et al.* A Sessão de Nomenclatura em Shenzhen (China) e as principais modificações no Código Internacional de Nomenclatura para Algas, Fungos e Plantas. *Rodriguésia*, [Rio de Janeiro], v. 68, n. 4, p. 1499-1503, 2017. Disponível em: <http://rodriguesia.jbrj.gov.br/FASCICULOS/rodrig68-4/30%20-%200297-2017.pdf>. Acesso em: 8 set. 2018.

PRINCÍPIO. *In*: MICHAELIS: moderno dicionário da língua portuguesa. Disponível em: <https://michaelis.uol.com.br/moderno-portugues/busca/portugues-brasileiro/princ%C3%ADpio/>. Acesso em: 2 jan. 2020.

RAMOS, S. E. *et al.* Cannibalism, growth performance, and body composition of giant trahira juveniles under different photoperiods. *Pesquisa agropecuária brasileira*, Brasília, DF, v. 53, n. 6, p. 664-672, June 2018. DOI: 10.1590/S0100-204X2018000600002. Available from: <https://seer.sct.embrapa.br/index.php/pab/article/view/25778/14196>. Access on: 20 Sept. 2019.

RDF WORKING GROUP. *Resource Description Framework (RDF)*. [S. l.]: W3C Semantic Web, 25 Feb. 2014. Available from: <https://www.w3.org/RDF/>. Access on: 23 Nov. 2019.

RIDE, W. D. L. Introduction. *In*: INTERNATIONAL COMMISSION ON ZOOLOGICAL NOMENCLATURE. *International Code of Zoological Nomenclature*. London: International Trust for Zoological Nomenclature, 1999. p. xix-xxix.

RILEY, J. *Understanding Metadata: what is Metadata, and what is it for?: a primer*. Baltimore: NISO, 2017. 45 p. Available from: https://groups.niso.org/apps/group_public/download.php/17446/Understanding%20Metadata.pdf. Access on: 16 Feb. 2019.

RIZZARDI, M. A. *et al.* Competição por recursos do solo entre ervas daninhas e culturas. *Ciência Rural*, Santa Maria, v. 31, n. 4, p. 707-714, ago. 2001. DOI: <http://dx.doi.org/10.1590/S0103-84782001000400026>. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-84782001000400026&lng=en&nrm=iso. Acesso em: 1 ago. 2019.

ROCHA, L. A picada da viúva-negra pode trazer consequências para o sistema nervoso central dos indivíduos. *In*: AGUIAR, R (ed.). *Atenção redobrada: IOC e Instituto Vital Brazil são parceiros em capacitação sobre os riscos relacionados a animais peçonhentos, como serpentes, escorpiões e aranhas*. *Comunicação/Instituto Oswaldo Cruz*, 17 jul. 2017. Disponível em: [http://www.fiocruz.br/ioc/cgi/cgilua.exe/sys/start.htm?infoid=2825&sid=32&tpl=printer view](http://www.fiocruz.br/ioc/cgi/cgilua.exe/sys/start.htm?infoid=2825&sid=32&tpl=printer_view). Acesso em: 19 set. 2019.

RODRIGUES, E. *et al.* *Os repositórios de dados científicos: estado da arte*. Minho: Porto: Repositório Científico de Acesso Aberto de Portugal, 2010. 54 p. Relatório.

RODRIGUES, W. A. Revisão da nomenclatura botânica: datas importantes na história da nomenclatura botânica pós-linneana. *Publicatio UEPG*, Ponta Grossa, v. 9, n. 3/4, p. 73-82, set./dez. 2003.

RUTLEDGE, K. *et al.* Biome. *In: NATIONAL GEOGRAPHIC SOCIETY. National Geographic Education: resource library.* [S. l.]: National Geographic, 2011. Available from: <https://www.nationalgeographic.org/encyclopedia/biome/>. Access on: 11 July 2019.

SADAVA, D. *et al.* *Vida: a ciência da biologia.* 8. ed. Tradução Carla Denise Bonan *et al.* Porto Alegre: Artmed, 2009. (Evolução, diversidade e ecologia, v. 2).

SAITO, M. L. *As plantas praguicidas: alternativa para o controle de pragas da agricultura.* Jaguariúna: Embrapa Meio Ambiente, [2015?]. Disponível em: <http://www.cnpma.embrapa.br/nova/mostra2.php3?id=77>. Acesso em: 2 ago. 2019.

SANCHES, F. N. [Foto]. *In: VALE, A. Joaninha é indicada para controle biológico de pragas. Embrapa [notícias],* Brasília, DF, 28 jun. 2016. Agroecologia e produção orgânica, p. 2. Disponível em: <https://www.embrapa.br/busca-de-noticias/-/noticia/13902532/joaninha-e-indicada-para-controle-biologico-de-pragas>. Acesso em: 20 set. 2019.

SANCHES, N. F.; CARVALHO, R. S. Nova metodologia e procedimentos para criação da joaninha predadora exótica *Cryptolaemus montrouzieri*. *In: CONGRESSO BRASILEIRO DE FRUTICULTURA, 21., 2010, Natal. Frutas: saúde, inovação e responsabilidade: anais.* Natal: Sociedade Brasileira de Fruticultura, 2010. Disponível em: <http://www.alice.cnptia.embrapa.br/handle/doc/874114>. Acesso em: 20 set. 2019.

SANTOS, J. L. C. *et al.* Ontologias para Interoperabilidade de Modelos e Sistemas de Informação de Biodiversidade. *In: IBEROAMERICAN MEETING OF ONTOLOGICAL RESEARCH, Gramado, 2011. Eletronic proceedings [...]* [Gramado]: CEUR, 2011. Não paginado. Disponível em: <http://ceur-ws.org/Vol-728/paper8.pdf>. Acesso em: 28 ago. 2018.

SARAIVA, A. M. Biodiversity Informatics in Brazil: A personal perspective. *Biodiversity Information Science and Standards*, [Sofia, Bulgária], v. 3, p. e37373, 26 Jun. 2019. DOI: 10.3897/biss.3.37373. Available from: <https://biss.pensoft.net/article/37373/>. Access on: 5 Nov. 2019.

SCHENKEL, R. XML for Beginners. *In: ORGANIZING AND SEARCHING INFORMATION WITH XML, 2003, [Saarbrücken]. Proceedings [...]* [Saarbrücken]: Saarland University, 2003. p. 1-55. Available from: <http://resources.mpi-inf.mpg.de/d5/teaching/ss03/xml-seminar/talks/xml%20for%20beginners.pdf>. Access on: 10 July 2019.

SCHUMANN, K. *Red tide in San Diego, CA.* 1 photo, color. *In: LANGLOIS, G. Red Tides in California.* Sea Grant California, San Diego, Sept. 2018. Available from: <https://caseagrants.ucsd.edu/extension-outreach/facts-and-resources/red-tides-in-california>. Access on: 19 Sept. 2019.

SECRETARIA DE EDUCAÇÃO (Paraná). Relações ecológicas: aula III. Curitiba, 22 jul. 2013. Disponível em:

<http://webcache.googleusercontent.com/search?q=cache:jW-NBe-i8VEJ:www.biologia.seed.pr.gov.br/modules/conteudo/conteudo.php%3Fconteudo%3D372+&cd=21&hl=pt-BR&ct=clnk&gl=br>. Acesso em: 18 set. 2019.

SILVA, E. L.; MENEZES, E. M. *Metodologia da pesquisa e elaboração de dissertação*. Florianópolis: Laboratório de Ensino à distância, 2001. 118 p.

SILVA, S. M.; FARINELLI, F.; ALMEIDA, M. B. Um roteiro para modelagem conceitual de sistemas de informação baseada em princípios ontológicos. ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 15., 2014, Belo Horizonte. *Além das nuvens: expandindo as fronteiras da ciência da informação*. Belo Horizonte: Associação Nacional de Pesquisa e Pós-graduação em Ciência da Informação, 2014. p. 4072-4095

SILVÉRIO, D. *et al.* *Amazônia em chamas: nota técnica do Instituto de Pesquisa Ambiental da Amazônia - IPAM*. Brasília, DF: IPAM, 2019. Disponível em: <https://ipam.org.br/wp-content/uploads/2019/08/NT-Fogo-Amazo%CC%82nia-2019.pdf>. Acesso em: 19 set. 2019.

SISTEMA DE INFORMAÇÃO SOBRE A BIODIVERSIDADE BRASILEIRA. *Termos Darwin Core*. Brasília, DF: Ministério da Ciência, Tecnologia, Inovações e Comunicações, [2018?]. Disponível em: <http://www.sibbr.gov.br/areas/index.php?area=publicar&subarea=termos-dwc>. Acesso em: 3 junho 2018.

SOARES, F. M.; HAMANAKA, R. Y.; MACULAN, B. C. M. S. Gestão de dados da biodiversidade: aplicação do padrão de metadados Darwin Core. *In: WORKSHOP DE INFORMAÇÃO, DADOS E TECNOLOGIA, 2.*, 2018, João Pessoa. *Anais WIDaT 2018*. João Pessoa: Ed. UFPB, 2018. p. 141-146. Disponível em: <http://www.ufpb.br/widat/contents/documentos/anaiswidat2018.pdf>. Acesso em: 19 set. 2019.

SORCI, G.; GARNIER, S. Parasitism. *In: JØRGENSEN, S. E.; FATH, B. F. Encyclopedia of Ecology*. Amsterdam: Elsevier, 2008. p. 2645-2650. DOI: 10.1016/B978-008045405-4.00814-4. Available from: <https://www.sciencedirect.com/science/article/pii/B9780080454054008144>. Access on: 24 Nov. 2019.

SPIELMANN, A. A. *Fungos liquenizados (liquens)*. Orientador: Marcelo Pinto Marcelli. 2006. 13 f. Relatório (Estágio em Micologia e Liquenologia) – Instituto de Botânica, São Paulo, 2006. Disponível em: http://www.biodiversidade.pgibt.ibot.sp.gov.br/Web/pdf/Fungos_Liquenizados_Spielmann_&_Marcelli.pdf. Acesso em: 20 set. 2019.

STEIN, R. T. *Ecologia geral*. Porto Alegre: SAGAH, 2018. 221 p. Revisão técnica: Diogo Ribeiro do Couto.

- STEINBERGER, M. Poluição urbana do ar por queimadas na Amazônia brasileira. *In: ENCONTRO DA ASSOCIAÇÃO BRASILEIRA DE ESTUDOS POPULACIONAIS*, 13., 2002, Ouro Preto. *Anais [...]*. Ouro Preto: Associação Brasileira de Estudos Populacionais, 2002. p. 1-13. Disponível em: <http://www.abep.org.br/~abeporgb/publicacoes/index.php/anais/article/viewFile/1173/1137>. Acesso em: 20 set. 2019.
- TASK FORCE ON METADATA. *Committee on Cataloging: Description & Access*. Final Report. [S. l.]: American Library Association, 2000. Available from: <https://www.libraries.psu.edu/tas/jca/ccda/tf-meta6.html>. Access on: 23 Sept. 2019.
- Telenomus podisi* (Ashmead, 1893). *In: GBIF SECRETARIAT. GBIF Backbone Taxonomy*. Copenhagen: GBIF.org, 2019. DOI: <https://doi.org/10.15468/39omei>. Available from <https://www.gbif.org/species/1401314>. Access on: 4 Nov. 2019.
- TEOREY, T. J. *Database modeling and design*. Burlington: Morgan Kaufmann, 1999. 366 p.
- TORRES, F. T. P. *et al.* Perfil dos incêndios florestais em unidades de conservação brasileiras no período de 2008 a 2012. *Floresta*, Curitiba, v. 46, n. 4, p. 531-542, out./dez. 2016. Disponível em: <https://revistas.ufpr.br/floresta/article/view/44199/30153>. Acesso em: 19 set. 2019.
- TURLAND, N. J. *et al.* (ed.). *International Code of Nomenclature for algae, fungi, and plants (Shenzhen Code) adopted by the Nineteenth International Botanical Congress Shenzhen, China, July 2017*. Glashütten: Koeltz Botanical Books, 2018. (Regnum Vegetabile, 159). Available from: <https://doi.org/10.12705/Code.2018>. Access on: 8 Sept. 2018.
- UNITED NATIONS EDUCATION, SCIENTIFIC AND CULTURAL ORGANIZATION. Representação da Unesco no Brasil. *Biodiversidade no Brasil*. [Brasília, DF], c2017. Disponível em: <http://www.unesco.org/new/pt/brasil/natural-sciences/environment/biodiversity/>. Acesso em: 19 set. 2019.
- UNIVERSIDADE DO VALE DO ITAJAÍ. *A teoria do neutralismo*. Itajaí, 2017. Disponível em: <https://www.studocu.com/pt-br/document/universidade-do-vale-do-itajai/biogeografia/tarefas/a-teoria-do-neutralismo/4833248/view>. Acesso em: 19 set. 2019.
- VALE, A. Joaquina é indicada para controle biológico de pragas. *Embrapa [notícias]*, Brasília, DF, 28 jun. 2016. Agroecologia e produção orgânica, p. 2. Disponível em: <https://www.embrapa.br/busca-de-noticias/-/noticia/13902532/joaquina-e-indicada-para-controle-biologico-de-pragas>. Acesso em: 20 set. 2019.
- VAN RIJCKEVORSEL, P. *A brief history of the Code*. Bratislava: IAPT, 2014a. Available from: <https://www.iapt-taxon.org/historic/history.htm>. Access on: 8 Aug. 2018.
- VAN RIJCKEVORSEL, P. *Overview of editions of the Code*. Bratislava: IAPT, 2014b. Available from: <https://www.iapt-taxon.org/historic/index.htm>. Access on: 8 Aug. 2018.

WALLS, L. R. *et al.* Semantics in Support of Biodiversity Knowledge Discovery: An Introduction to the Biological Collections Ontology and Related Ontologies. *PLoS ONE*, [San Francisco], v. 9, n. 3, p. e89606, Mar. 2014.

DOI:10.1371/journal.pone.0089606. Available from:

<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0089606>. Access on: 8 Sept. 2018.

WALSH, N. *What is XML?* [S. l.]: O'Reilly & Associates, 1998. Available from:

http://www.ce.unipr.it/people/bianchi/Teaching/IntelligenzaArtificiale/rdf_pl/XML-RDF/xmlguide1.html. Access on: 10 July 2019.

WESSELS, C. *Roar and peace?*: Incredible images show battle for dominance between mighty lions. [S. l.]: Caters News Agency, c2019. Available from:

<https://www.catersnews.com/stories/animals/roar-and-peace-incredible-images-show-battle-for-dominance-between-mighty-lions/>. Access on: 1 Aug. 2019.

WHEELER, Q. D. (ed.). *The New Taxonomy*. New York: CRC Press, 2008. xi, 237 p.

WIECZOREK, J. *et al.* *Darwin Core Terms*: A quick reference guide. [San Francisco]: TDWG, 2009a. Available from: <https://dwc.tdwg.org/terms>. Access on: 26 Feb. 2019.

WIECZOREK, J. *et al.* Darwin Core: An Evolving Community-Developed Biodiversity Data Standard. *PLoS ONE*, [San Francisco], v. 7, n. 1, p. e29715, Jan. 2012.

Available from: doi:10.1371/journal.pone.0029715. Access on: 8 Sept. 2018.

WIECZOREK, J. *et al.* *Simple Darwin Core*. [San Francisco]: TDWG, 2009b.

Available from: <https://dwc.tdwg.org/simple/#5-are-there-any-rules>. Access on: 4 Nov. 2019.

WORLD WILDLIFE FUND. *History*. Washington, DC, © 2018. Available from:

<https://www.worldwildlife.org/about/history>. Access on: 10 May 2018.

YOUNG, P. Nodules made by *Rhizobium leguminosarum* on the roots of common vetch. 1 photo, color. *In*: GARNER, D. *Understanding the personalities of bacteria*. University of New York [News], New York, 14 Jan. 2015. Available from:

<https://www.york.ac.uk/news-and-events/news/2015/research/bacteria-personalities/>. Access on: 20 Sept. 2019.

ZENG, M. L. *Metadata basics*. Version 2.0. [Kent], 2015. Available from:

<https://marciazeng.slis.kent.edu/metadatabasics/types.htm>. Access on: 20 Sept. 2019.

ZENG, M. L.; QIN, J. *Metadata*. New York: Neal-Scguman Publishers, 2008.

ZERMOGLIO, P. Simple Darwin Core. *In*: *TDWG Darwin Core Task Group [wiki]*,

[San Fransisco], 3 Feb. 2018. Available from: <https://github.com/tdwg/dwc-qa/wiki/Simple-Darwin-Core>. Access on: 30 Jan. 2019.

APÊNDICE A – LISTA DE EXTENSÕES ESTÁVEIS DO DARWIN CORE

(Continua)

TÍTULO	ATRIBUTOS	DEFINIÇÃO
GGBN Amplification Extension	Definição	Suporte para Amplificações de DNA como uma extensão para a classe de dados de amostragem <i>MaterialSample</i> no DwC. Destina-se a ser uma relação do tipo um-para-muitos (1-n) para a classe <i>MaterialSample</i> .
	Disponível em	http://terms.tdwg.org/wiki/GGBN_Amplification_Vocabulary
	Propriedades	34
	Nome da classe	Amplification
	Name space	http://data.ggbn.org/schemas/ggbn/terms/
	RowType	http://data.ggbn.org/schemas/ggbn/terms/Amplification
GGBN DNA Cloning Extension	Definição	Suporte para Clonagem de DNA como uma extensão para a classe de dados de amostragem <i>MaterialSample</i> no DwC. Destina-se a ser uma relação do tipo um-para-muitos (1-n) para a classe <i>MaterialSample</i> .
	Disponível em	http://terms.tdwg.org/wiki/GGBN_DNA_Cloning_Vocabulary
	Propriedades	16
	Nome da classe	Cloning
	Name space	http://data.ggbn.org/schemas/ggbn/terms/
	RowType	http://data.ggbn.org/schemas/ggbn/terms/Cloning
GGBN Gel Image Extension	Definição	Suporte para propriedades de imagem em gel como uma extensão para a classe de dados de amostragem <i>MaterialSample</i> no DwC. Destina-se a ser uma relação do tipo um-para-muitos (1-n) para a classe <i>MaterialSample</i> .
	Disponível em	http://terms.tdwg.org/wiki/GGBN_Gel_Image_Vocabulary
	Propriedades	12
	Nome da classe	GellImage
	Name space	http://data.ggbn.org/schemas/ggbn/terms/
	RowType	http://data.ggbn.org/schemas/ggbn/terms/GellImage

APÊNDICE A – LISTA DE EXTENSÕES ESTÁVEIS DO DARWIN CORE

(Continuação)

TÍTULO	ATRIBUTOS	DEFINIÇÃO
GGBN Loan Extension	Definição	Suporte para propriedades de empréstimo de informações sobre espécimes, tecidos ou amostras de DNA como uma extensão para a classe de dados de amostragem <i>MaterialSample</i> no DwC. Destina-se a ser uma relação do tipo um-para-um (1-1) para a classe <i>MaterialSample</i> .
	Disponível em	http://terms.tdwg.org/wiki/GGBN_Loan_Vocabulary
	Propriedades	8
	Nome da classe	Loan
	RowType	http://data.ggbn.org/schemas/ggbn/terms/Loan
GGBN Material Sample Extension	Definição	Suporte para propriedades da amostra do material (por exemplo, tecidos, DNA, RNA) como uma extensão para a classe de dados de amostragem <i>MaterialSample</i> no DwC. Destina-se a ser uma relação do tipo um-para-um (1-1) para a classe <i>MaterialSample</i> .
	Disponível em	http://terms.tdwg.org/wiki/GGBN_Material_Sample_Vocabulary
	Propriedades	22
	Nome da classe	MaterialSample
	RowType	http://data.ggbn.org/schemas/ggbn/terms/MaterialSample
GGBN Permit Extension	Definição	Suporte para todos os tipos de permissões legais relacionadas a aspectos de aquisição de amostras, empréstimo e uso como uma extensão para a classe de dados de amostragem <i>MaterialSample</i> no DwC. Destina-se a ser uma relação do tipo um-para-muitos (1-n) para a classe <i>MaterialSample</i> .
	Disponível em	http://terms.tdwg.org/wiki/GGBN_Permit_Vocabulary
	Propriedades	5
	Nome da classe	Permit
	RowType	http://data.ggbn.org/schemas/ggbn/terms/Permit

APÊNDICE A – LISTA DE EXTENSÕES ESTÁVEIS DO DARWIN CORE

(Continuação)

TÍTULO	ATRIBUTOS	DEFINIÇÃO
GGBN Preparation Extension	Definição	Suporte para propriedades de preparação do espécime, da amostra de tecido ou extração de DNA (tratado como preparação) como uma extensão para uma classe do tipo <i>MaterialSample</i> . Destina-se a ser uma relação do tipo um-para-muitos (1-n) para a classe <i>MaterialSample</i> .
	Disponível em	http://terms.tdwg.org/wiki/GGBN_Preparation_Vocabulary
	Propriedades	6
	Nome da classe	Preparation
	Name space	http://data.ggbn.org/schemas/ggbn/terms/
	RowType	http://data.ggbn.org/schemas/ggbn/terms/Preparation
GGBN Preservation Extension	Definição	Suporte para todos os tipos de preservação de amostra em uma coleção física como uma extensão para a classe de dados de amostragem <i>MaterialSample</i> no DwC. Destina-se a ser uma relação do tipo um-para-muitos (1-n) para a classe <i>MaterialSample</i> .
	Disponível em	http://terms.tdwg.org/wiki/GGBN_Preservation_Vocabulary
	Propriedades	4
	Nome da classe	Preservation
	Name space	http://data.ggbn.org/schemas/ggbn/terms/
	RowType	http://data.ggbn.org/schemas/ggbn/terms/Preservation
EOL Media Extension 1.0	Definição	Esta extensão importa metadados do Audubon Core, Dublin Core e outros para coletar informação sobre texto e multimídia. Ele foi projetado para conter todos os metadados que precisam ser indexados pela Encyclopedia of Life (EOL), mas essa extensão é, com sorte, geral o suficiente para ser útil a todos os provedores e consumidores de texto e mídia.
	Disponível em	http://eol.org/info/cp_archives
	Propriedades	31
	Nome da classe	EOLMediaExtension
	Name space	http://www.eol.org/schema/transfer#
	RowType	http://eol.org/schema/media/Document

APÊNDICE A – LISTA DE EXTENSÕES ESTÁVEIS DO DARWIN CORE

(Continuação)

TÍTULO	ATRIBUTOS	DEFINIÇÃO
EOL References Extension 1.0	Definição	Essa extensão importa metadados da BIBO (http://bibliontology.com/), Dublin Core e outros para coletar informações sobre citações e referências bibliográficas.
	Disponível em	http://eol.org/info/cp_archives
	Propriedades	18
	Nome da classe	EOLReferencesExtension
	Name space	http://www.eol.org/schema/transfer#
	RowType	http://eol.org/schema/reference/Reference
Germplasm accession (v20140515) ⁴⁵	Definição	O Germplasm Vocabulary fornece um conjunto de termos (complementando os termos do Darwin Core) para descrever os acessos do banco de genes. Esses termos são mantidos pela comunidade temática de recursos genéticos de plantas para alimentação e agricultura (PGRFA). A maioria desses termos é importada da Multi-Crop Passport Descriptor List (MCPD) mantida pela Bioversity International e pela Organização das Nações Unidas para Agricultura e Alimentação (FAO). Alguns termos também foram desenvolvidos pelo Programa Cooperativo Europeu para Recursos Genéticos de Plantas (ECPGR).
	Disponível em	http://purl.org/germplasm/germplasmTerm
	Propriedades	37
	Nome da classe	GermplasmAccession
	Name space	http://purl.org/germplasm/germplasmTerm#
	RowType	http://purl.org/germplasm/germplasmTerm#GermplasmAccession

⁴⁵ Extensão indisponível para acesso quando consultada em 30 set. 2019.

APÊNDICE A – LISTA DE EXTENSÕES ESTÁVEIS DO DARWIN CORE

(Continuação)

TÍTULO	ATRIBUTOS	DEFINIÇÃO
Trait measurement score (v20140515)	<p>Definição</p> <p>Disponível em</p> <p>Propriedades</p> <p>Nome da classe</p> <p>Name space</p> <p>RowType</p>	<p>Medições de traços ou os chamados dados de Caracterização e Avaliação (para recursos genéticos de plantas para alimentação e agricultura (PGRFA). O termo de vocabulário de germoplasma <i>g.germplasmID</i> é um identificador permanente para o material de germoplasma (acesso ao banco de genes, amostra, cultivar, etc.) que foi o objeto do experimento de medição. Por favor, informe identificadores permanentes estáveis e resolvíveis para o termo <i>g.germplasmID</i> ao vincular a pontuação de medição de característica para uma ocorrência de germoplasma descrita em um arquivo externo Darwin Core. Ao vincular a descritores de medição de característica e ensaios de medição descritos externamente ao arquivo Darwin Core, inclua identificadores persistentes estáveis e resolvíveis para <i>g.measurementTraitID</i> e <i>g.measurementTrialID</i> - resolução para informações sobre cada entidade.</p> <p>http://purl.org/germplasm/germplasmTerm</p> <p>18</p> <p>MeasurementScore</p> <p>http://purl.org/germplasm/germplasmTerm#</p> <p>http://purl.org/germplasm/germplasmTerm#MeasurementScore</p>
Trait descriptor (v20140515)	<p>Definição</p> <p>Disponível em</p> <p>Propriedades</p> <p>Nome da classe</p> <p>Name space</p> <p>RowType</p>	<p>Descritores de traços descrevendo métodos e protocolos seguidos ao fazer medições de características ou os chamados dados de Caracterização e Avaliação (C e E) para recursos genéticos de plantas para alimentação e agricultura (PGRFA).</p> <p>http://purl.org/germplasm/germplasmTerm</p> <p>9</p> <p>MeasurementTrait</p> <p>http://purl.org/germplasm/germplasmTerm#</p> <p>http://purl.org/germplasm/germplasmTerm#MeasurementTrait</p>
Trait measurement trial (v20140515)	<p>Definição</p> <p>Disponível em</p> <p>Propriedades</p> <p>Nome da classe</p> <p>Name space</p> <p>RowType</p>	<p>Ensaio de medição (campo ou estufa) para recolher medições de características ou os chamados dados de Caracterização e Avaliação (C e E) para recursos genéticos de plantas para alimentação e agricultura (PGRFA).</p> <p>http://purl.org/germplasm/germplasmTerm</p> <p>10</p> <p>MeasurementTrial</p> <p>http://purl.org/germplasm/germplasmTerm#</p> <p>http://purl.org/germplasm/germplasmTerm#MeasurementTrial</p>

APÊNDICE A – LISTA DE EXTENSÕES ESTÁVEIS DO DARWIN CORE

(Continuação)

TÍTULO	ATRIBUTOS	DEFINIÇÃO
Taxon Description	Definição	DwC Taxon extension para trocar descrições simples de táxons baseados em texto/parágrafo. Não é adequado para descrições e chaves estruturadas, mas é útil para criar páginas de espécies.
	Disponível em	http://rs.gbif.org/extension/gbif/1.0/description.xml
	Propriedades	11
	Nome da classe	Description
	Name space	http://rs.gbif.org/terms/1.0
	RowType	gbif:Description
Species Distribution	Definição	Distribuição geográfica de um táxon.
	Disponível em	http://rs.gbif.org/extension/gbif/1.0/distribution.xml
	Propriedades	14
	Nome da classe	Distribution
	Name space	http://rs.gbif.org/terms/1.0
	RowType	gbif:Distribution
Alternative Identifiers	Definição	Identificadores alternativos para metadados de Taxons, e bases de dados do DwC.
	Disponível em	http://rs.gbif.org/extension/gbif/1.0/identifier.xml
	Propriedades	5
	Nome da classe	Identifier
	Name space	http://rs.gbif.org/terms/1.0
	RowType	gbif:Identifier

APÊNDICE A – LISTA DE EXTENSÕES ESTÁVEIS DO DARWIN CORE

(Continuação)

TÍTULO	ATRIBUTOS	DEFINIÇÃO
Simple Multimedia	Definição	Extensão simples para troca de metadados sobre recursos de multimídia, em particular links para arquivos de imagem, vídeo e áudio.
	Disponível em	http://web.archive.org/web/20080225232621/http://www.dartmouth.edu/comp/about/projects/digitalmedia/images/resources/metadata-schemas.html
	Propriedades	15
	Nome da classe	Multimedia
	Name space	http://rs.gbif.org/terms/1.0
	RowType	gbif:Multimedia
Literature References	Definição	Bibliografia, ou seja, lista de referências bibliográficas. Por exemplo, um táxon ou ocorrência/espécime.
	Disponível em	http://rs.gbif.org/extension/gbif/1.0/references.xml
	Propriedades	13
	Nome da classe	Reference
	Name space	http://rs.gbif.org/terms/1.0
	RowType	gbif:Reference
GBIF Relevé	Definição	Suporte para medições de levantamento de área de vegetação (213eleve) auxiliares àquelas relacionadas usando o Event core com extensão de Ocorrência. Observa-se que todas as medidas de cobertura estão em porcentagem.
	Disponível em	http://rs.gbif.org/extension/gbif/1.0/releve_2016-05-10.xml
	Propriedades	20
	Nome da classe	Releve
	Name space	http://rs.gbif.org/terms/1.0/
	RowType	http://rs.gbif.org/terms/1.0/Releve

APÊNDICE A – LISTA DE EXTENSÕES ESTÁVEIS DO DARWIN CORE

(Continuação)

TÍTULO	ATRIBUTOS	DEFINIÇÃO
Species Profile	Definição	Um perfil básico de espécies com características além da descrição textural que são cobertas pela extensão da descrição.
	Disponível em	http://rs.gbif.org/extension/gbif/1.0/speciesprofile.xml
	Propriedades	14
	Nome da classe	SpeciesProfile
	Name space	http://rs.gbif.org/terms/1.0
	RowType	gbif:SpeciesProfile
Types and Specimen	Definição	Uma extensão para espécimes e tipos, incluindo espécimes tipo, espécie tipo e gêneros tipo e espécimes simples não relacionados a tipos.
	Disponível em	http://rs.gbif.org/extension/gbif/1.0/typesandspecimen.xml
	Propriedades	19
	Nome da classe	TypesAndSpecimen
	Name space	http://rs.gbif.org/terms/1.0
	RowType	gbif:TypesAndSpecimen
Vernacular Names	Definição	Extensão para o <i>core Taxa</i> que lista nomes vernáculos para um táxon científico.
	Disponível em	http://rs.gbif.org/extension/gbif/1.0/vernacularname.xml
	Propriedades	14
	Nome da classe	VernacularName
	Name space	http://rs.gbif.org/terms/1.0
	RowType	gbif:VernacularName

APÊNDICE A – LISTA DE EXTENSÕES ESTÁVEIS DO DARWIN CORE

(Continuação)

TÍTULO	ATRIBUTOS	DEFINIÇÃO
Extended Measurement Or Facts	Definição	Suporte para medições genéricas ou fatos, versão estendida ligada a ocorrências. Essa extensão (eMoF) foi desenvolvida para ser usada em combinação com o <i>Event Core</i> , mas também é compatível com outros núcleos. Quando usado com <i>Event Core</i> , permite criar um link adicional entre o eMoF e a extensão de ocorrência. O eMoF pode armazenar medições ou fatos relacionados a uma ocorrência biológica, medições ou fatos ambientais e atributos do método de amostragem. Essa extensão também oferece a opção de fornecer identificadores para referenciar um vocabulário para os campos <i>measurementType</i> , <i>measurementValue</i> e <i>measurementUnit</i> .
	Disponível em	http://rs.tdwg.org/dwc/terms/index.htm#measureindex
	Propriedades	13
	Nome da classe	ExtendedMeasurementOrFact
	Name space	http://rs.iobis.org/obis/terms/
RowType	http://rs.iobis.org/obis/terms/ExtendedMeasurementOrFact	
Audubon Media Description	Definição	O Audubon Core é um conjunto de vocabulários criados para representar metadados de recursos multimídia e coleções de biodiversidade. Esses vocabulários visam representar informações que ajudarão a determinar se um dado recurso ou coleção estará apto para alguma aplicação específica da ciência da biodiversidade antes de adquirir a mídia. Entre outros, os vocabulários abordam questões como a gestão da mídia e coleções, descrições de seu conteúdo, sua cobertura taxonômica, geográfica e temporal, e as formas apropriadas de recuperá-las, atribuí-las e reproduzi-las.
	Disponível em	http://terms.tdwg.org/wiki/Audubon_Core_Term_List
	Propriedades	96
	Nome da classe	Multimedia
	Name space	http://rs.tdwg.org/ac/terms/
RowType	http://rs.tdwg.org/ac/terms/Multimedia	
Darwin Core Event	Definição	A categoria de informação relativa a um evento de amostragem.
	Disponível em	http://rs.tdwg.org/dwc/terms/index.htm#Event
	Propriedades	95
	Nome da classe	Event
	Name space	http://rs.tdwg.org/dwc/terms/
RowType	dwc:Event	

APÊNDICE A – LISTA DE EXTENSÕES ESTÁVEIS DO DARWIN CORE

(Continuação)

TÍTULO	ATRIBUTOS	DEFINIÇÃO
Darwin Core Identification History	Definição	Suporte para múltiplas identificações/determinações de ocorrências de espécies como espécimes. Todas as identificações, incluindo a atual, devem ser listadas, enquanto a corrente também deve ser repetida no núcleo de ocorrência para acesso simples.
	Disponível em	http://rs.tdwg.org/dwc/terms/index.htm#Identification
	Propriedades	38
	Nome da classe	Identification
	Name space	http://rs.tdwg.org/dwc/terms/
	RowType	dwc:Identification
Darwin Core Measurement Or Facts	Definição	Suporte para medições genéricas ou fatos como definido no Darwin Core.
	Disponível em	http://rs.tdwg.org/dwc/terms/index.htm#measureindex
	Propriedades	9
	Nome da classe	MeasurementOrFact
	Name space	http://rs.tdwg.org/dwc/terms/
	RowType	dwc:MeasurementOrFact
Darwin Core Occurrence	Definição	A categoria de informação relativa à evidência de uma ocorrência na natureza, em uma coleção ou em um conjunto de dados (amostra, observação, etc.). Atualizado em 2 de abril de 2014 com materialSampleID.
	Disponível em	http://rs.tdwg.org/dwc/terms/index.htm#Occurrence
	Propriedades	162
	Nome da classe	Occurrence
	Name space	http://rs.tdwg.org/dwc/terms/
	RowType	dwc:Occurrence

APÊNDICE A – LISTA DE EXTENSÕES ESTÁVEIS DO DARWIN CORE

(Continuação)

TÍTULO	ATRIBUTOS	DEFINIÇÃO
Darwin Core Resource Relationship	Definição	Suporte para relações genéricas de recursos de acordo com os termos do Darwin Core.
	Disponível em	http://rs.tdwg.org/dwc/terms/index.htm#reindex
	Propriedades	7
	Nome da classe	ResourceRelationship
	Name space	http://rs.tdwg.org/dwc/terms/
	RowType	dwc:ResourceRelationship
Darwin Core Taxon	Definição	A categoria de informações referentes a nomes taxonômicos, usos de nomes de taxons ou conceitos de taxon. Atualizado em novembro de 2011 com termos recém-ratificados.
	Disponível em	http://rs.tdwg.org/dwc/terms/index.htm#Taxon
	Propriedades	44
	Nome da classe	Taxon
	Name space	http://rs.tdwg.org/dwc/terms/
	RowType	dwc:Taxon
ChronometricAge	Definição	Extensão para o <i>core</i> Ocorrência para capturar informações sobre a idade cronológica.
	Disponível em	https://github.com/VertNet/dwc-chronology/blob/master/standard/vocabularies/term_versions.csv
	Propriedades	14
	Nome da classe	ChronometricAge
	Name space	http://zooarchnet.org/dwc/terms/
	RowType	http://zooarchnet.org/dwc/terms/ChronometricAge

APÊNDICE A – LISTA DE EXTENSÕES ESTÁVEIS DO DARWIN CORE

(Conclusão)

TÍTULO	ATRIBUTOS	DEFINIÇÃO
ChronometricDate	Definição	Extensão para o <i>core</i> Ocorrência para capturar informações de datação cronométrica.
	Disponível em	https://github.com/VertNet/dwc-chronology/blob/master/standard/vocabularies/term_versions.csv
	Propriedades	14
	Nome da classe	ChronometricDate
	Name space	http://zooarchnet.org/dwc/terms/
	RowType	http://zooarchnet.org/dwc/terms/ChronometricDate

Fonte: elaborado pelo proponente (2019)

Nota: informações traduzidas com algumas adaptações de GBIF (2013)

APÊNDICE B – EXTENSÕES EM DESENVOLVIMENTO

(Continua)

TÍTULO	ATRIBUTOS	DEFINIÇÃO
MlxS Sample	Definição	Termos baseados em amostragem do MlxS estendendo uma clássica <i>Darwin Core Occurrence</i> .
	Disponível em	http://www.gensc.org/gc_wiki/index.php/MlxS
	Propriedades	28
	Nome da classe	Sample
	Name space	http://gensc.org/ns/mixs/terms/
	RowType	http://gensc.org/ns/mixs/terms/Sample
Plinian Distribution Extension	Definição	Definição geográfica das espécies.
	Disponível em	http://purl.org/plic/extension/3.2.1/distribution.xml
	Propriedades	5
	Nome da classe	Distribution
	Name space	http://purl.org/plic/terms/3.2.1/
	RowType	http://purl.org/plic/terms/3.2.1/Distribution
Plinian Endemicity Extension	Definição	Um marcador que indica que um organismo vive exclusivamente em um território específico e não pode ser encontrado em nenhum outro lugar. Lista de áreas estruturadas e categorizadas pelo grau de segurança.
	Disponível em	http://purl.org/plic/extension/3.2.1/endemicity.xml
	Propriedades	2
	Nome da classe	Endemism
	Name space	http://purl.org/plic/terms/3.2.1/
	RowType	http://purl.org/plic/terms/3.2.1/Endemicity

APÊNDICE B – EXTENSÕES EM DESENVOLVIMENTO

(Continuação)

TÍTULO	ATRIBUTOS	DEFINIÇÃO
Plinian Legislation Extension	Definição	Lei proposta a nível nacional ou regional ou grupo de leis.
	Disponível em	http://purl.org/plic/extension/3.2.1/legislation.xml
	Propriedades	6
	Nome da classe	Legislation
	Name space	http://purl.org/plic/terms/3.2.1/
	RowType	http://purl.org/plic/terms/3.2.1/Legislation
Plinian ManagementAnd Conservation Extension	Definição	Gestão: ações direcionadas à conservação ou restauração de espécies. Conservação: intervenções realizadas para preservar espécies.
	Disponível em	http://purl.org/plic/extension/3.2.1/managementAndConservation.xml
	Propriedades	6
	Nome da classe	ManagementAndConservation
	Name space	http://purl.org/plic/terms/3.2.1/Legislation
	RowType	http://purl.org/plic/terms/3.2.1/ManagementAndConservation
Plinian Core Simple Extension	Definição	O Plinian Core é um conjunto de conceitos que define os atributos básicos necessários para integrar e recuperar as informações sobre espécies de organismos requeridas por usuários especializados em biodiversidade, bem como usuários de outras áreas. Esta extensão deve ser usada juntamente com os núcleos de <i>taxon</i> , <i>simple images</i> , <i>vernacular names</i> e <i>references extension</i> .
	Disponível em	http://purl.org/plic/extension/3.2.1/pliniancore.xml
	Propriedades	47
	Nome da classe	Plinian__ore
	Name space	http://purl.org/plic/terms/3.2.1/
	RowType	http://purl.org/plic/terms/3.2.1/PlinianCore

APÊNDICE B – EXTENSÕES EM DESENVOLVIMENTO

(Continuação)

TÍTULO	ATRIBUTOS	DEFINIÇÃO
Plinian Synonym Extension	Definição	Nomes diferentes para este táxon. Esse conceito é um campo de espaço reservado.
	Disponível em	http://purl.org/plic/extension/3.2.1/synonyms.xml
	Propriedades	8
	Nome da classe	Synonyms
	Name space	http://purl.org/plic/terms/3.2.1/
	RowType	http://purl.org/plic/terms/3.2.1/Synonyms
Plinian ThreatStatus Extension	Definição	Informações sobre o estado de conservação do táxon.
	Disponível em	http://purl.org/plic/extension/3.2.1/threatStatus.xml
	Propriedades	2
	Nome da classe	ThreatStatus
	Name space	http://purl.org/plic/terms/3.2.1/
	RowType	http://purl.org/plic/terms/3.2.1/ThreatStatus
Plinian Uses Extension	Definição	Formas em que as espécies são utilizadas pelas pessoas. Incluindo Folclore. Também aplicado a animais e fungos.
	Disponível em	http://purl.org/plic/extension/3.2.1/uses.xml
	Propriedades	16
	Nome da classe	Uses
	Name space	http://purl.org/plic/terms/3.2.1/
	RowType	http://purl.org/plic/terms/3.2.1/Uses

APÊNDICE B – EXTENSÕES EM DESENVOLVIMENTO

(Continuação)

TÍTULO	ATRIBUTOS	DEFINIÇÃO
CoL Name Relations	Definição	Suporte para relações nomenclaturais de nomes a serem usadas em um arquivo DwC do <i>core Taxon</i> , especificando uma relação direcionada do nome do <i>core</i> para um nome relacionado.
	Disponível em	https://github.com/Sp2000/colplus/blob/master/docs/NAMES.md
	Propriedades	5
	Nome da classe	NameRelations
	Name space	http://rs.col.plus/terms/
	RowType	http://rs.col.plus/terms/NameRelations
Invasive Species Distribution	Definição	DwC Taxon extension para trocar dados de distribuições de áreas de espécies invasoras. Baseia-se na extensão de distribuição do GBIF, mas acrescenta invasividade e usa outros vocabulários
	Disponível em	https://tools.gbif.org/dwca-validator/extension.do?id=http://rs.gbif.org/issg/terms/Distribution
	Propriedades	12
	Nome da classe	Distribution
	Name space	http://rs.gbif.org/issg/terms/
	RowType	http://rs.gbif.org/issg/terms/Distribution
Invasive Species Pathways	Definição	DwC Táxon extension para trocar dados sobre rotas de espécies invasoras.
	Disponível em	https://tools.gbif.org/dwca-validator/extension.do?id=http://rs.gbif.org/issg/terms/Pathways
	Propriedades	4
	Nome da classe	Pathways
	Name space	http://rs.gbif.org/issg/terms/
	RowType	http://rs.gbif.org/issg/terms/

APÊNDICE B – EXTENSÕES EM DESENVOLVIMENTO

(Continuação)

TÍTULO	ATRIBUTOS	DEFINIÇÃO
Taxon Concept Relationship	Definição	Uma extensão do <i>Darwin Core Archive</i> (DwCA) para especificar as relações entre conceitos de táxon, com base no elemento 'TaxonRelationshipAssertion' do TDWG <i>Taxon Concept Transfer</i> (XML) <i>Schema</i> . Os tipos de relacionamento incluem tipos de teoria de conjuntos (<i>included_in</i> , etc.) e relacionamentos pais-filhos taxonômicos. Os conceitos de táxon são especificados no módulo 'Taxon' do DwCA, onde um registro deve ter: a) um nome (e idealmente também um <i>scientificNameID</i>), b) um <i>nameAccordingTo</i> (e idealmente também um <i>nameAccordingToID</i>) e c) um <i>taxonConceptID</i> . O último é usado como a chave para <i>subjectTaxonConceptID</i> e <i>objectTaxonConceptID</i> neste módulo.
	Disponível em	http://rs.gbif.org/sandbox/extension/tcrel.xml
	Propriedades	6
	Nome da classe	TaxonConceptRelationship
	Name space	http://rs.gbif.org/sandbox/terms/tcs/tcrel.rdf#
	RowType	http://rs.gbif.org/sandbox/terms/tcs/tcrel.rdf#TaxonConceptRelationship
Chromosomes Count	Definição	Informação sobre o número de cromossomos de uma espécie.
	Disponível em	http://rs.gbif.org/sandbox/extension/chromosomes.xml
	Propriedades	4
	Nome da classe	Chromosomes
	Name space	http://rs.gbif.org/terms/1.0
	RowType	http://rs.gbif.org/terms/1.0/Chromosomes
Ellenbergs Ecological Indicators for European Plants	Definição	O nome abreviado de Ellenberg, <i>Pointer Values</i> , para os "Indicadores Ecológicos de Plantas na Europa Central" é um método de classificação para plantas da Europa Central, descrito em detalhes por Heinz Ellenberg em meados dos anos 70, de acordo com seu "comportamento" ecológico e propriedades botânicas. Os <i>pointer values</i> de acordo com Ellenberg são derivados de observações e experiências ecológicas e botânicas e são agora amplamente confirmados por análises de locais e investigações ecofisiológicas ou parâmetros confirmados para espécies de plantas individuais.
	Disponível em	http://de.wikipedia.org/wiki/Zeigerwerte_nach_Ellenberg
	Propriedades	10
	Nome da classe	Ellenberg
	Name space	http://rs.gbif.org/terms/1.0
	RowType	http://rs.gbif.org/terms/1.0/Ellenberg

APÊNDICE B – EXTENSÕES EM DESENVOLVIMENTO

(Continuação)

TÍTULO	ATRIBUTOS	DEFINIÇÃO
Web Links	Definição	Links de páginas da Web relacionadas. Os links podem ser classificados com palavras-chave (dc: subject) e devem ser usados para páginas da web tenham interfaces amigáveis.
	Disponível em	http://rs.gbif.org/sandbox/extension/links.xml
	Propriedades	5
	Nome da classe	Links
	Name space	http://rs.gbif.org/terms/1.0
	RowType	http://rs.gbif.org/terms/1.0/Links
Extension for mapping NBN eXchange Format (NXF) files to DwC-A	Definição	Extensão NBN Darwin Core do Reino Unido para mapear arquivos NBN eXchange Format (NXF) para DwC-A. Esta extensão é, provavelmente, aplicável apenas para a Rede Nacional de Biodiversidade do Reino Unido e seus fornecedores de dados, e contém algumas peculiaridades específicas do Reino Unido. Esta extensão é um rascunho e deve ser considerada instável.
	Disponível em	http://rs.gbif.org/sandbox/extension/NBNeXchangeFormat.xml
	Propriedades	8
	Nome da classe	nxf
	Name space	http://rs.nbn.org.uk/dwc/nxf/0.1/terms/
	RowType	http://rs.nbn.org.uk/dwc/nxf/0.1/terms/nxfOccurrence
OBIS Marine BioGeography (MBG) Common Terms Extension	Definição	Suporte para propriedades de OBIS MBG como uma extensão para o <i>core Occurrence</i> do DwC. Destina-se a ser uma relação do tipo um-para-um com o <i>core Occurrence</i> .
	Disponível em	http://rs.obis.org/obis/extension/mbg/terms/index.htm#mbgcommontermsindex
	Propriedades	36
	Nome da classe	MBGCommonTerms
	Name space	http://rs.obis.org/obis/extension/mbg/terms/
	RowType	http://rs.obis.org/obis/extension/mbg/terms/MBGCommonTerms

APÊNDICE B – EXTENSÕES EM DESENVOLVIMENTO

(Continuação)

TÍTULO	ATRIBUTOS	DEFINIÇÃO
OBIS Marine BioGeography (MBG) Fish Abundance Extension	Definição	Suporte para propriedades de OBIS MBG <i>Fish Abundance</i> como uma extensão para o <i>core Occurrence</i> do DwC. Destina-se a ser uma relação do tipo um-para-um com o <i>core Occurrence</i> .
	Disponível em	http://rs.obis.org/obis/extension/mbg/terms/index.htm#mbgfishabundanceindex
	Propriedades	18
	Nome da classe	MBGFishAbundance
	Name space	http://rs.obis.org/obis/extension/mbg/terms/
	RowType	http://rs.obis.org/obis/extension/mbg/terms/MBGFishAbundance
OBIS Marine BioGeography (MBG) Visual Line Transect Survey Extension	Definição	Suporte para as propriedades VLTS do MBG OBIS como uma extensão para <i>core Occurrence</i> do DwC. Pretende ser uma relação de um para um com o núcleo de Ocorrência. Destina-se a ser uma relação do tipo um-para-um com o <i>core Occurrence</i> .
	Disponível em	http://rs.obis.org/obis/extension/mbg/terms/index.htm#mbgvltindex
	Propriedades	21
	Nome da classe	MBGVLTS
	Name space	http://rs.obis.org/obis/extension/mbg/terms/
	RowType	http://rs.obis.org/obis/extension/mbg/terms/MBGVLTS
Audubon Media Description	Definição	Esses termos são metadados dependentes de representação, referentes a representações digitais específicas de um recurso.
	Disponível em	http://terms.tdwg.org/wiki/Audubon_Core_Term_List#Service_Access_Point_Vocabulary
	Propriedades	13
	Nome da classe	Multimedia
	Name space	http://rs.tdwg.org/ac/terms/
	RowType	http://rs.tdwg.org/ac/terms/Multimedia#ServiceAccessPointVocabulary

APÊNDICE B – EXTENSÕES EM DESENVOLVIMENTO

(Conclusão)

TÍTULO	ATRIBUTOS	DEFINIÇÃO
Darwin Core Material Sample	Definição	A categoria de informações pertencentes aos resultados físicos de um evento de amostragem (ou subamostragem). Em coleções biológicas, a amostra de material é tipicamente coletada e preservada ou destrutivamente processada. Criado em 2 Abr 2014 com todos os termos ratificados do Simple Darwin Core.
	Disponível em	http://rs.tdwg.org/dwc/terms/index.htm#MaterialSample
	Propriedades	160
	Nome da classe	MaterialSample
	Name space	http://rs.tdwg.org/dwc/terms/
	RowType	dwc:MaterialSample
GISIN Species Status	Definição	Em desenvolvimento.
	Disponível em	https://tools.gbif.org/dwca-validator/extension.do?id=http://www.gisin.org/IASProfile/SpeciesStatus
	Propriedades	34
	Nome da classe	SpeciesStatus
	Name space	http://www.gisin.org/IASProfile
	RowType	http://www.gisin.org/IASProfile/SpeciesStatus

Fonte: elaborado pelo proponente (2019)

Nota: informações traduzidas com algumas adaptações de GBIF (2013)

APÊNDICE C – PESQUISA BIBLIOGRÁFICA

Primeiramente, definiu-se os temas no campo das Ciências Biológicas que seriam fundamentais para o desenvolvimento da pesquisa. Decidiu-se por delimitar o estado da arte dos códigos internacionais de nomenclatura de espécimes, que são esquemas de codificação amplamente recomendados para a descrição de dados no DwC, dos padrões de metadados para descrição da biodiversidade e sobre os princípios da agrobiodiversidade brasileira. Para buscar literatura sobre os temas dos códigos e metadados, foram geradas as seguintes estratégias de busca iniciais: a) *international code of nomenclature*; b) *biodiversity metadata*.

Optou-se por utilizar os termos em inglês pelo fato de a maior parte da literatura buscada ser produzida nesse idioma e, também, os códigos e os padrões de metadados serem editados e atualizados em inglês.

Em seguida, foram efetuadas buscas em diversas bases de dados multidisciplinares e especializadas. Para localizar as bases de dados especializadas, foi utilizado o guia Bases de Dados em Ciências Biológicas⁴⁶, publicado pela Biblioteca Universitária da Universidade Federal de Santa Catarina (BU-UFSC). Todas as bases elencadas no guia estão disponíveis via Portal de Periódicos da Capes. Dentre as bases do guia, foram selecionadas as que apresentam conteúdo específico das Ciências Biológicas, excetuando-se as bases multidisciplinares neste primeiro momento. As bases selecionadas foram:

- a) BioOne;
- b) Ecological Society of America: Journals;
- c) Aquatic Science & Fisheries Abstracts (ASFA) 1: Biological Sciences & Living Resources.

A busca inicial teve alta revocação, pois foram recuperados textos que tratavam sobre a aplicação dos códigos de nomenclatura para descrição de espécies específicas ou outros grupos taxonômicos e esses foram descartados, pois não apresentam o conteúdo pretendido como base para a revisão de literatura. A alta revocação

⁴⁶ Disponível em: <http://portal.bu.ufsc.br/bases-de-dados-por-area-de-conhecimento/base-de-dados-em-ciencias-biologicas/>. Acesso em: 17 jul. 2018.

permaneceu mesmo após a aplicação de filtros de busca, como: a) período de tempo; b) aspas duplas; c) texto completo disponível; d) filtrar por textos em inglês, português ou espanhol. Todavia, a partir da busca inicial foi possível constatar que existem seis códigos de nomenclatura de espécies. Posteriormente, a existência dos seis códigos foi confirmada no site especializado Biocycloedia⁴⁷. Os seis códigos são:

- a) International Code of Nomenclature for algae, fungi, and plants (ICN);
- b) International Code of Nomenclature for Cultivated Plants (ICNCP);
- c) International Code of Nomenclature of Prokaryotes (ICNP);
- d) International Code of Phytosociological Nomenclature (ICPN);
- e) International Code of Zoological Nomenclature (ICZN₂);
- f) International Committee on Taxonomy of Viruses (ICTV code).

Em seguida, a busca foi retomada com os nomes específicos de cada código, nas bases de dados especializadas (mencionadas anteriormente), nos websites de cada um dos códigos e nas seguintes bases multidisciplinares:

- a) Portal de Periódicos da Capes⁴⁸ (busca geral por assunto);
- b) PubMed⁴⁹;
- c) International Information System for the Agricultural Sciences and Technology⁵⁰ (AGRIS);
- d) Academic Search Premier⁵¹ (ASP).

As buscas preliminares com a expressão *international code of nomenclature*

⁴⁷ Disponível em: https://biocycloedia.com/index/nomenclature_codes.php. Acesso em: 15 abr. 2019.

⁴⁸ Disponível em: <http://www-periodicos-capes-gov-br.ez27.periodicos.capes.gov.br/>. Acesso em: 18 jul. 2018.

⁴⁹ Disponível em: <https://www.ncbi.nlm.nih.gov/pubmed/>. Acesso em: 18 jul. 2018.

⁵⁰ Disponível em: <http://agris.fao.org>. Acesso em: 18 jul. 2018.

⁵¹ Disponível em: <https://www.ebsco.com/products/research-databases/academic-search-premier>. Acesso em: 18 jul. 2018.

permitiram encontrar novos termos e vocabulário sobre os códigos de nomenclatura, o que levou à reformulação das expressões de busca.

O *International Code of Nomenclature for algae, fungi, and plants* tem como sinônimo a expressão *Shenzhen code*, que foi utilizada para construir a *query* “*International Code of Nomenclature for algae, fungi, and plants*” OR “*Shenzhen code*”. Para refinar a busca e diminuir a revocação nas bases de dados consultadas, adotou-se o filtro “apenas no título”, a fim de recuperar documentos que contivessem as palavras da *query* em seus títulos. O título é o melhor descritor do conteúdo do documento na maioria das vezes, uma vez que, em geral, refletem o objetivo do estudo. Logo, se as palavras e expressões buscadas estiverem presentes no título, há grande chance de os documentos recuperados serem relevantes à pesquisa. Os resultados das buscas foram semelhantes em todas as bases de dados.

Para buscar bibliografia sobre o *International Code of Nomenclature for Cultivated Plants*, foi elaborada a *query*: “*International Code of Nomenclature for Cultivated Plants*” OR ICNCP OR “*Cultivated Plant Code*”. A base de dados AGRIS apresentou os melhores resultados de busca, o que já era esperado, uma vez que plantas cultiváveis estão presentes nos estudos da Agricultura, um dos assuntos principais da base de dados.

Em seguida, foram efetuadas buscas para levantar literatura sobre o ICZN₂. A *query* foi composta da seguinte forma: ‘*International Code of Zoological Nomenclature*’ OR ‘*ICZN*’ e buscada nas bases de dados.

Como mencionado anteriormente, além das bases de dados científicas, os próprios *websites* dos códigos de nomenclatura foram consultados para obter-se literatura sobre os mesmos. Buscou-se literatura apenas sobre o ICN, ICNCP e ICZN₂, pois os outros códigos de classificação não fazem parte do escopo deste estudo. Outro recurso utilizado para levantamento de bibliografia foi a consulta a especialista do domínio, para indicação de textos sobre o assunto.

Iniciou-se a pesquisa bibliográfica sobre o assunto *biodiversity metadata*. De acordo com o *Digital Curation Centre* (DCC, 2018), três padrões de metadados são utilizados mundialmente para descrever a biodiversidade:

- a) *Darwin Core* (DwC);
- b) *Ecological Metadata Language* (EML);
- c) *Access to Biological Collections Data schema* (ABCD).

A *query* foi então reformulada para permitir a recuperação de bibliografia específica sobre cada padrão de metadados. A busca foi executada nas mesmas bases de dados utilizadas na pesquisa sobre os códigos de nomenclatura. Os sites específicos de cada padrão de metadados também foram fontes de consulta. Os resultados relevantes foram utilizados como insumos para a fundamentação teórica.

Para aprimorar os conhecimentos do campo da sistemática e consolidar as definições de conceitos, foi cursada a disciplina Tópicos em Zoologia II: taxonomia e conservação da biodiversidade, no Programa de Pós-Graduação em Zoologia, do Instituto de Ciências Biológicas da UFMG. Foi ofertada a disciplina Representação da informação sobre Biodiversidade para os cursos Biblioteconomia e Ciências Biológicas da UFMG em parceria com um pesquisador do campo de taxonomia e sistemática. Ambas as disciplinas forneceram bibliografia útil para o referencial teórico desta pesquisa e discussões fundamentais sobre os conceitos da Biodiversidade.

APÊNDICE D – TRADUÇÃO DOS QUADROS 20, 21 E 22

Tradução do Quadro 25

Parasitism Class	
Definition	Interação ecológica desarmônica em que um organismo, denominado parasito, se desenvolve às custas de outro organismo, denominado hospedeiro, prejudicando-o.
Comments	
Examples	Tênias no intestino humano; cipó-chumbo em árvores de laranjeira.

Fonte: elaborado pelo proponente (2019)

Tradução do Quadro 26

parasiteOf Property	
Definition	Ser vivo que sobrevive às custas de outro ser vivo, prejudicando-o.
Comments	A melhor prática recomendada é inserir o nome científico completo da espécie ou subespécie hospedeira, com autoria e data se possível, de acordo com os Códigos Internacionais de Nomenclatura.
Examples	<i>Euschistus heros</i> (Fabricius, 1794)

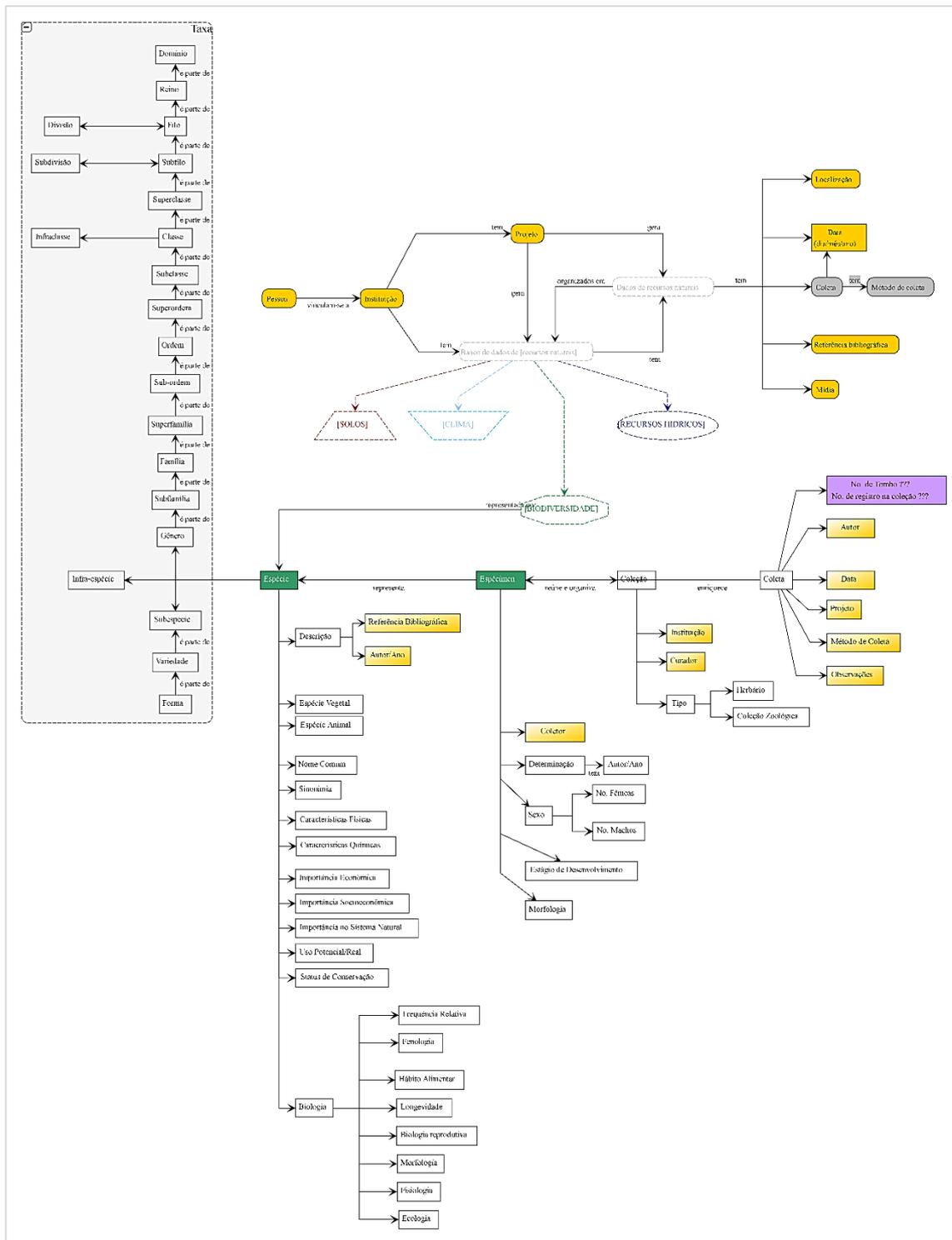
Fonte: elaborado pelo proponente (2019)

Tradução do Quadro 27

hostOf Property	
Definition	Ser vivo que provê abrigo para outro ser vivo.
Comments	A melhor prática recomendada é usar o nome científico completo do parasito, comensal ou mutualista para o qual o ser vivo serve como hospedeiro, com autoria e data se possível, de acordo com os Códigos Internacionais de Nomenclatura.
Examples	<i>Telenomus podisi</i> (Ashmead, 1893)

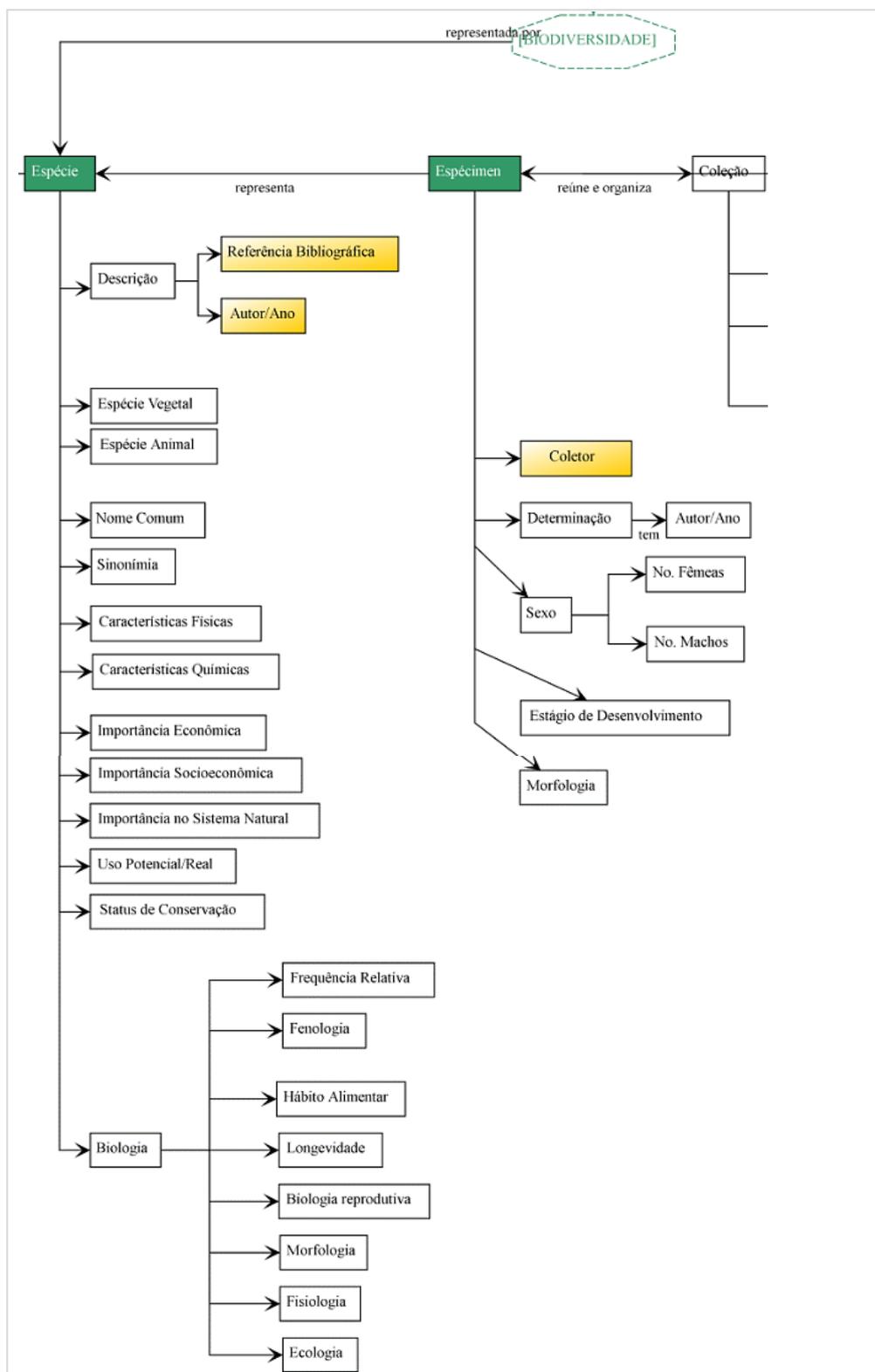
Fonte: elaborado pelo proponente (2019)

ANEXO A – MODELO <ONTO.BIODIVERSIDADE> ESTRUTURA GERAL

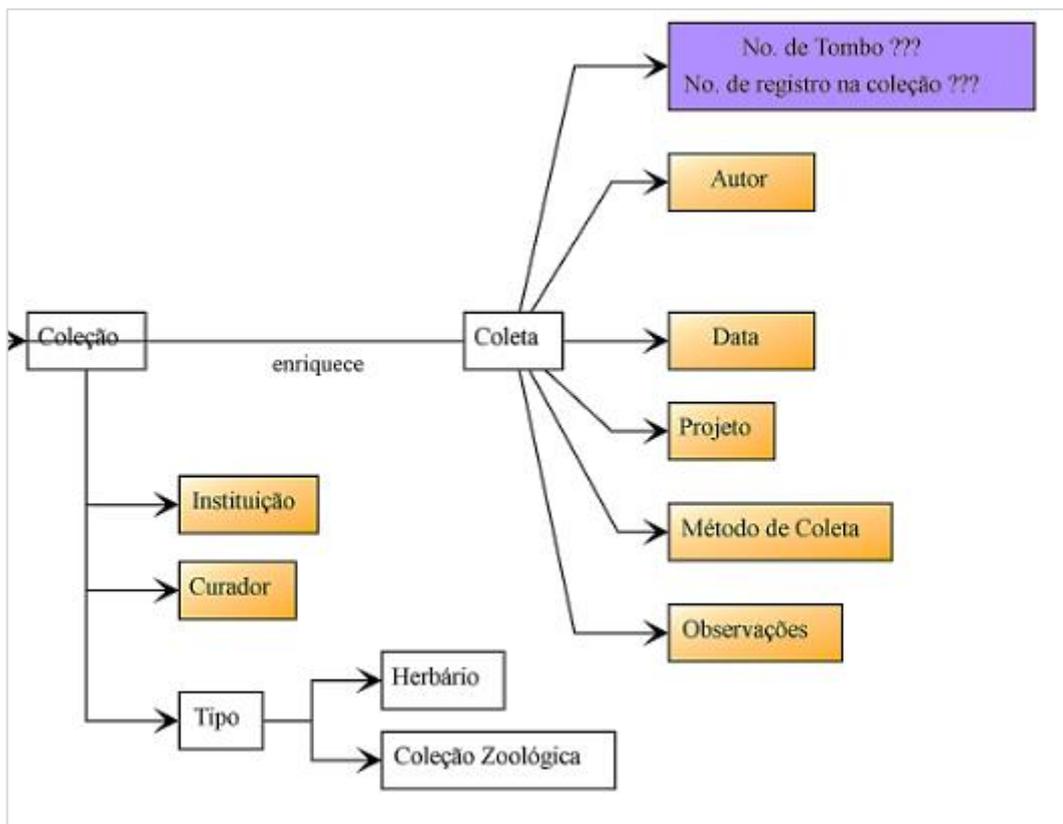


Fonte: Pierozzi Junior (2012)

ANEXO B – MODELO <ONTO.BIODIVERSIDADE> ESPÉCIMEN E ESPÉCIE

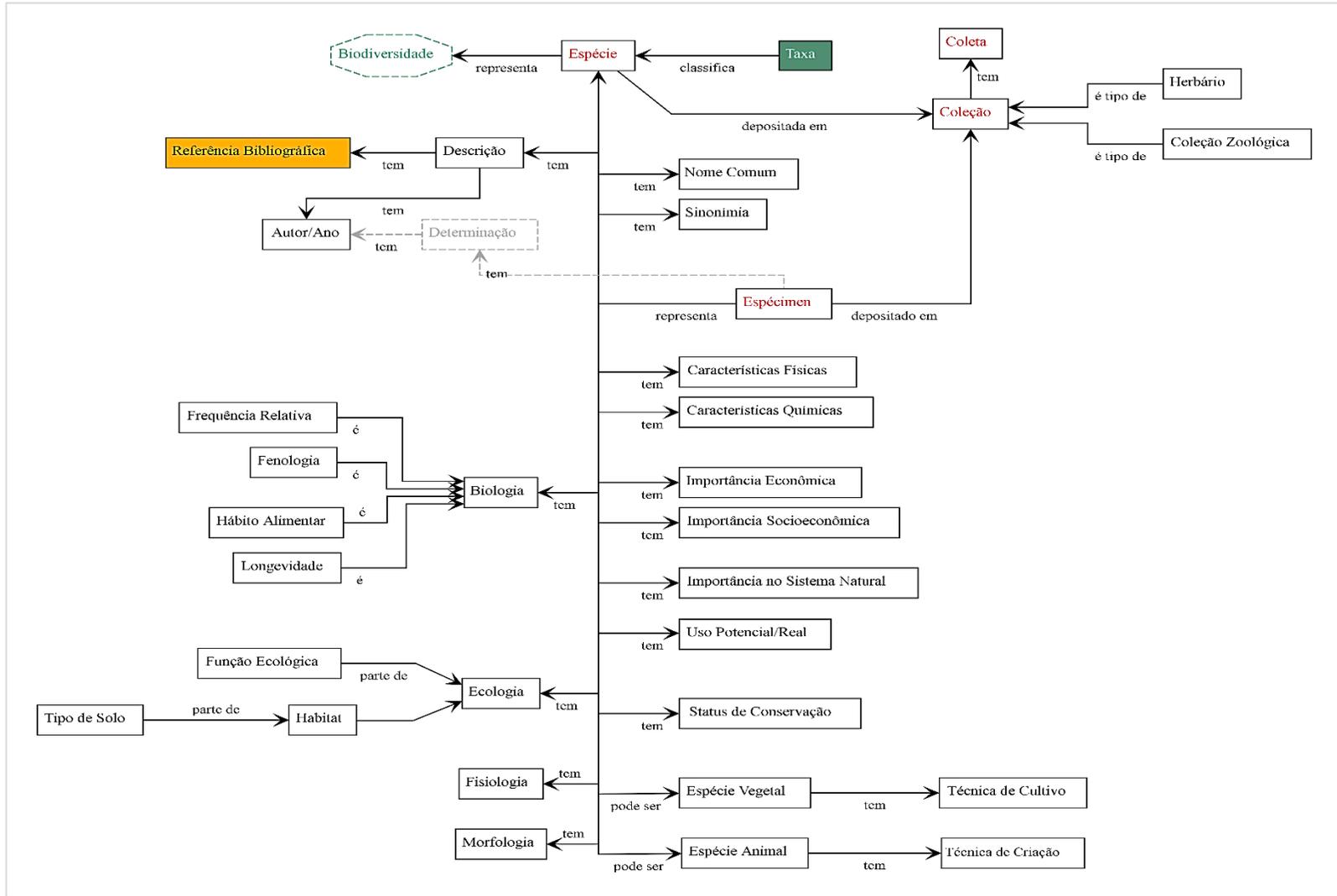


Fonte: Pierozzi Junior (2012)

ANEXO C – MODELO <ONTO.BIODIVERSIDADE> CLASSE COLEÇÃO

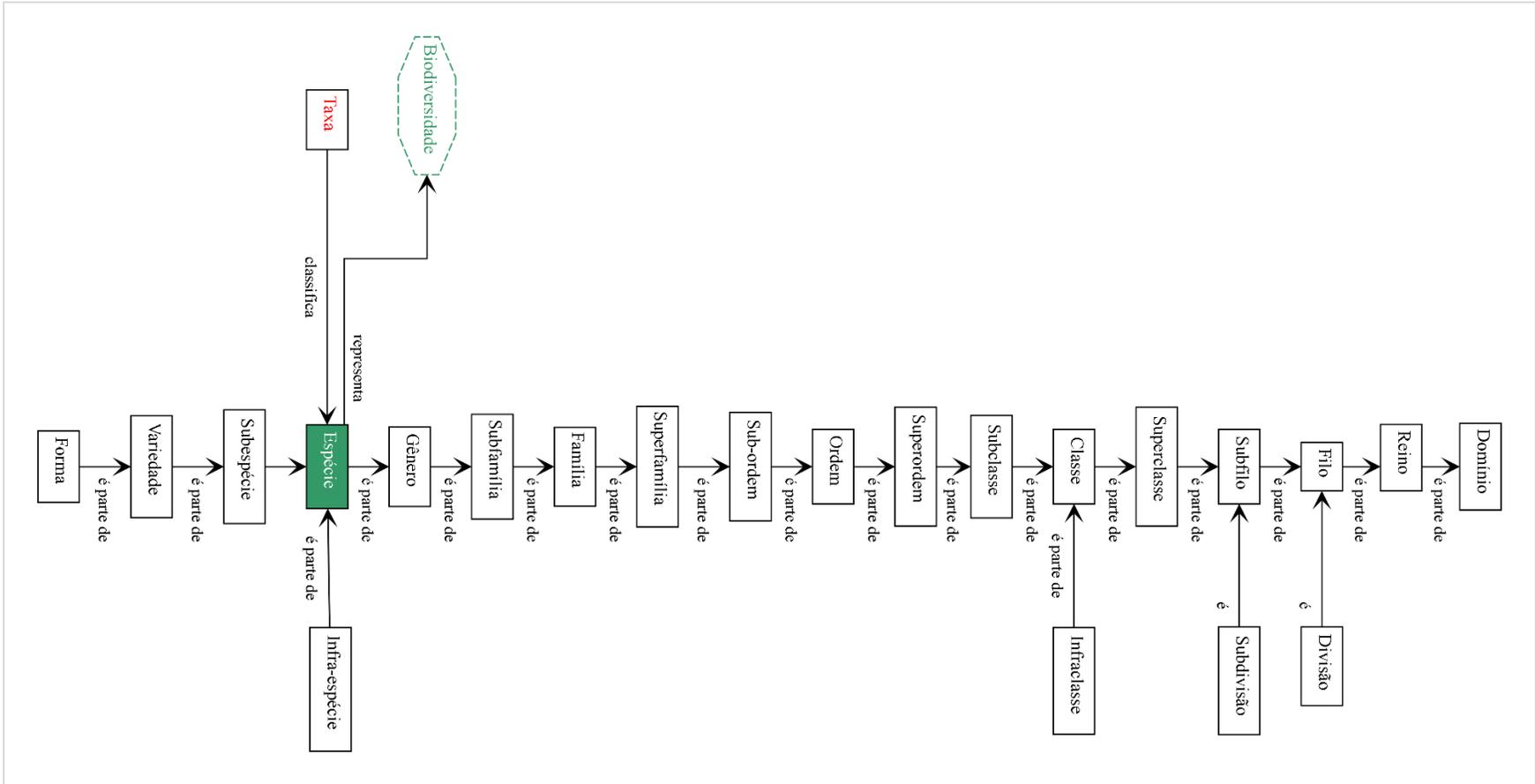
Fonte: Pierozzi Junior (2012)

ANEXO D – MODELO <ONTO.BIODIVERSIDADE> ECOLOGIA



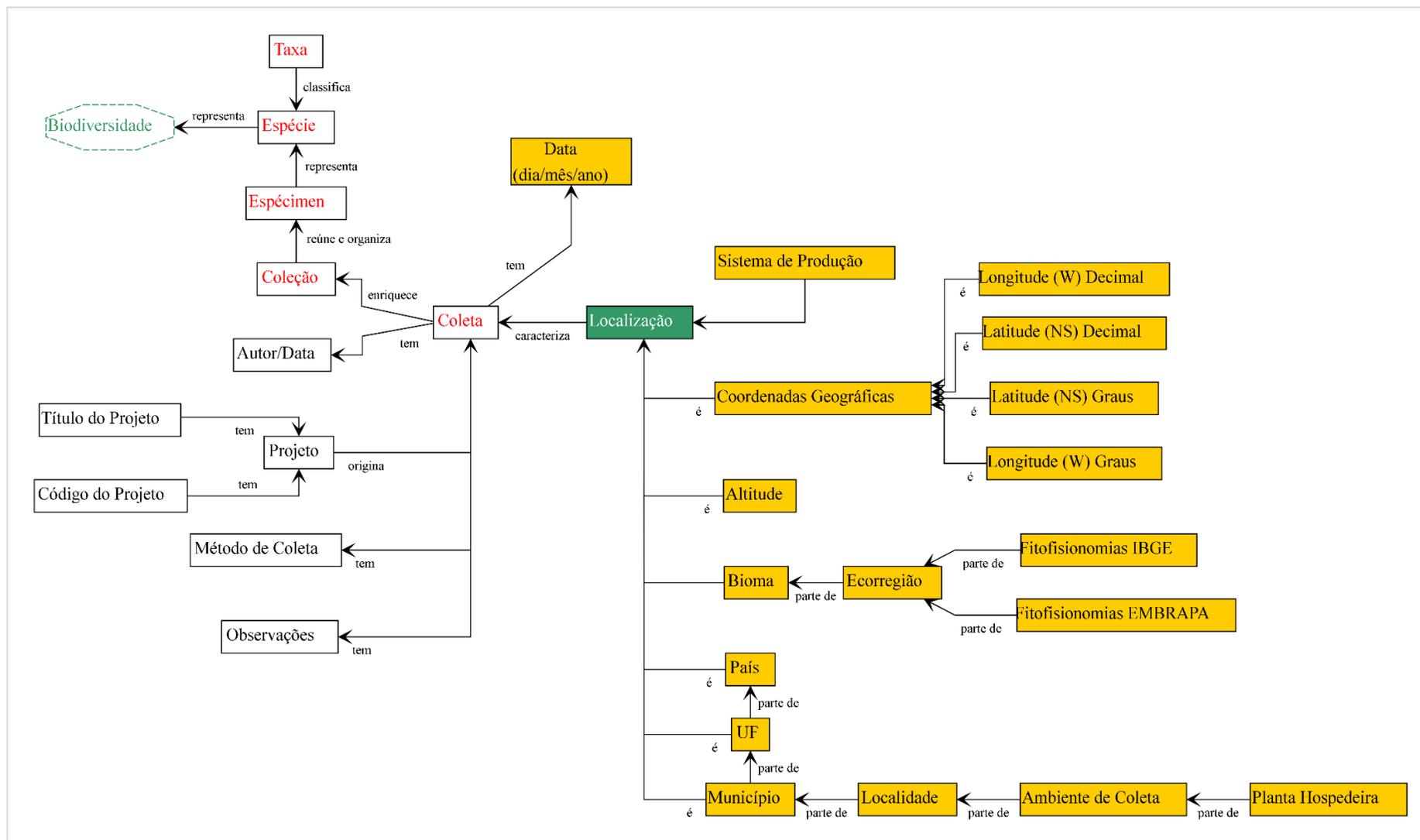
Fonte: Pierozzi Junior (2012)

ANEXO E – MODELO <ONTO.BIODIVERSIDADE> CLASSE TAXA



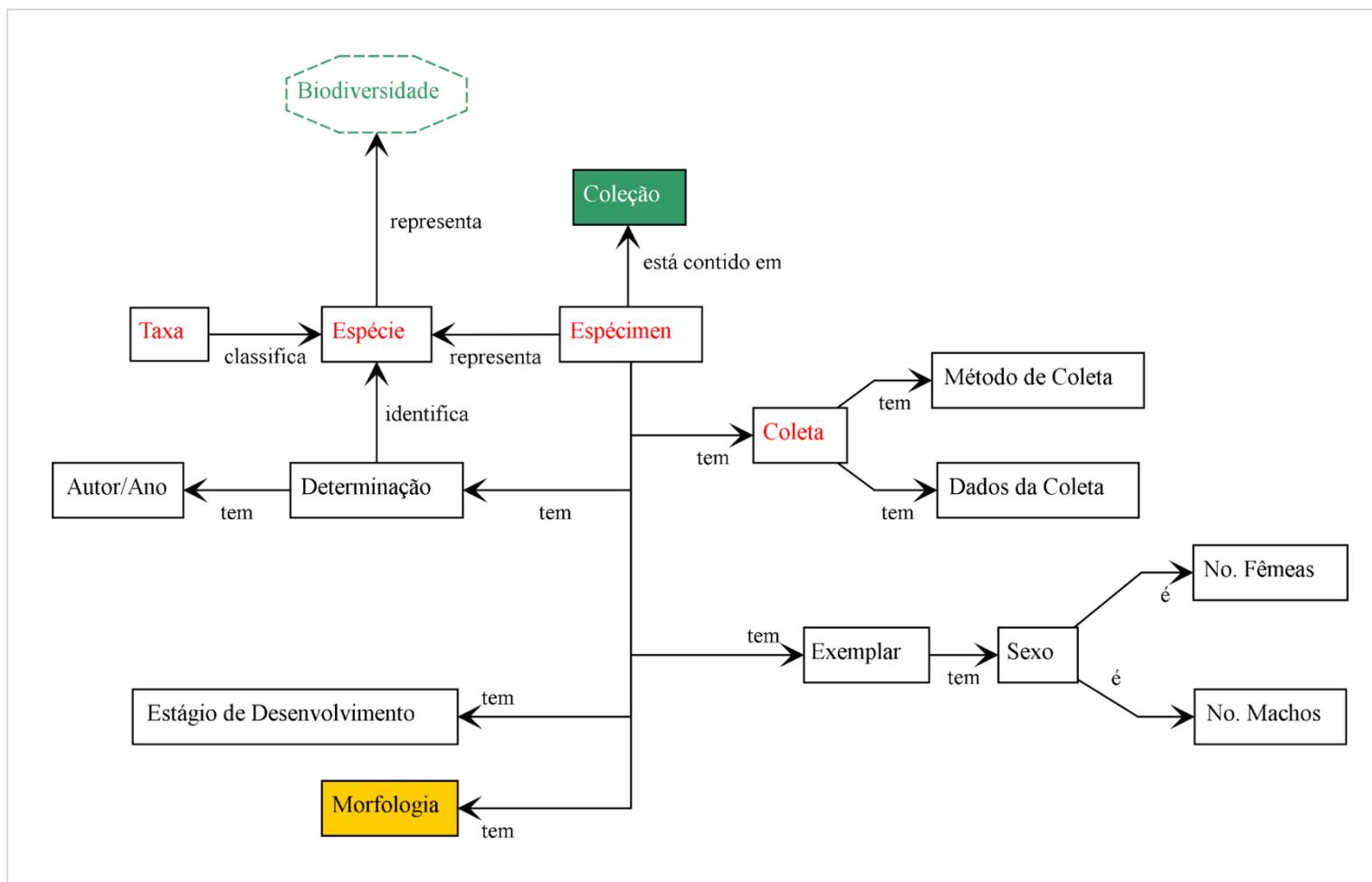
Fonte: Pierozzi Junior (2012)

ANEXO F – MODELO <ONTO.BIODIVERSIDADE> CLASSE COLETA



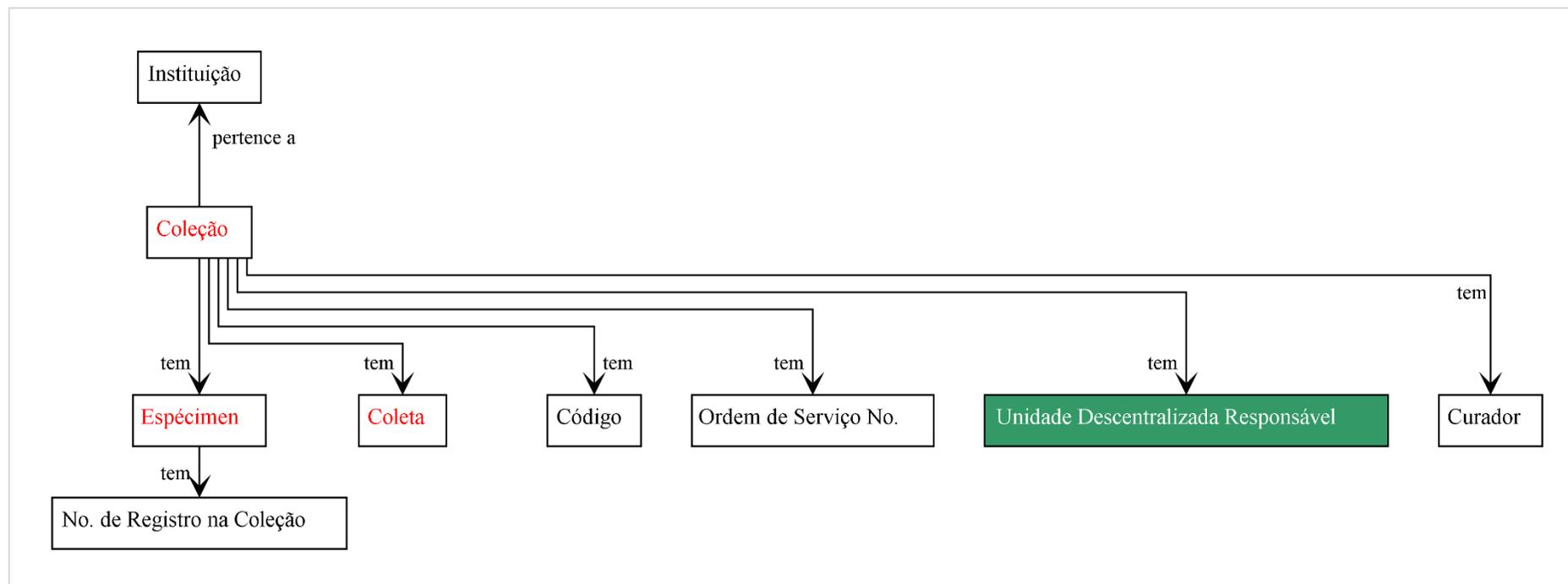
Fonte: Pierozzi Junior (2012)

ANEXO G – MODELO <ONTO.BIODIVERSIDADE> CLASSE ESPÉCIMEN



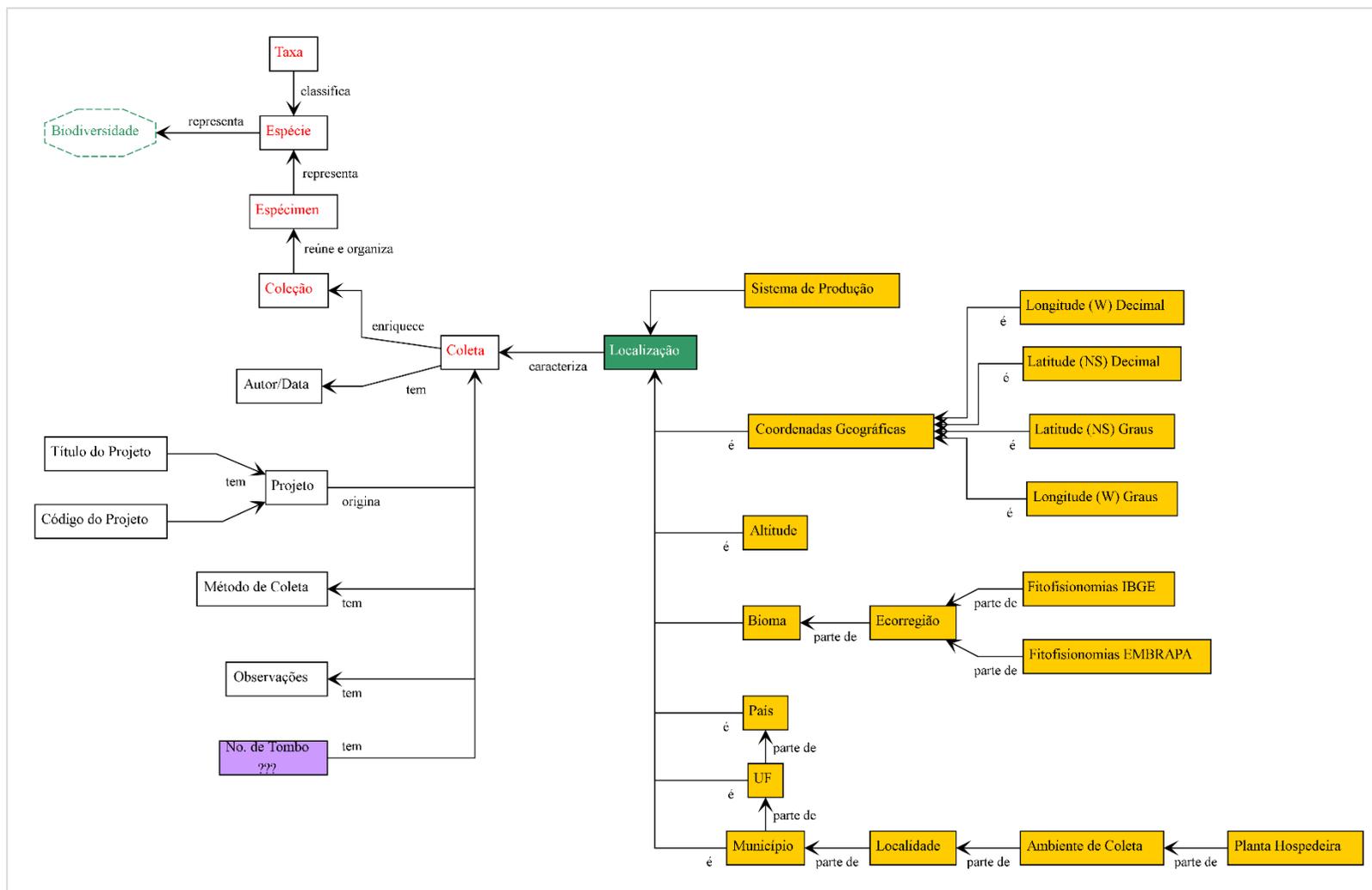
Fonte: Pierozzi Junior (2012)

ANEXO H – MODELO <ONTO.BIODIVERSIDADE> COLEÇÃO



Fonte: Pierozzi Junior (2012)

ANEXO I – MODELO <ONTO.BIODIVERSIDADE> CLASSE LOCALIZAÇÃO



Fonte: Pierozzi Junior (2012)