



ARQUITETURA DE DADOS E MECANISMOS DE PUBLICAÇÃO DE INFORMAÇÃO SOBRE BIODIVERSIDADE

CONSULTOR: JOÃO MONNERAT LANNA – BIÓLOGO MSC (CRBIO 70187/04-D)

Contrato N^o: 002103-2020

Rio de Janeiro-RJ, 20 de novembro de 2022

Versão Interna

Produto elaborado no âmbito do Projeto Pró-Espécies



APRESENTAÇÃO

O acesso eficiente aos dados é fundamental para a tomada de decisões públicas ágeis e assertivas. No contexto da conservação de espécies ameaçadas, o acesso a informações atualizadas, padronizadas e disponíveis representa um desafio para tomadores de decisão que precisam dirigir políticas públicas no sentido de se evitar o declínio das populações ou indivíduos remanescentes.

Este documento detalha as atividades desenvolvidas ao longo de 190 dias, envolvendo reuniões e oficinas participativas com instituições responsáveis pela produção e gestão de dados sobre a biodiversidade. Buscou-se, a partir de um estudo de caso de demanda de dados por parte do Departamento de Espécies do Ministério do Meio Ambiente, desenvolver uma **proposta de padrão de dados** e de uma **arquitetura de repositórios digitais** para publicação baseando-se nos princípios FAIR de gestão e publicação de dados. O estudo de caso “Estratégia Nacional para Conservação de Espécies Ameaçadas de Extinção” serviu de modelo para estas propostas.

Participaram do processo membros do DESP/MMA, CNCFlora/JBRJ, SALVE/ICMBIO e SIBBr/MCTI. A equipe executora deste trabalho envolve dois consultores e uma empresa de moderação especializada. Os consultores são **João Lanna**, contratado pela WWF-Brasil para execução deste contrato; e **Eduardo Dalcin**, servidor do Instituto de Pesquisas Jardim Botânico do Rio de Janeiro (JBRJ/MMA), convidado para participação neste trabalho por meio do [Ofício Nº 6128/2021/MMA](#) e [DESPACHO Nº 3771/2022/GABINETE/JBRJ](#). A facilitação, moderação e documentação das oficinas e reuniões é realizada pela empresa [Vallie](#).

O presente trabalho foi realizado no âmbito do projeto GEF Pró-Espécies: Todos Contra a Extinção. O Pró-Espécies é coordenado pelo Ministério do Meio Ambiente - MMA, por meio do Departamento de Espécies, tendo o Fundo Brasileiro para a Biodiversidade – Funbio como agência implementadora do projeto e o WWF-Brasil como agência executora.



APÊNDICE

1. INTRODUÇÃO	1
1.1. Princípios FAIR e Licenças de Uso	3
2. OFICINAS E ETAPAS DE PREPARAÇÃO DO TRABALHO	5
2.1. Ofertas e Demandas de Dados no MMA.	5
2.2. Sistematização de dados referentes à Estratégia Nacional	7
2.3. Consolidação de Demandas do MMA e Ofertas por Instituições Parceiras	7
2.4. Harmonização da Lista de Termos com Classes e Termos de Padrões Existentes	19
3. PROPOSTA DE PADRÃO DE DADOS PARA ESPÉCIES AMEAÇADAS DE EXTINÇÃO	21
3.1. Acesso e Uso da Proposta de Padrão de Dados MMA	23
4. MECANISMOS DE PUBLICAÇÃO DE DADOS DE BIODIVERSIDADE	24
4.1 Serviços para Máquinas ou Serviços WEB	25
4.2 Repositórios Genéricos	27
4.2.1. Ckan	27
4.2.2. Geonode	28
4.3 Repositórios Temáticos	29
4.3.1 IPT	30
5. CONCLUSÕES E RECOMENDAÇÕES	32
6. REFERÊNCIAS	34

1. INTRODUÇÃO

Espécies estão sendo extintas em uma taxa 1.000 vezes maior que as observadas em registros fósseis, com previsões de aumento destas taxas para 10.000 (DE VOS, *et al.* 2015). A compreensão dos processos que levam estas espécies à extinção é importante para o direcionamento de políticas públicas para conservação (MURRAY, *et al.* 2014).

Este trabalho se enquadra no contexto do uso de dados e informações sobre espécies ameaçadas pelo corpo de Analistas Ambientais do Departamento de Espécies do Ministério do Meio Ambiente (**DESP/MMA**), dados estes produzidos pelas instituições vinculadas ao MMA responsáveis pela avaliação de risco de extinção da fauna e flora: o Sistema de Avaliação do Risco de Extinção da Biodiversidade do Instituto Chico Mendes de Conservação da Biodiversidade (**SALVE/ICMBio**, <https://salve.icmbio.gov.br/#/>) e o Centro Nacional de Conservação da Flora do Jardim Botânico do Rio de Janeiro (**CNCFlora/JBRJ**, <http://cncflora.jbrj.gov.br/portal>), respectivamente.

No contexto do DESP/MMA, o uso de informações sobre biodiversidade é um desafio devido à ausência de uma infraestrutura voltada para a gestão deste tipo de dados. Este cenário torna-se ainda mais evidente pela diversidade de fontes de dados, provindos de um conjunto diverso de instituições, que limita a capacidade de análise e síntese de informações que sejam acessíveis, atuais e padronizadas pelo corpo de analistas ambientais do MMA (LANNA, *et al.* 2021).

O diagnóstico desenvolvido por Lanna, *et al.* (2021) apontou as vantagens da criação de uma **Infraestrutura Ministerial de Gestão de Dados e Informações sobre Biodiversidade** para o cumprimento das missões ministeriais. Neste mesmo diagnóstico foi apresentada uma **Proposta de Plano de Ação** para gestão e integração de dados e informações sobre biodiversidade no MMA e instituições vinculadas, com um conjunto de objetivos, metas e ações baseadas nas principais demandas observadas no diagnóstico (Figura 1).

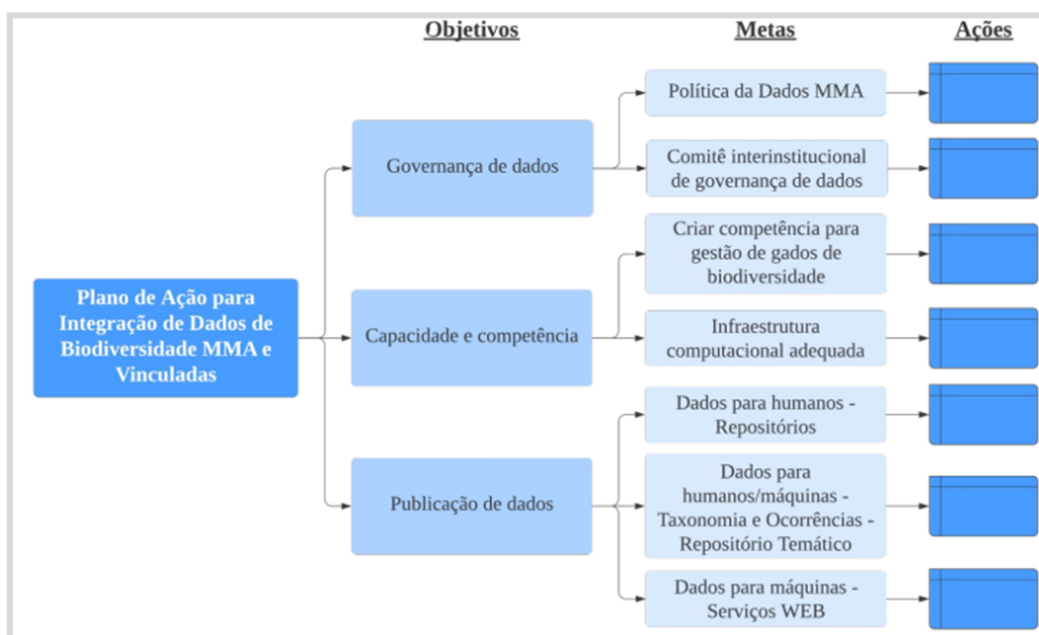


Figura 1. Proposta de Plano de Ação para Gestão e Integração de Dados e Informações sobre Biodiversidade no MMA e Instituições Vinculadas desenvolvido por Lanna, *et al.* (2021).

Este trabalho visa aplicar um conjunto de ações incluídas nos objetivos de Melhorias de Capacidade e Competência, bem como na Eficiência na Publicação de Dados pelo MMA e vinculadas, no contexto de dados provenientes das avaliações de risco de extinção de espécies no âmbito nacional. Dados estes que provêm das instituições vinculadas ao MMA responsáveis pela avaliação de risco de extinção da fauna e flora, SALVE/ICMBio e CNCFlora/JBRJ, respectivamente.

Neste contexto, foi desenvolvida uma estratégia de trabalho participativo envolvendo membros do DESP/MMA, CNCFlora/JBRJ e SALVE/ICMBio, responsáveis pela gestão de dados e informações. Este trabalho envolveu um conjunto de oficinas que, ao longo de 190 dias, buscaram produzir uma proposta de padrão de dados e mecanismos de publicação a partir da escolha de um estudo de caso de uso de dados de espécies ameaçadas pelo time de analistas ambientais do DESP/MMA. Ressalta-se a relevância neste trabalho do nó brasileiro do *Global Biodiversity Information Facility* (GBIF, <https://www.gbif.org/>), o Sistema de Informação sobre a Biodiversidade Brasileira (SiBBR, <https://www.sibbr.gov.br/>), cujos membros participaram conjuntamente ao longo das oficinas e da execução deste trabalho.

O **Plano de Trabalho**, bem como toda a **documentação relacionada e produtos** desta atividade, podem ser acessados no **GitHub** deste trabalho (<https://github.com/jlanna/MMA-Ameacadas>).

1.1. Princípios FAIR e Licenças de Uso

Os princípios FAIR foram elaborados em 2014 por um grupo heterogêneo de pessoas do setor acadêmico e da iniciativa privada em um workshop em Leiden, na Holanda, para tratar dos obstáculos na busca e reutilização de dados, não somente de biodiversidade. Este grupo deliberou que o descobrimento, acesso, integração e reuso apropriado, assim como uma correta citação dos dados deveria ser baseada em uma série de acordos comunitários de princípios e práticas vastamente aceitos (Wilkinson et al., 2016).

A iniciativa está sendo implementada globalmente por meio do [GO FAIR](#). No Brasil, o escritório [GO FAIR Brasil](#) está implementando diversas frentes e aguarda o lançamento de uma frente para a biodiversidade.

Os princípios FAIR resumem muito bem as discussões sobre gestão e compartilhamento de dados que vêm sendo feitas nas últimas décadas e que direcionam os princípios da Informática da Biodiversidade, sobre os quais este trabalho se pauta (Bisby, 2000; Gadelha et al., 2020; Soberón & Peterson, 2004). Estes princípios indicam que os dados devem ser, da sigla em inglês, *Findable* (Buscáveis), *Accessible* (Acessíveis), *Interoperable* (Interoperáveis) e *Reusable* (Reutilizáveis), tanto para pessoas quanto para máquinas. A Figura 2 ilustra os principais pontos relacionados à gestão de dados seguindo os princípios FAIR.

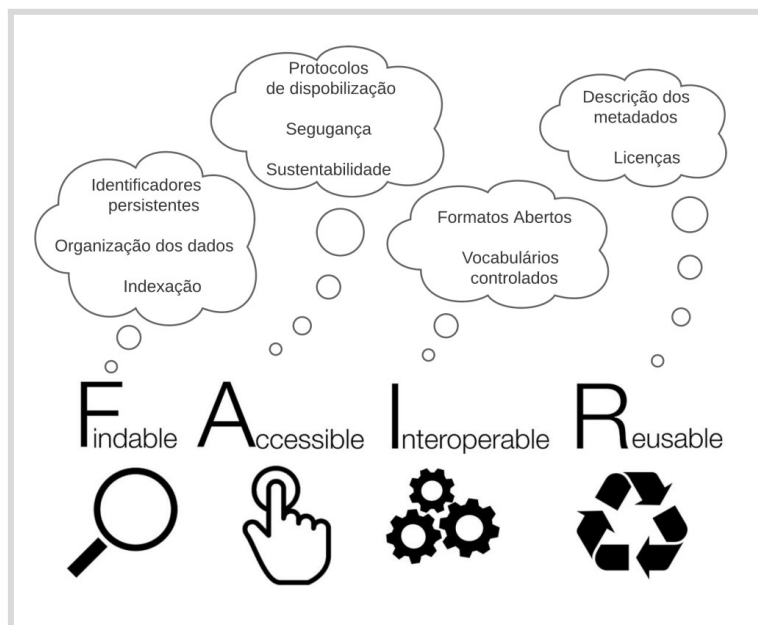


Figura 2. Princípios FAIR de compartilhamento de dados.

Uma parte considerável destes pontos será tratada ao longo deste trabalho. Destaca-se, no entanto, o princípio da aplicação de licenças de uso que funcionam como instrumentos de direito de autor. Neste contexto, aplicando os princípios FAIR de dados abertos, destacam-se as licenças de uso [Creative Commons](https://creativecommons.org/) (CC), criadas no sentido de buscar um equilíbrio nos selos tradicionais de “todos direitos reservados”. Desta forma, as licenças CC fornecem a autores individuais ou empresas e instituições, uma forma padronizada de atribuir autorizações de direito de autor aos seus trabalhos criativos.

2. OFICINAS E ETAPAS DE PREPARAÇÃO DO TRABALHO

Neste tópico são descritas as etapas de execução do trabalho que culminou na construção de uma proposta de Padrão de Dados e de Mecanismos para Espécies Ameaçadas e Mecanismos de Publicação de dados de espécies ameaçadas no contexto do MMA. Ressalta-se que a facilitação, moderação e documentação das oficinas e reuniões foi realizada pela empresa [Vallie](#), contratada pela [WWF-Brasil](#) para contribuir na execução deste trabalho.

2.1. Ofertas e Demandas de Dados no MMA.

Previamente ao início das atividades deste trabalho foram levantados recursos de informação de biodiversidade que são publicados por instituições nacionais, vinculadas ou não ao MMA (Tabela 1). Estes recursos foram baixados e tiveram seus campos mapeados na ferramenta de gestão de planilhas seguindo estrutura relacional [Grist](#), gerando a **Lista de Termos dos Recursos**, que será tratada mais adiante e está disponível na página de produtos no [GitHub](#) deste trabalho.

A primeira oficina deste trabalho correspondeu à **Oficina de Demandas por Dados de Espécies Ameaçadas do DESP/MMA**, que ocorreu no dia 23/07/2022. Esta foi planejada visando uma troca de experiências entre a equipe de consultores e o time do DESP/MMA, de maneira que todos os integrantes do Departamento foram convidados.

Tabela 1. Recursos de informação de biodiversidade levantados previamente à realização da primeira oficina deste trabalho, Oficina de Demandas de Dados do DESP/MMA.

INSTITUIÇÃO	RECURSO	LINK
CNCFlora/JBRJ	Conjunto de dados SIS Connect IUCN	https://cloud.jbrj.gov.br/s/AHfKz3PK5aPsYb2
	API do CNCFlora	https://github.com/edalcin/MMA-Ameacadas/tree/main/recursos/cncflora/api
Catálogo da Fauna/MCTI	API do Catálogo da Fauna	https://github.com/edalcin/MMA-Ameacadas/tree/main/recursos/fauna/api
Flora de Funga do Brasil/JBRJ	API da Flora e Funga do Brasil	https://servicos.jbrj.gov.br/v2/flora/
	IPT da Flora e Funga do Brasil	http://ipt.jbrj.gov.br/jbrj/resource?r=lista_especies_flora_brasil
SiBBr/MCTI	API do SiBBr	https://github.com/edalcin/MMA-Ameacadas/blob/main/recursos/sibbr/atlas/species.md

O principal produto desta oficina foi um conjunto de estudos de casos de demandas por dados do DESP/MMA, de onde um estudo de caso seria selecionado como foco nas etapas seguintes deste trabalho. Três estudos de caso foram apresentados na oficina pela equipe do DESP/MMA: 1 – CITES – propostas de inclusão de espécies nos anexos da Convenção; 2 – [Análise de Efetividade da Estratégia Nacional de Conservação de Espécies Ameaçadas de Extinção \(EN\)](#); 3 – Planos de Recuperação de Espécies Aquáticas Ameaçadas. Cada um destes estudos de caso foi detalhado com relação aos dados demandados para sua construção, bem como os principais desafios no levantamento destes dados pela equipe do DESP/MMA.

Posteriormente à oficina, seguido de avaliações dos conjuntos de dados envolvendo cada uma dos estudos de casos tratados e reuniões com membros do DESP responsáveis pela coordenação deste trabalho, **optou-se por adotar a EN como estudo de caso**. Principalmente pela EN envolver o conjunto mais amplo de dados relacionados a espécies ameaçadas que, se geridos adequadamente, podem por sua vez suprir diversas demandas de consumo de dados do DESP/MMA.

2.2. Sistematização de dados referentes à Estratégia Nacional

Posteriormente à oficina foram reunidos documentos tabulados contendo os dados relacionados à EN. Foram agrupadas três planilhas que reúnem os dados utilizados para a construção da EN, cujos nomes seguem: Dados_BI_especiesameacadas_20.02.2020_paraequipe.csv (Dados BI); Estratégia Nacional - níveis de proteção.csv (Níveis de Proteção), Tabelão estratégia_ATUALIZADA Pedro_29.08.2019_bi.csv (Tabelão EN).

Estes dados foram agrupados e adicionados à ferramenta **Grist**, gerando a **Lista de Termos Gerados na Oficina de Demandas**, que serviu de subsídio para a próxima etapa deste trabalho. Esta lista está disponível na página de produtos no [GitHub](#) deste trabalho.

2.3. Consolidação de Demandas do MMA e Ofertas por Instituições Parceiras

Nesta etapa do trabalho foi realizada, na data 11/07/2022, a **Oficina de Consolidação de Termos Demandados e Presentes nos Recursos**. Visou-se tratar dos dados produzidos pelas instituições vinculadas ao MMA, responsáveis pela avaliação de risco de extinção da fauna e flora, SALVE/ICMBio e CNCFlora/JBRJ, respectivamente. Foram convidados os membros destas instituições envolvidos em papéis técnicos ou de coordenação de atividades envolvendo gestão de informação e/ou desenvolvimento de sistemas de informação sobre biodiversidade. Também foram convidados os membros do SIBBr e os membros do DESP/MMA, participantes da primeira oficina.

O principal objetivo desta oficina foi discutir e validar as listas de termos relativos à EN compiladas durante e após a primeira oficina. A **Lista de Termos dos Recursos**, contendo os termos presentes nos recursos de informação publicados pelas instituições vinculadas ao MMA e a **Lista de Termos Gerada na Oficina de Demandas**, contendo termos relativos à EN, já sistematizados.

Estas listas foram comparadas entre si buscando medir a sobreposição entre a atual oferta de dados pelas instituições vinculadas com as demandas ministeriais. Observou-se que 15 dos 28 termos presentes na **Lista de Termos Gerada na Oficina de Demandas** não possuem conteúdo relacionado nos recursos de informação atualmente ofertados e, portanto, precisam ser sistematizados e publicados (tabela 2). Ressalta-se que os dados de avaliações de risco de extinção relacionados à fauna não estão publicados sob nenhum mecanismo até o momento.

Nos casos em que as demandas por dados relativos à EN não estão sendo supridas pelas instituições vinculadas ou outra instituição, realizou-se uma discussão direcionando ações de prospecção e/ou foram assumidos compromissos de publicação destes dados pelas instituições presentes na oficina. A tabela 2 apresenta os compromissos de publicação de dados assumidos por cada instituição presente na oficina.

A Oficina de Consolidação de Termos Demandados e Presentes nos Recursos teve como produtos duas listas de termos sistematizadas de acordo com sua adequação aos padrões de dados usuais. A **Lista de Termos Preliminares**, que contém os termos cujo conteúdo se adequa aos padrões de dados de biodiversidade usuais e a **Lista de Termos que Necessitam de Sistematização**, contendo os termos que não se adequam a estes padrões de dados. Estas listas estão disponíveis na página de produtos no [GitHub](#) deste trabalho.

Tabela 2. A comparação da **Lista de Termos Gerada na Oficina de Demandas** com a **Lista de Termos dos Recursos** gerou esta tabela que indica as carências de dados relativos ao estudo de caso da EN. As responsabilidades de publicação de dados assumidas durante a **Oficina de Consolidação de Termos Demandados e Presentes nos Recursos** também estão detalhadas na última coluna.

Termo	Descrição MMA	Domínio	Disponibilidade	Responsável
ID_taxon	Identificados único do táxon		disponível	CNCFlora/JBRJ e SALVE/ICMBio
Fauna/Flora / Grupão	Nome vernacular relativo ao reino biológico ao qual o táxon pertence. Campo de vocabulário controlado	Fauna; Flora	disponível	CNCFlora/JBRJ e SALVE/ICMBio
Grupo / Grupo taxonômico	Nome do grupo biológico ao qual o táxon pertence segundo critérios do MMA. Campo de vocabulário controlado	Anfibios; Angiospermas; Aves; Briófitas; Gimnospermas; Invertebrados Aquáticos; Invertebrados Terrestres; Mamíferos; Peixes Continentais; Peixes Marinhos; Pteridófitas; Répteis	disponível	CNCFlora/JBRJ e SALVE/ICMBio
Família			disponível	CNCFlora/JBRJ e SALVE/ICMBio

Termo	Descrição MMA	Domínio	Disponibilidade	Responsável
Espécie (Simplificado) / Nome científico	Nome científico sem autor		disponível	CNCFlora/JBRJ e SALVE/ICMBio
Nome Comum			disponível	CNCFlora/JBRJ e SALVE/ICMBio
Categoria de Ameaça	Nome da categoria de ameaça por extenso		disponível	CNCFlora/JBRJ e SALVE/ICMBio
Sigla Categoria de Ameaça			disponível	CNCFlora/JBRJ e SALVE/ICMBio
Critério			disponível	CNCFlora/JBRJ e SALVE/ICMBio
Bioma	Bioma brasileiro em que a espécie ocorre.	Amazônia; Caatinga; Cerrado; Ilha Oceânica; Marinho; Mata Atlântica; Pampa; Pantanal	disponível	CNCFlora/JBRJ e SALVE/ICMBio
Principais Ameaças			necessita sistematizar	CNCFlora/JBRJ e SALVE/ICMBio
Presença em Áreas Protegidas	Presença ou ausência em áreas protegidas	Sim; Não	necessita sistematizar	CNCFlora/JBRJ e SALVE/ICMBio

Termo	Descrição MMA	Domínio	Disponibilidade	Responsável
Plano de Ação Nacional para Conservação (PAN)			necessita sistematizar	CNCFlora/JBRJ e SALVE/ICMBio
Ordenamento Pesqueiro			necessita sistematizar	DESP/MMA
Nível de Proteção na Estratégia Nacional / NÍVEL PÓS-OFICINA	Os Níveis de Proteção variam de 0 a 5, em que 0 demonstra ausência de medidas de conservação e 5 demonstra que a espécie está suficientemente coberta por medidas de conservação		necessita sistematizar	DESP/MMA
Espécie exclusiva do Brasil		Sim; Não; Informação não disponível	disponível	CNCFlora/JBRJ e SALVE/ICMBio
Estados de Ocorrência			disponível	CNCFlora/JBRJ e SALVE/ICMBio
Fonte da Imagem			necessita sistematizar	DESP/MMA
Créditos	Autor da imagem		necessita sistematizar	DESP/MMA

Termo	Descrição MMA	Domínio	Disponibilidade	Responsável
Medida: Área distribuição UC+TI+AM	Esta medida se refere ao grau de proteção de habitat da espécie em Unidades de Conservação, Terras Indígenas e Áreas Militares. Os cálculos utilizados nesta medida consideram o grau de efetividade de cada área protegida calculado pelo SAMGe/ICMBio e o grau de sobreposição da área de ocorrência da espécie com estas áreas protegidas. Detalhes do cálculo encontram-se no documento intitulado "Descricao_metodologia_analiseefetividade_no v18".	0; 1; 2; 3	necessita sistematizar	UC + TI = CNCFlora/JBRJ e SALVE/ICMBio; AM = DESP/MMA (shapes), CNCFlora/JBRJ e SALVE/ICMBio (sobreposição de ocorrências).
Peso da Medida: Área distribuição UC+TI+AM	Peso proporcional ao potencial de contribuição da medida para melhorar o status de conservação da espécie	2,5	peso (não necessita sistematizar)	DESP/MMA

Termo	Descrição MMA	Domínio	Disponibilidade	Responsável
Medida: Caverna	<p>As análises espaciais de ocorrência das espécies em cavernas foram descritas no item anterior. A lista das espécies ameaçadas que ocorrem em cavernas foi fornecida pelo CECAV/ICMBio.</p> <p>A sugestão da reunião de especialistas foi considerar com um peso maior as cavernas classificadas como de máxima relevância. Entretanto, a ocorrência de espécies ameaçadas é um dos critérios que classifica a caverna como de máxima relevância. Portanto, todas as cavernas deste exercício são necessariamente de máxima relevância.</p> <p>Devido à dificuldade de determinar a extensão das cavernas, não foi possível calcular a área de distribuição das espécies que está coberta por cavernas, e sim presença e ausência. Dessa forma, foi atribuída Classe 2 neste critério para todas as espécies analisadas que apresentaram sobreposição com cavernas.</p>	0; 2	necessita sistematizar	SALVE/ICMBio
Peso da Medida: Caverna	Peso proporcional ao potencial de contribuição da medida para melhorar o status de conservação da espécie	2	peso (não necessita sistematizar)	DESP/MMA

Termo	Descrição MMA	Domínio	Disponibilidade	Responsável
Medida: PAN	<p>A efetividade dos PAN foi avaliada de acordo com seu grau de implementação:</p> <ul style="list-style-type: none"> • Classe 1: espécies que estão em PAN não avaliado; espécies em PAN encerrados, aguardando novo ciclo; espécies em PAN com menos de 40% de implementação; • Classe 2: espécies que estão em PAN com grau de implementação igual ou maior que 40%; • Classe 3: espécies em PAN encerrado que atingiu seu objetivo e não necessita de novo ciclo. 	0; 1; 2; 3	necessita sistematizar	CNCFlora/JBRJ e SALVE/ICMBio
Peso da Medida: PAN	Peso proporcional ao potencial de contribuição da medida para melhorar o status de conservação da espécie	2	peso (não necessita sistematizar)	DESP/MMA
Medida: Ordenamento pesqueiro	<p>A equipe do DESP/MMA realizou um levantamento das normas de ordenamento pesqueiro vigentes, considerando também moratórias e planos de recuperação. O critério considerado foi o seguinte:</p> <ul style="list-style-type: none"> • Classe 1: espécie é indiretamente beneficiada pela medida; • Classe 2: ordenamento direcionado às práticas que afetam diretamente as espécies • Classe 3: moratória ou citada nominalmente no ordenamento (exclusivo exemplo portaria da moratória da Piracatinga) 	0; 1; 2; 3	necessita sistematizar	DESP/MMA

Termo	Descrição MMA	Domínio	Disponibilidade	Responsável
Peso da Medida: Ordenamento pesqueiro	Peso proporcional ao potencial de contribuição da medida para melhorar o status de conservação da espécie	1	peso (não necessita sistematizar)	DESP/MMA
Medida: Portaria 443	<p>A Portaria MMA nº 443/2014, que estabelece a lista das espécies ameaçadas da flora, foi considerada como uma medida de conservação pelo fato de proibir coleta, corte, transporte, armazenamento, manejo, beneficiamento e comercialização das espécies listadas. Todavia, o Art. 3º permite uma exceção para as espécies da categoria Vulnerável (VU), desde que regulamentada. As espécies madeireiras desta categoria que estão no bioma Amazônia foram regulamentadas, o que permite seu corte. Portanto, essas espécies receberam uma pontuação menor nesse critério:</p> <ul style="list-style-type: none"> • Classe 1: espécies madeireiras, da categoria VU, que ocorrem no bioma amazônico; • Classe 2: demais espécies da flora, listadas na Portaria MMA nº 443/2014. 		necessita sistematizar	DESP/MMA
Peso da Medida: Portaria 443	Peso proporcional ao potencial de contribuição da medida para melhorar o status de conservação da espécie	2	peso (não necessita sistematizar)	DESP/MMA

Termo	Descrição MMA	Domínio	Disponibilidade	Responsável
Medida: Portaria 445	<p>A inclusão da Portaria MMA nº 445/2014 segue o mesmo princípio que o da Portaria MMA nº 443/2014, considerando que a lista das espécies de peixes e invertebrados aquáticos ameaçados também oferece proteção diferenciada a estas espécies em relação às espécies não listadas considerando as restrições de uso previstas. O critério considerado foi o seguinte:</p> <ul style="list-style-type: none"> • Classe 1: espécies com interesse socioeconômico que se beneficiariam com a Portaria MMA 445/2014. • Classe 2: CR e EN de interesse econômico, com captura proibida. 		necessita sistematizar	DESP/MMA
Peso da Medida: Portaria 445	Peso proporcional ao potencial de contribuição da medida para melhorar o status de conservação da espécie	2	peso (não necessita sistematizar)	DESP/MMA
Medida: Anexo CMS	<p>Anexo CMS foi considerado como medida de conservação segundo o seguinte critério:</p> <ul style="list-style-type: none"> • Classe 2: espécies listadas no Anexo II da CMS • Classe 3: espécies listadas no Anexo I da CMS 	0; 2; 3	necessita sistematizar	DESP/MMA; SiBBR/MCTI
Peso da Medida: Anexo CMS	Peso proporcional ao potencial de contribuição da medida para melhorar o status de conservação da espécie	1	peso (não necessita sistematizar)	DESP/MMA

Termo	Descrição MMA	Domínio	Disponibilidade	Responsável
Medida: Anexos CITES	Anexos da CITES foram considerados como medidas de conservação segundo o seguinte critério: • Classe 3: espécies listadas nos Anexos da CITES	0; 3	necessita sistematizar	SiBBr/MCTI
Peso da Medida: Anexos CITES	Peso proporcional ao potencial de contribuição da medida para melhorar o status de conservação da espécie	1	peso (não necessita sistematizar)	DESP/MMA
SOMA CLASSES	Soma dos valores (após multiplicação na coluna vizinha) relativos às medidas de conservação listadas com exceção da Portaria 445 e Anexo CMS. A portaria 445		peso (não necessita sistematizar)	DESP/MMA

2.4. Harmonização da Lista de Termos com Classes e Termos de Padrões Existentes

Nesta etapa do trabalho, foi realizada a harmonização da **Lista de Termos que Necessitam de Sistematização** e da **Lista de Termos Preliminar** com classes e termos de padrões de dados usuais para a biodiversidade.

No escopo deste trabalho foram considerados dois padrões de dados, **Darwin Core** (WIECZOREK *et al*, 2012) e **Dublin Core** (WEIBEL; KOCH, 2000), e os padrões de metadados **Ecological Metadata Language (EML)** (VITOUSEK *et al.*, 2005) e padrão **INDE** para dados e metadados geoespaciais (<https://www.inde.gov.br/CatalogoMetadados>). Para uma descrição detalhada destes e de outros padrões de dados e metadados, recomenda-se o livro desenvolvido no âmbito do Ministério do Meio Ambiente, **Diretrizes para a Integração de Dados de Biodiversidade** (SILVA *et al.*, 2015).

A Oficina de Elaboração da Lista de Termos Acordados foi realizada na data de 16/08/2022 e contou com os mesmos participantes da oficina anterior. O principal objetivo desta oficina foi validar a harmonização das listas de termos relativos à EN com classes e padrões de dados usuais, realizada pela equipe de consultores deste trabalho. Estas listas representam os principais resultados do produto anterior deste trabalho: **Lista de Termos Preliminares**, contendo os termos efetivamente presentes nos padrões usuais e a **Lista de Termos que Necessitam Sistematização**, contendo termos ausentes nos padrões de dados usuais.

Portanto, a lista com os termos harmonizados com padrões usuais foi trabalhada de maneira participativa ao longo da oficina. Termos cujos conteúdos não estavam sendo ofertados de maneira direta nos recursos das instituições vinculadas tiveram destaque nestas discussões, como por exemplo

“unidadesFederativas” e “medida: área distribuição UC+TI+AM”. Nestes exemplos, e nos demais que surgiram, discutiu-se a melhor maneira de se organizar as ofertas de acordo com a disponibilidade de informações existentes, criando listas de demandas e responsabilidades para cada instituição envolvida.

3. PROPOSTA DE PADRÃO DE DADOS PARA ESPÉCIES AMEAÇADAS DE EXTINÇÃO

A importância do uso de padrões e protocolos que sejam adequados ao contexto e sigam semântica correspondente se baseia na redução de risco de erros, na interoperabilidade, escalabilidade de sistemas e consumo de dados por terceiros. No contexto da biodiversidade, a principal referência para definição de padrões de dados é a organização *Biodiversity Information Standards* ([TDWG](#)) e os protocolos e diretrizes de publicação são tratados pelos principais grupos de trabalho no tema.

Apresenta-se a Proposta de Padrão de Dados do MMA para Espécies Ameaçadas de Extinção (**Anexo 1**). Esta proposta de padrão reúne 41 termos e visa servir de modelo de entrega de dados ao MMA pelas instituições vinculadas, responsáveis pela elaboração das listas nacionais de espécies ameaçadas de extinção da fauna e da flora, SALVE/ICMBIO e CNCFlora/JBRJ, respectivamente. Porém, encoraja-se seu uso também por instituições estaduais que publicam ou desejam publicar o mesmo tipo de listas.

Os esforços para elaboração desta proposta envolveram a harmonização das listas de termos da EN com vocabulários e padrões de dados usuais, onde termos com semânticas similares foram agrupados em uma estrutura de classes e termos com metadados associados. Esta harmonização serviu de referência de estrutura de uma proposta de padrão de dados do MMA, ou seja, seguiu-se a estrutura básica dos padrões usuais, mas foi construído um vocabulário específico para uso do MMA, com semântica definida de acordo com as demandas ministeriais.

Apesar da definição de um vocabulário específico para as demandas do MMA, optou-se por manter uma relação direta com relação ao **identificador** (Anexo 1) para termos cuja semântica é idêntica à adotada pelos padrões de dados e metadados de biodiversidade considerados (Darwin Core e Dublin Core). Neste mesmo contexto, optou-se por não considerar uma relação de identificador com vocabulários de uso restrito, como as listas de termos da *International Union for Conservation of Nature* (IUCN) ou os vocabulários utilizados pelas instituições

vinculadas ao MMA, uma vez que estes não possuem uma estrutura semântica publicamente definida.

Os termos desta proposta foram organizados no seguinte conjunto de classes:

- **Táxon:** reúne os dados relacionados à taxonomia, enquanto a classe Nome Vernacular reúne dados relacionados aos nomes populares das espécies;
- **Distribuição:** reúne os termos relativos à distribuição geográfica da espécie em questão;
- **Recursos Associados:** reúne os termos relativos a recursos associados ao táxon, como imagens e documentos
- **Avaliação:** reúne os termos relativos à Análise de Efetividade da Estratégia Nacional de Conservação de Espécies Ameaçadas de Extinção, desenvolvida pelo DESP/MMA.

A Proposta de Padrão de Dados para Espécies Ameaçadas de Extinção foi sistematizada na ferramenta de gestão de planilhas, seguindo uma estrutura relacional [Grist](#). O **Anexo 1** contém os termos relacionados à Proposta exportados na data 23/11/2022.

O conteúdo presente nas tabelas foi organizado nos seguintes campos:

- **termo:** nomes dos termos acordados para este trabalho, seguindo os nomes dos termos equivalentes no padrão de dados de maior relevância;
- **classe:** A classe em que um termo se encontra é relacionada ao tipo de dado principal (para o padrão Darwin Core, Occurrence ou Táxon), sempre apontando para um registro no arquivo principal;
- **identificador:** identificador único do termo;
- **descrição:** Descrição do termo dentro do contexto de uso do MMA;
- **domínio:** para termos controlados foi feita a listagem do conteúdo indicado;
- **exemplo:** exemplo de conteúdo para o termo;
- **observação** - observações sobre o termo ou recomendações de domínio.

3.1. Acesso e Uso da Proposta de Padrão de Dados MMA

Visando à adequação aos princípios FAIR de publicação de dados e metadados, a Proposta de Padrão de Dados do MMA foi incluída no diretório GitHub deste trabalho (<https://github.com/jlanna/MMA-Ameacadas>). De uso amplo e comum por desenvolvedores de sistemas e ferramentas web, principalmente de código aberto, o GitHub se apresenta como uma ferramenta ideal para a documentação deste tipo de trabalho. Ressalta-se que este diretório estará acessível de forma permanente e poderá ser complementado e ter a Proposta de Padrão de Dados atualizada à medida que estudos complementares sejam desenvolvidos, incluindo, por exemplo, outros estudos de caso além da EN.

O uso desta proposta de padrão pode ser feito da mesma maneira como ocorre com outros padrões usuais de dados de biodiversidade como o DwC-A e o Dublin Core. Desta forma, recomenda-se que a publicação dos dados referentes à **EN**, seja realizada com os campos mapeados de acordo com esta proposta de padrão, seguindo recomendações e domínios específicos, conforme indicado nos metadados relativos a cada termo (Figura 3).

bioma	
identificador	https://github.com/jlanna/MMA-Ameacadas/blob/main/padrao/proposta-padrao-mma.md#bioma
Definição	Bioma brasileiro em que a espécie ocorre.
Comentários	Não existe um termo de domínio específico nos padrões internacionais usuais.
Domínio	Amazônia , Mata Atlântica , Cerrado , Caatinga , Pampa , Pantanal , Marinho
Exemplos	Pantanal

Figura 3. Metadados relacionados ao termo “bioma”, presente na Proposta de Padrão de Dados do MMA, disponível em <https://github.com/jlanna/MMA-Ameacadas/blob/main/padrao/proposta-padrao-mma.md>, na data de 12/12/2022.

4. MECANISMOS DE PUBLICAÇÃO DE DADOS DE BIODIVERSIDADE

Definimos como mecanismos de publicação as formas como se é possível publicar dados de forma aberta e acessível na web, buscando seguir os princípios FAIR de publicação de dados. No contexto deste trabalho, a publicação de dados de biodiversidade envolve diferentes formatos com distintos níveis de complexidade com relação à extração e uso, de maneira que estes possam ser acessíveis a analistas de dados; de forma automatizada para máquinas; e para pessoas sem grandes domínios de ciência dos dados. Desta maneira, trataremos de formas de publicação de dados para humanos e para máquinas, em diferentes contextos.

Recomenda-se que a publicação e compartilhamento de dados e metadados de biodiversidade seja realizada por meio de ferramentas e/ou repositórios de dados de uso comum, de preferência seguindo uma arquitetura de microserviços (FOWLER, *et al.* 2014) e hospedados em infraestruturas computacionais estáveis. Visando uma correta indexação destes recursos, é imprescindível a oferta de identificadores únicos permanentes, por meio do qual cada recurso possa ser citado na literatura e indexado nas ferramentas web de busca.

As discussões sobre mecanismos de publicação de dados de biodiversidade na infraestrutura do MMA foram realizadas ao longo da **Oficina de Definição dos Mecanismos de Publicação de Dados**, executada na data de 02/09/2022, contando com a presença dos mesmos convidados das demais oficinas. Ao longo desta oficina foram apresentados e discutidos os diversos mecanismos atualmente disponíveis para publicação de dados de biodiversidade.

Ao longo da oficina direcionou-se as discussões para três grupos principais de repositórios; os **repositórios genéricos**, os **serviços para máquinas** e os **repositórios temáticos**. Considerou-se repositórios genéricos aqueles não restritos a nenhuma área de conhecimento específica, podendo agregar dados de qualquer tipo. Serviços para máquinas são aplicações web que permitem a troca de informações entre componentes de softwares seguindo um conjunto de definições e protocolos. Como repositórios temáticos, foram consideradas ferramentas de

publicação de dados voltadas para uma área do conhecimento específica, se utilizando de formatos e padrões de dados específicos.

Logo abaixo são detalhadas as ferramentas que foram escolhidas como ideais para publicação de dados no contexto deste trabalho, seguindo o estudo de caso da EN. Um levantamento mais amplo de ferramentas pode ser consultado em outros trabalhos realizados no âmbito do MMA ao longo dos últimos anos, como Silva *et al.* (2015) e Lanna (2021).

4.1 Serviços para Máquinas ou Serviços WEB

Os serviços WEB, também chamados pela nomenclatura em inglês *Web services* ou também pela nomenclatura genérica API (*application programming interface*), servem para expor os dados de forma automatizada para serem consumidos por outros sistemas. Logo, um sistema que possui um serviço WEB está ofertando seus dados de forma simplificada e facilmente integrável por outros sistemas e muitas vezes com acesso imediato às suas bases de dados. Logo, as informações ofertadas via serviço WEB podem ser a versão atual do banco de dados, resolvendo assim a questão de dados desatualizados e de versionamento (Figura 4).

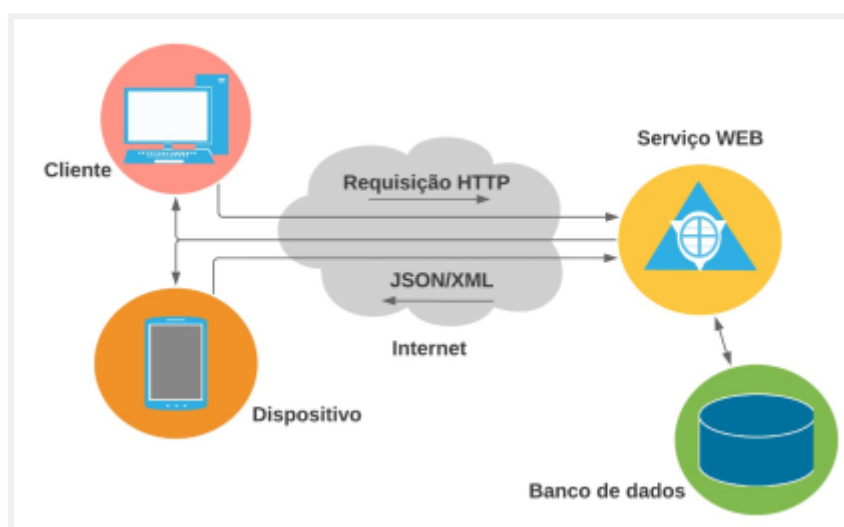


Figura 4. Detalhamento de um serviço web e seu fluxo de informações na internet.

Para melhor indexação, é recomendado que os sistemas de informação, via serviços WEB, ofereçam dados seguindo os padrões usuais para os tipos de dados em questão como o DWc-A ou o Padrão MMA, por exemplo. Também se recomenda que haja uma documentação na Web, servindo como um manual de uso para os usuários.

Os formatos de dados ofertados por estes serviços geralmente são o XML ou JSON, que possuem estruturas facilmente compreensíveis “por máquinas”, possibilitando a plena integração de sistemas.

Citando como exemplo o serviço WEB da Flora e Funga do Brasil, disponível na página de serviços do JBRJ (<http://servicos.jbrj.gov.br/>). Se fizermos a consulta para listar as famílias botânicas que ocorrem no Brasil, podemos utilizar a URL que segue para ter a última versão das informações presentes do banco de dados da Flora e Funga do Brasil em formato integrável por sistemas (<http://servicos.jbrj.gov.br/flora/families>). O resultado resumido desta consulta segue abaixo:

```
{
  "success":true,
  "result":[
    "Acanthaceae",
    "Achariaceae",
    "Achatocarpaceae"
  [...]
    "Zingiberaceae",
    "Zygophyllaceae",
    null
  ]
}
```

O serviço WEB da Flora e Funga do Brasil citado acima, assim como todos os serviços oferecidos pelo JBRJ, segue o pacote Swagger (<https://swagger.io/>), um conjunto de programas de código aberto para desenvolvimento de APIs. O Swagger foi escolhido por ser uma ferramenta moderna e genérica para exportação de dados, de fácil customização e manutenção.

4.2 Repositórios Genéricos

Como repositórios genéricos foram escolhidas as ferramentas já em uso por instituições vinculadas ao Governo Federal, incluindo as instituições vinculadas ao MMA responsáveis pela publicação das listas de espécies brasileiras ameaçadas de extinção. Estas ferramentas possuem como semelhança o fato de seguirem os princípios de ferramentas de código livre e baseadas na web, com possibilidade de associação de metadados, autoria e licença de uso para cada recurso digital (DALCIN *et al.* 2019).

4.2.1. Ckan

O Ckan (<https://ckan.org/>), acrônimo de *Comprehensive Knowledge Archive Network*, foi criado para o armazenamento, gestão e publicação de dados abertos. É utilizado em vários sites do governo federal, inclusive no Portal Brasileiro de Dados Abertos (<https://dados.gov.br/home>).

É indicado para publicação de dados estruturados (planilhas, tabelas, etc.), e utiliza o padrão Dublin Core para catalogação dos metadados de cada recurso. No contexto deste trabalho ele é indicado para publicação de dados para serem utilizados por “humanos”, ou seja, para serem baixados em formatos acessíveis como XLSX, CSV, entre outros, por pessoas com pouco domínio no uso e tratamento de dados. Também pode ser utilizado para publicação de tabelas e planilhas não padronizadas e/ou que sirvam de material suplementar de documentos/artigos.

O CKan aceita diversos formatos de dados, seguindo qualquer tipo de vocabulário ou padrão de dados. Porém, no contexto deste trabalho, é recomendado que os recursos de informação contendo os dados da EN sejam publicados seguindo a Proposta de Padrão de Dados do MMA em formatos de arquivos acessíveis por diversas ferramentas, preferencialmente em mais de uma formato (e.g. .CSV, .XLS, .TSV).

4.2.2. Geonode

O GeoNode (<https://geonode.org/>) é uma plataforma livre para a catalogação e publicação de dados espaciais, em formato vetorial ou matricial. Os dados e metadados catalogados são oferecidos também como web services para acesso em outras aplicações. No contexto deste trabalho o GeoNode é indicado para publicação e compartilhamento de dados espaciais, como shapefiles e outros formatos, preferencialmente seguindo o padrão INDE de metadados.

O JBRJ utiliza o Geonode como uma das ferramentas que compõem sua arquitetura de repositórios digitais, sendo utilizado como sistema de informação espacial e também como repositórios de dados espaciais (DALCIN et al., 2019). A Empresa Brasileira de Pesquisa Agropecuária (Embrapa) utiliza o GeoNode como ferramenta em sua plataforma de dados espaciais [Geoinfo](#), criada a partir da adesão da Embrapa à Infraestrutura Nacional de Dados Espaciais (Inde) (Figura 5).

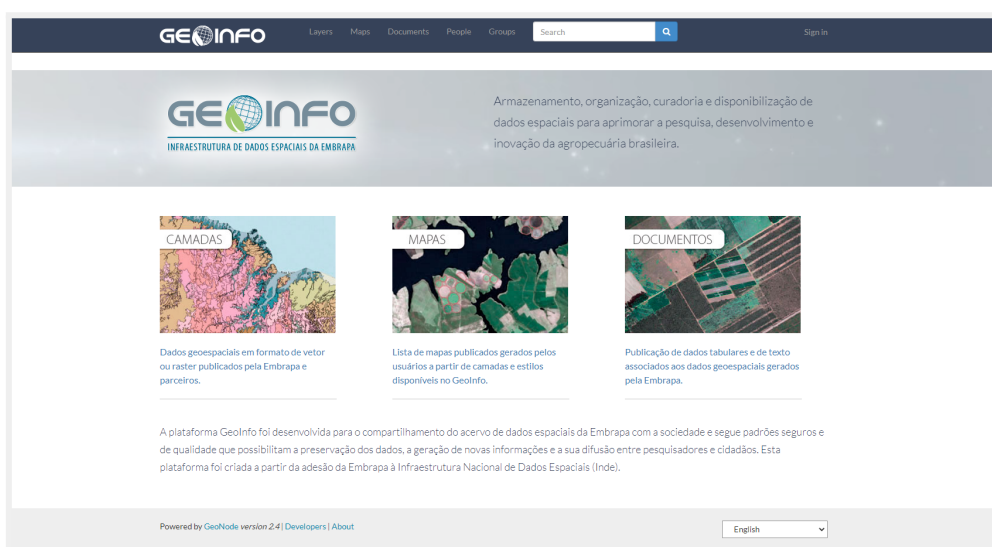


Figura 5. Geoinfo, plataforma de dados espaciais da Embrapa, que tem a ferramenta GeoNode como componente.

Apesar do estudo de caso adotado neste trabalho não incluir dados espaciais, reconhece-se a relevância de se tratar de mecanismos de publicação deste tipo de dados para a tomada de decisão baseada em dados. Portanto, destaca-se o GeoNode como um modelo de ferramenta importante para gestão e publicação de dados espaciais.

4.3 Repositórios Temáticos

Foram considerados repositórios temáticos aqueles voltados para publicação de dados de uma área do conhecimento em especial, baseados em formatos e vocabulários, ou padrões de dados específicos. Como ferramenta desta categoria foi escolhido o *Integrated Publishing Toolkit* (IPT, <https://www.gbif.org/pt/ipt>) como modelo para publicação de dados seguindo o padrão Darwin Core.

No contexto deste trabalho é importante ressaltar que por este tipo de ferramenta exigir o uso de padrões de dados específicos, muitas vezes não é possível mapear a totalidade dos termos presentes em conjuntos de dados mais abrangentes, como é o caso da EN. Porém, mesmo com estas limitações, recomenda-se uma análise mais profunda da ferramenta para produtores de dados devido à sua relevância e para consideração de seu eventual uso.

4.3.1 IPT

O IPT é uma ferramenta aberta e gratuita para a publicação e compartilhamento de conjuntos de dados de biodiversidade (Robertson et al. 2014; SILVA *et al.*, 2015). Pode ser instalada e mantida pela instituição publicadora de dados em servidor próprio ou em uma instituição agregadora de dados, como o SiBBr, e funciona como uma ferramenta baseada na Web. Desenvolvida pelo GBIF, com o apoio de diversas instituições de biodiversidade, é a ferramenta recomendada para a publicação de dados na rede do GBIF e no nó brasileiro do GBIF, o SiBBr.

Os dados podem ser inseridos em diversos formatos na ferramenta (consulta via SQL em banco de dados relacional, e via arquivos .xls ou .csv), contanto que os campos de dados sejam mapeados no formato DwC. Juntamente com um arquivo de metadados em formato EML, o recurso pode ser empacotado em um arquivo Darwin Core Archive (DwC-A), pronto para ser publicado e automaticamente exportado para portais agregadores de dados de biodiversidade, como o GBIF e SiBBr (Figura 6).

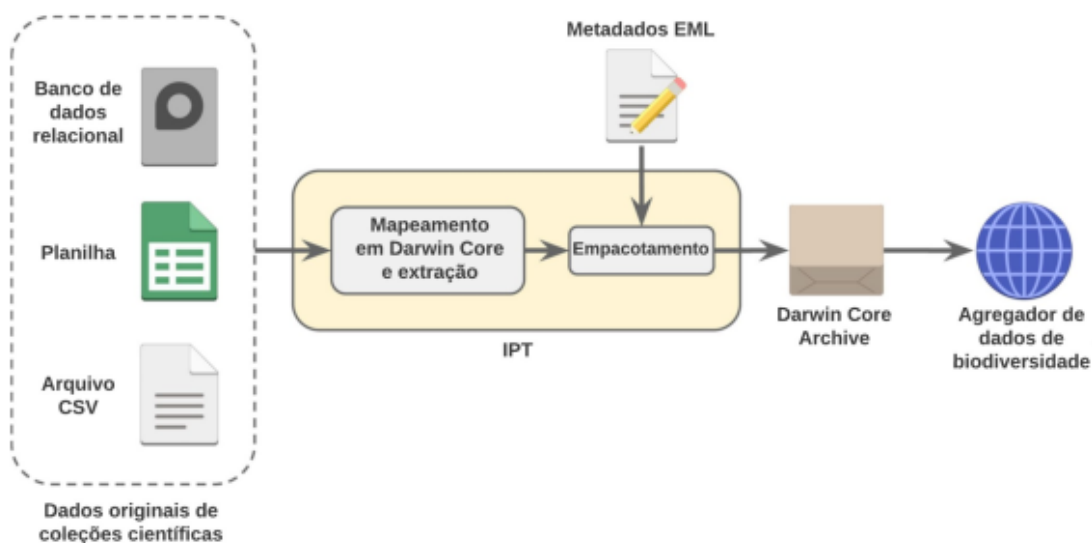


Figura 6. Esquema de entrada, empacotamento e publicação de dados através da ferramenta IPT. Retirado de Gadelha et al. (2020).

Como nó brasileiro do GBIF, o SiBBr possui uma instalação do IPT (<https://ipt.sibbr.gov.br/sibbr/>) disponível para a publicação de dados de parceiros e também funciona como um agregador de [dados por meio dos IPTs de instituições parceiras](#). Desta maneira propõe-se o uso da infraestrutura do SiBBr para publicação direta de dados, por meio de sua instalação de IPT, ou que esta publicação seja feita a partir de provedor próprio e os dados sejam agregados ao SiBBr como forma de divulgação e uso público desses recursos de informação.

5. CONCLUSÕES E RECOMENDAÇÕES

Os dados relativos a espécies ameaçadas abordados neste trabalho não se enquadram completamente nos padrões de dados existentes. Esta é uma condição esperada, no sentido de que padrões de dados são construídos para atender a um conjunto de demandas específicas. O padrão Darwin Core, por exemplo, mesmo sendo desenvolvido para dados de ocorrências e checklists de espécies, não abrange todos os campos presentes em uma ficha de catálogo de uma coleção biológica brasileira comum. Campos como Bioma e Município estão restritos a uma lógica regional de concepção da paisagem e logo não se aplicam a padrões globalmente abrangentes. Neste sentido, é crucial e estratégico para uma gestão eficiente de dados, que sejam desenvolvidas soluções para demandas regionais, neste caso, nacionais, sem deixar de contribuir e se integrar às redes globais de dados de biodiversidade.

Complementarmente, o estudo de caso adotado neste trabalho não representa a totalidade de dados relacionados a espécies ameaçadas. Portanto, é importante a compreensão deste trabalho como um exercício aplicado no sentido de se desenvolver um padrão abrangente para dados sobre biodiversidade no âmbito do MMA. Neste sentido, buscou-se desenvolver uma metodologia de trabalho que possibilite sua replicação futura e a construção de versões estendidas do Padrão de Dados do MMA e Mecanismos de Publicação compatíveis.

Reforça-se a importância da adoção dos produtos deste trabalho nas práticas de gestão de dados no MMA e instituições vinculadas responsáveis pela avaliação de risco de extinção para espécies da fauna e flora. Recomenda-se que esta proposta de padrão de dados, bem como os mecanismos de publicação aqui propostos, sejam incorporados às práticas institucionais das instituições envolvidas neste trabalho de maneira ágil e permanente..

O processo de construção da Análise de Efetividade da Estratégia Nacional de Conservação de Espécies Ameaçadas será otimizado com a publicação eficiente dos dados sobre espécies ameaçadas seguindo os modelos propostos neste trabalho. A publicação destes dados seguindo o padrão de dados do MMA e os

princípios F.A.I.R. (dados “encontráveis, acessíveis, interoperáveis e reproveitáveis”) possibilitará um fluxo automatizado de dados entre máquinas, gerando informações atuais e baseadas nos melhores dados disponíveis. Resumidamente, estes dados poderiam ser organizados, mapeados segundo o padrão proposto e exportados dos bancos de dados dos respectivos projetos via serviços WEB. Complementarmente, a exportação destes dados publicados via serviço podem ser convertidos em formatos amigáveis para o público e publicados em repositórios apropriados, de acesso público e aberto, como recursos de informação temporalmente versionados.

Por fim, recomenda-se esforços específicos no sentido de incluir dados espaciais em trabalhos futuros relacionados a este tema.

6. REFERÊNCIAS

Bisby, F. A. (2000). The Quiet Revolution: Biodiversity Informatics and the Internet. *Science*, 289(5488), 2309–2312.
<https://doi.org/10.1126/science.289.5488.2309>

Dalcin, E., Lanna, J., Queiroz, N., Forzza, R. C. (2019). Arquitetura para publicação de dados sobre biodiversidade em instituições de pesquisa. *Liinc em Revista*, vol. 15, no. 2, p. 52–64.

De Vos, J. M., Joppa, L. N., Gittleman, J. L., Stephens, P. R. & Pimm, S. L. (2015). Estimating the normal background rate of species extinction. *Conserv. Biol.* 29, 452–462.

Fowler, Martin, and J. Lewis. “Microservices a definition of this new architectural term.” (2014) URL: <https://martinfowler.com/articles/microservices.html>. Acesso em 13/12/2022.

Gadelha, L. M. R., de Siracusa, P. C., Dalcin, E. C., de Siqueira, M. F., et al. (2020). A survey of biodiversity informatics: Concepts, practices, and challenges. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 11(1), 1–41. <https://doi.org/10.1002/widm.1394>

Lanna, J. M. (2021) A gestão de dados e informações sobre biodiversidade no Ministério do Meio Ambiente e Instituições Vinculadas (2021). Produto do contrato NO: 002103-2020 assinado pela WWF-Brasil e elaborado no âmbito do Projeto Pró-Espécies.

Murray, K. A., Verde Arregoitia, L. D., Davidson, A., Di Marco, M. & Di Fonzo, M. M. I. (2014). Threat to the point: improving the value of comparative extinction risk analysis for conservation action. *Glob. Chang. Biol.* 20, 483–494.

Robertson, T., Doring, M., Guralnick, R., Desmet, P., et al. (2014). The GBIF integrated publishing toolkit: Facilitating the efficient publishing of biodiversity

data on the internet. PLoS ONE, vol. 9, no. 8.
<https://doi.org/10.1371/journal.pone.0102623>.

Silva, D. L., Corrêa, P. L. P., Juárez, K. M., & Fonseca, R. L. (2015). Diretrizes para a Integração de Dados de Biodiversidade. MMA

Soberón, J., & Peterson, A. T. (2004). Biodiversity informatics: Managing and applying primary biodiversity data. Philosophical Transactions of the Royal Society B: Biological Sciences, 359(1444), 689–698.
<https://doi.org/10.1098/rstb.2003.1439>

Vitousek, P. M., Aber, J. D., Howarth, R. W., Likens, G. E., A, M. P., Schindler, D. W., Schlesinger, W. H., & Tilman, D. G. (2005). Maximizing the Value of Ecological Data with Structured Metadata: An Introduction to Ecological Metadata Language (EML) and Principles for Metadata Creation. Ecological Society of America, 86(3), 158–168.

Weibel, S. L., & Koch, T. (2000). The Dublin Core Metadata Initiative. D-Lib Magazine, 6(12). <https://doi.org/10.1045/december2000-weibel>

Wieczorek, J., Bloom, D., Guralnick, R., Blum, S., Döring, M., Giovanni, R., Robertson, T., & Vieglais, D. (2012). Darwin core: An evolving community developed biodiversity data standard. PLoS ONE, 7(1).
<https://doi.org/10.1371/journal.pone.0029715>

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J. W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Mons, B. (2016). Comment: The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3(March).
<https://doi.org/10.1038/sdata.2016.18>

Anexo 1. Proposta de Padrão de Dados para Espécies Ameaçadas de Extinção desenvolvida no contexto do estudo de caso da Análise de Efetividade da Estratégia Nacional de Conservação de Espécies Ameaçadas de Extinção.

termo	classe	identificador	descrição	domínio	exemplo	observação
taxonID	táxon	https://dwc.tdwg.org/list/#dwc_taxonID	Um identificador único permanente para o nome científico associado ao táxon		8fa58e08-08de-4ac1-b69c-1235340b7001, 32567, https://www.gbif.org/species/212	
reino	táxon	https://dwc.tdwg.org/list/#dwc_kingdom	O nome científico completo do Reino em que o táxon é classificado.	Plantae, Animalia, Fungi	Plantae	Nas bases originais do MMA constam "Fauna" e "Flora" no domínio. Contudo, visando a compatibilidade semântica entre sistemas, sugerimos a adoção dos termos "Plantae" e "Animalia" conforme listado.
grupo	táxon		O nome "informal" para o grupo de organismos em que o táxon é classificado. Neste caso, utiliza-se a classificação adotada pelo DESM/MMA de domínio fechado.	Anfíbios, Angiospermas, Aves, Briófitas, Gimnospermas, Invertebrados Aquáticos, Invertebrados Terrestres, Mamíferos, Peixes Continentais, Peixes	Aves	É uma definição de subconjuntos de taxa conforme necessidades específicas do MMA. O termo equivalente adotado "gmp:taxonomicCoverage:commonName" é indicado para nomes comuns aplicáveis, mas pode também ser uma descrição geral para grupos de

termo	classe	identificador	descrição	domínio	exemplo	observação
				Marinhos, Pteridófitas, Répteis		organismos, se apropriados. E.g. pau-brasil ou peixes continentais.
familia	táxon	https://dwc.tdwg.org/list/#dwc_family	O nome científico completo da família em que o táxon é classificado.		Felidae, Monocleacea e	
genero	táxon	https://dwc.tdwg.org/list/#dwc_genus	O nome científico completo do gênero em que o táxon é classificado.	Nomes válidos para o ranque "Gênero" na Flora e Funga do Brasil e Catálogo da Fauna, para a "Família" correspondente	Hura	
epitetoEspecifico	táxon	https://dwc.tdwg.org/list/#dwc_specificEpithet	O nome do epíteto específico da espécie.	Nomes válidos para o ranque "Epíteto Específico" na Flora e Funga do Brasil e Catálogo da Fauna, para o "Gênero" correspondente	crepitans	
epitetoInfraespecifico	táxon	https://dwc.tdwg.org/list/#dwc_infraspecificEpithet	O nome do epíteto infra específico da espécie, excluindo qualquer designação de ranque taxonômico.	Nomes válidos para o ranque "Infra Espécie" na Flora e Funga do Brasil e Catálogo da Fauna,	alba	

termo	classe	identificador	descrição	domínio	exemplo	observação
				para a "Espécie" correspondente		
nomeVernacular	nomeVernacular	https://dwc.tdwg.org/list/#dwc_vernacularName	Nome vernacular/comum/popular popularmente dado ao táxon descrito.		pau-brasil	
idioma	nome_vernacular	https://dwc.tdwg.org/terms/#dwc:language	Idioma relacionado ao campo <code>mma:nomeVernacular</code> .	Recomenda-se como melhores práticas o uso de vocabulários controlados como RFC 5646.	pt (para Português), en (para Inglês), es (para Espanhol)	
localidade	nome_vernacular	https://dwc.tdwg.org/terms/#dwc:locality	Localidade onde é utilizado o <code>mma.nomeVernacular</code> .		Sudeste, Bahia, SP, RJ e SP	
bioma	distribuição	https://github.com/jlanna/MMA-Ameacadas/blob/main/propostaPadrao.md#bioma	Bioma brasileiro em que a espécie ocorre.	Amazônia, Mata Atlântica, Cerrado, Caatinga, Pampa, Pantanal, Marinho	Pantanal	Não existe um termo de domínio específico nos padrões internacionais usuais. O termo equivalente dwc apontado foi adotado conforme demandas internas dos produtores.
endemismo	distribuição	https://github.com/jlanna/MMA-Ameacadas/blob/main/propostaPadrao.md#endemismo	Endemismo para Brasil.			Os termos base listados possuem domínios específicos

termo	classe	identificador	descrição	domínio	exemplo	observação
		A-Ameacadas/blob/main/propostaPadrao.md#endemismo				que devem ser considerados para eventual mapeamento. No caso do termo ``dwc'', recomenda-se o uso de um json detalhando nome de campo e conteúdo.
unidades Federativas	distribuição	https://dwc.tdwg.org/list/#dwc_locationID	Unidades Federativas onde o táxon ocorre.	AC, AL, AP, AM, BA, CE, ES, GO, MA, MT, MS, MG, PA, PB, PR, PE, PI, RJ, RN, RS, RO, RR, SC, SP, SE, TO, DF	MG, RJ	
tipo	registro associado	https://dwc.tdwg.org/terms/#dwc:type	A natureza ou gênero do recurso	foto, vídeo e demais formatos de arquivos multimídia	foto	
licença	registro associado	https://dwc.tdwg.org/list/#dwc:license	Um documento legal dando permissão oficial de uso do recurso.			
codigoInstituicaoProprietaria	registro associado	http://rs.tdwg.org/dwc/terms/ownerInstitutionCode	O nome (ou acrônimo) em uso pela instituição possuidora de direitos sobre o objeto ou informação referida no registro.			

termo	classe	identificador	descrição	domínio	exemplo	observação
creditos	registro associado	https://www.dublin-core.org/specifications/dublin-core/dcmi-terms/#rightsHolder	Uma pessoa ou organização possuidora ou detentora de direitos sobre o recurso.			
categoria	avaliação	https://github.com/jlanna/MMA-Ameacadas/blob/main/propostaPadrao.md#categoria	Código da categoria de ameaça definida.	CR, DD, EN, EW, EX, LC, NE, NR, NT, VU	CR	
critério	avaliação	https://github.com/jlanna/MMA-Ameacadas/blob/main/propostaPadrao.md#critério	Uma string tokenizada dos critérios de avaliação considerados na avaliação da espécie. Deve conter tokens de cada critério individual, separados por vírgula. Por exemplo: A seguinte string de critérios "B1ab(ii,iii,iv,v)+2ab(ii,iii,iv,v)" deve ser representada da seguinte forma: "B1a,B1bii,B1biii,B1biv,B1bv,B2a,B2bii,B2biii,B2biv,B2bv". A vírgula também pode ser substituída por barra vertical " ".		B2ab(iii), A2c,B1ab(iii)+2ab(iii)	
ameacas	avaliação	https://github.com/jlanna/MMA-Ameacadas/blob/main/propostaPadrao.md#ameacas	Descrição da ameaça.		Agropecuária ; Extração	Os termos equivalentes listados possuem domínios

termo	classe	identificador	descrição	domínio	exemplo	observação
		A-Ameacadas/b lob/main/prop staPadrao.md#a meacas			Direta: Caça/Pesca; Queimadas	específicos que devem ser considerados para eventual mapeamento.
presença em áreas protegidas	avaliação	https://github.com/jlanna/MM-A-Ameacadas/blob/main/propstaPadrao.md#presenca-em-areas-protegidas	Presença ou ausência em áreas protegidas, como Unidades de Conservação, Terras Indígenas e Áreas Militares Entretanto, devido a ausência de dados para Terras Indígenas e Áreas Militares, o termo poderia se restringir a UC e ser nomeado 'presença em unidades de conservacao'	Sim; Não		Catálogo de UCs pode ser uma fonte de dados importante para flora. Os termos equivalentes dwc:MeasurementOrFact e dwc:measurementValue devem ser usados concomitantemente, o primeiro para descrever o campo e o segundo para o valor.
plano de ação nacional para conservação (pan)	avaliação	https://github.com/jlanna/MM-A-Ameacadas/blob/main/propstaPadrao.md#plano-de-acao-nacional-para-conservacao-(pan)				Os termos equivalentes listados possuem domínios específicos que devem ser considerados para eventual mapeamento.
ordenamento pesqueiro	avaliação	https://github.com/jlanna/MM-A-Ameacadas/b				Os termos equivalentes listados possuem domínios específicos que

termo	classe	identificador	descrição	domínio	exemplo	observação
		lob/main/propostaPadrao.md#ordenamento-pe-squeiro				devem ser considerados para eventual mapeamento.
nível de protecao na estrategia nacional	avaliação	https://github.com/jlanna/MMA-Ameacadas/blob/main/propostaPadrao.md#nivel-de-protecao-na-estrategia-nacional	Os Níveis de Proteção variam de 0 a 5, em que 0 demonstra ausência de medidas de conservação e 5 demonstra que a espécie está suficientemente coberta por medidas de conservação.			
medida: area distribuicao UC+TI+AM	avaliação	https://github.com/jlanna/MMA-Ameacadas/blob/main/propostaPadrao.md#medida-area-distribuicao-UC+TI+AM	Esta medida se refere ao grau de proteção de habitat da espécie em Unidades de Conservação, Terras Indígenas e Áreas Militares. Os cálculos utilizados nesta medida consideram o grau de efetividade de cada área protegida calculado pelo SAMGe/ICMBio e o grau de sobreposição da área de ocorrência da espécie com estas áreas protegidas. Detalhes do cálculo encontram-se no documento intitulado "Descricao_metodologia_analiseefetividade_nov18".	0; 1; 2; 3		

termo	classe	identificador	descrição	domínio	exemplo	observação
peso: area distribuicao UC+TI+AM	avaliação	https://github.com/jlanna/MMA-Ameacadas/blob/main/propostaPadrao.md#peso_area-distribuicao-UC+TI+AM	Peso proporcional ao potencial de contribuição da medida para melhorar o status de conservação da espécie, que para esta medida de conservação representa o valor 2,5	2,5		
medida: caverna	avaliação	https://github.com/jlanna/MMA-Ameacadas/blob/main/propostaPadrao.md#medida_caverna	<p>As análises espaciais de ocorrência das espécies em cavernas foram descritas no item anterior. A lista das espécies ameaçadas que ocorrem em cavernas foi fornecida pelo CECAV/ICMBio.</p> <p>A sugestão da reunião de especialistas foi considerar com um peso maior as cavernas classificadas como de máxima relevância. Entretanto, a ocorrência de espécies ameaçadas é um dos critérios que classificam a caverna como de máxima relevância. Portanto, todas as cavernas deste exercício são necessariamente de máxima relevância.</p> <p>Devido à dificuldade de determinar a extensão das cavernas, não foi possível calcular a área de</p>	0; 2		

termo	classe	identificador	descrição	domínio	exemplo	observação
			distribuição das espécies que está coberta por cavernas, e sim presença e ausência. Dessa forma, foi atribuída Classe 2 neste critério para todas as espécies analisadas que apresentaram sobreposição com cavernas.			
peso: caverna	avaliação	https://github.com/jlanna/MMA-Ameacadas/blob/main/propostaPadrao.md#peso_caverna	Peso proporcional ao potencial de contribuição da medida para melhorar o status de conservação da espécie.	2		
medida: PAN	avaliação	https://github.com/jlanna/MMA-Ameacadas/blob/main/propostaPadrao.md#medida_PAN	<p>A efetividade dos PAN foi avaliada de acordo com seu grau de implementação:</p> <ul style="list-style-type: none"> • Classe 1: espécies que estão em PAN não avaliado; espécies em PAN encerrados, aguardando novo ciclo; espécies em PAN com menos de 40% de implementação; • Classe 2: espécies que estão em PAN com grau de implementação igual ou maior que 40%; • Classe 3: espécies em PAN encerrado que atingiu seu objetivo e não necessita de novo ciclo. 	0; 1; 2; 3		

termo	classe	identificador	descrição	domínio	exemplo	observação
peso: PAN	avaliação	https://github.com/jlanna/MMA-Ameacadas/blob/main/propostaPadrao.md#peso_PAN	Peso proporcional ao potencial de contribuição da medida para melhorar o status de conservação da espécie.	2		
medida: ordenamento pesqueiro	avaliação	https://github.com/jlanna/MMA-Ameacadas/blob/main/propostaPadrao.md#medida_ordenamento-pesqueiro	<p>A equipe do DESP/MMA realizou um levantamento das normas de ordenamento pesqueiro vigentes, considerando também moratórias e planos de recuperação. O critério considerado foi o seguinte:</p> <ul style="list-style-type: none"> • Classe 1: espécie é indiretamente beneficiada pela medida; • Classe 2: ordenamento direcionado às práticas que afetam diretamente as espécies; • Classe 3: moratória ou citada nominalmente no ordenamento (exclusivo exemplo portaria da moratória da Piracatinga). 	0; 1; 2; 3		
peso: ordenamento pesqueiro	avaliação	https://github.com/jlanna/MMA-Ameacadas/blob/main/propostaPadrao.md#peso_ordename	Peso proporcional ao potencial de contribuição da medida para melhorar o status de conservação da espécie.	1		

termo	classe	identificador	descrição	domínio	exemplo	observação
		nto-pesqueiro				
medida: portaria 443	avaliação	https://github.com/jlanna/MMA-Ameacadas/blob/main/propostaPadrao.md#medida_portaria-443	<p>A Portaria MMA nº 443/2014, que estabelece a lista das espécies ameaçadas da flora, foi considerada como uma medida de conservação pelo fato de proibir coleta, corte, transporte, armazenamento, manejo, beneficiamento e comercialização das espécies listadas.</p> <p>Todavia, o Art. 3º permite uma exceção para as espécies da categoria Vulnerável (VU), desde que regulamentada. As espécies madeireiras desta categoria que estão no bioma Amazônia foram regulamentadas, o que permite seu corte. Portanto, essas espécies receberam uma pontuação menor nesse critério:</p> <ul style="list-style-type: none"> • Classe 1: espécies madeireiras, da categoria VU, que ocorrem no bioma amazônico; • Classe 2: demais espécies da flora, listadas na Portaria MMA nº 443/2014. 			
peso:	avaliação	https://github.com/jlanna/MMA-Ameacadas/blob/main/propostaPadrao.md#medida_portaria-443	Peso proporcional ao potencial de	2		

termo	classe	identificador	descrição	domínio	exemplo	observação
portaria 443		om/jlanna/MM A-Ameacadas/b lob/main/propo staPadrao.md#p eso_portaria-44 3	contribuição da medida para melhorar o status de conservação da espécie.			
medida: portaria 445	avaliação	https://github.com/jlanna/MM-A-Ameacadas/blob/main/propostaPadrao.md#medida_portaria-445	<p>A inclusão da Portaria MMA nº 445/2014 segue o mesmo princípio que o da Portaria MMA nº 443/2014, considerando que a lista das espécies de peixes e invertebrados aquáticos ameaçados também oferece proteção diferenciada a estas espécies em relação às espécies não listadas considerando as restrições de uso previstas. O critério considerado foi o seguinte:</p> <ul style="list-style-type: none"> • Classe 1: espécies com interesse socioeconômico que se beneficiariam com a Portaria MMA 445/2014; • Classe 2: CR e EN de interesse econômico, com captura proibida. 			
peso: portaria 445	avaliação	https://github.com/jlanna/MM-A-Ameacadas/blob/main/propostaPadrao.md#medida_portaria-445	Peso proporcional ao potencial de contribuição da medida para melhorar o status de conservação	2		

termo	classe	identificador	descrição	domínio	exemplo	observação
		lob/main/propostaPadrao.md#peso_portaria-445	da espécie.			
medida: anexo CMS	avaliação	https://github.com/jlanna/MMA-Ameacadas/blob/main/propostaPadrao.md#medida_anexo-CMS	Anexo CMS foi considerado como medida de conservação segundo o seguinte critério: <ul style="list-style-type: none"> • Classe 2: espécies listadas no Anexo II da CMS; • Classe 3: espécies listadas no Anexo I da CMS. 	0; 2; 3		
peso: anexo CMS	avaliação	https://github.com/jlanna/MMA-Ameacadas/blob/main/propostaPadrao.md#peso_anexo-CMS	Peso proporcional ao potencial de contribuição da medida para melhorar o status de conservação da espécie.	1		
medida: anexos CITES	avaliação	https://github.com/jlanna/MMA-Ameacadas/blob/main/propostaPadrao.md#medida_anexos-CITES	Anexos da CITES foram considerados como medidas de conservação segundo o seguinte critério: <ul style="list-style-type: none"> • Classe 3: espécies listadas nos Anexos da CITES. 	0; 3		
peso:	avaliação	https://github.com/jlanna/MMA-Ameacadas/blob/main/propostaPadrao.md#peso_anexo-CMS	Peso proporcional ao potencial de	1		

termo	classe	identificador	descrição	domínio	exemplo	observação
anexos CITES		om/jlanna/MM A-Ameacadas/b lob/main/prop staPadrao.md#p eso_anexos-CIT ES	contribuição da medida para melhorar o status de conservação da espécie.			
soma classes	avaliação	https://github.com/jlanna/MM-A-Ameacadas/blob/main/propostaPadrao.md#soma-classes	Soma dos valores (após multiplicação na coluna vizinha) relativos às medidas de conservação listadas com exceção da Portaria 445 e Anexo CMS.			