

GESTÃO DE DADOS DA BIODIVERSIDADE: aplicação do padrão de metadados Darwin Core

Filipi Miranda Soares, Raíssa Yuri
Hamanaka, Benildes C. Moreira dos
Santos Maculan

INTRODUÇÃO

No atual contexto do fenômeno big data, em que são gerados grandes volumes de dados não-estruturados em meio digital, torna-se de extrema importância a questão da organização do dado para sua posterior recuperação.

A efetiva gestão de dados de pesquisa e sua posterior reutilização por demais pesquisadores, depende de um conjunto de práticas durante a coleta, o processamento e a análise dos mesmos

Essas práticas compõem o ciclo de vida dos dados (CVD), que é formado pelos seguintes elementos: planejar, coletar, assegurar, descrever, preservar, descobrir, integrar e analisar

O foco deste estudo foi o elemento “descrição” do CVD, definido como “dados que são descritos com precisão e minuciosamente usando os padrões de metadados apropriados” (STRASSER et al., 2012, p. 3, tradução nossa).

Para explicitar o elemento descrição (e seus fatores) do CVD, analisaram-se os metadados de um registro de ocorrência do repositório de dados do Portal da Biodiversidade em comparação com o padrão de metadados Darwin Core (DwC).

OBJETIVO

Analisar a aplicação do padrão de metadados Darwin Core (DwC) em um registro de ocorrência do Portal da Biodiversidade

PROCEDIMENTOS METODOLÓGICOS

O universo de pesquisa é todos os registros no Portal da Biodiversidade e como amostra foi selecionado o registro de ocorrência da espécie *Rhinella granulosa* (Spix, 1824), leigamente conhecida como sapo da caatinga.

O procedimento adotado foi correlacionar os metadados do DwC com o significado atribuído a cada campo do registro e avaliar o seu uso. Assim, as categorias de análise selecionadas para análise foram os metadados adotados para descrição do registro de ocorrência, que foram correlacionadas com o conjunto de metadados do DwC.

RESULTADOS

A análise determinou que o registro selecionado possui 34 metadados, apresentados em formato de folha de dados (CSV), baixado do Portal da Biodiversidade, que serão apresentados na ordem que aparecem no registro.

Conforme o objetivo do estudo, foram analisados 34 campos de um registro de ocorrência do Portal da Biodiversidade em correlação com os campos do padrão de metadados DwC. Na maioria das análises, em cerca de 65% dos casos, houve correspondência do campo do registro com o campo do padrão de metadados

CONSIDERAÇÕES FINAIS

Acredita-se que a padronização dos dados coletados e armazenados em repositórios permite a democratização dos dados ao garantir o acesso efetivo aos mesmos.

REFERÊNCIAS

COSTA, Maíra Murrieta. **Diretrizes para uma política de gestão de dados científicos no Brasil**. 2017. 288 f. Tese (Doutorado em Ciência da Informação) – Faculdade de Ciência da Informação, Universidade de Brasília, Brasília, DF, 2017.

DARWIN CORE TASK GROUP. Biodiversity Information Standards. Darwin Core. 2015. Available from: <<http://rs.tdwg.org/dwc/>>. Accessed on: 8 Sept. 2018.

FAIR principles for data stewardship. **Nature Genetics**, London, v. 48, n. 4, p. 343, Apr. 2016.

SANT'ANA, Ricardo César Gonçalves. Ciclo de vida dos dados: uma perspectiva a partir da Ciência da Informação. **Inf. Inf.**, Londrina, v. 21, n. 2, p. 116-142, maio/ago. 2016.

SAYÃO, Luis Fernando; SALES, Luana Faria. Curadoria digital: um novo patamar para preservação de dados digitais de pesquisa. **Inf. & Soc.: Est.**, João Pessoa, v. 22, n. 3, p. 179-191, set./dez. 2012.

STRASSER, Carly et al. Primer on data management: what you always wanted to know: but were afraid to ask. **DataONE**, Oakland, CA, p. 1-11, fev. 2012.