

# AQUISIÇÃO E ORGANIZAÇÃO DE DADOS GEOESPACIAIS PARA USO NO PROJETO AQUA-BNDES: REGISTRO DE METADADOS E ATRIBUTOS

*Fábio Camargo Rodrigues<sup>1,2</sup>, Sandro Eduardo Marschhausen Pereira<sup>2</sup>, Celso Vainer Manzatto<sup>2</sup>, e Luciana Spinelli de Araújo<sup>2</sup>*

<sup>1</sup>Centro Universitário de Jaguariúna, <sup>2</sup>Embrapa Meio Ambiente; fabiocr899@gmail.com; sandro.pereira@embrapa.br; celso.manzatto@embrapa.br; luciana.spinelli@embrapa.br

## RESUMO

Este trabalho tem por objetivo analisar metadados geoespaciais de dados geoespaciais coletados em ambiente web. Além disso, busca descrever seus atributos, exclusivamente, a partir do site de origem do dado. O contexto deste trabalho foi a sistematização de uma base de dados geoespaciais como suporte à criação de uma rede de pesquisa e monitoramento da aquicultura brasileira. Esta rede, visa fortalecer a indústria e promover o desenvolvimento sustentável (econômico e ambiental) desse setor. Assim, foi compilada uma lista com 15 metadados, com base no Perfil de Metadados Geoespaciais Brasileiro (Perfil MGB) e na Norma Brasileira de Referências (NBR 6023). Como resultado, foram analisados 38 dados de 4 instituições. Dos quais, identificou-se a lista inteira de metadados em 68% dos dados, todos os atributos foram identificados em 60% dos dados e foi possível preencher completamente tanto a lista de metadados quanto a de atributos em 47% dos dados.

**Key words** — Rede Aquicultura, Perfil Brasileiro de Metadados Geoespaciais (Perfil MGB), Dados científicos, Geoinformação, Ciência agrícola.

## ABSTRACT

This work aims to analyze geospatial metadata collected through a web environment. Moreover, the possibility of describing its attributes, only from the web site in which it was collected. Its context is a systematization of a geospatial database to offer support to the creation of an Aquaculture Research and Monitoring Net. Which aims to strengthen the Industry and promote its sustainable development (economical and environmental). Thus, a list with 15 metadata was compiled based on the Brazilian Metadata Standard (Perfil MGB) and the Brazilian Reference Standard (NBR 6023). As a result, 38 data from 4 institutions were analyzed. From which, we were able to identify the whole metadata list in 68% of the data. We could describe all of the attributes in 60% of the data. Finally we could identify the whole metadata list and describe all of the attributes in 47% of the data.

**Key words** — Aquaculture Net, Brazilian Metadata Standard (Perfil MGB), Scientific data, Geoinformation, Agricultural Science.

## 1. INTRODUÇÃO

O pescado representa, atualmente, a proteína animal mais produzida no mundo, sendo o Brasil o 13º produtor mundial de pescado cultivado [1]. Estima-se que até 2025 haverá um crescimento de mais de 100% em relação a 2015-2016 [1]. Assim, para manter as taxas de expansão anual com sustentabilidade é fundamental adotar um manejo produtivo eficiente com otimização da gestão ambiental da aquicultura [2]. Além de sensibilizar os agentes da sua cadeia produtiva para a adoção de Boas Práticas de Manejo [3, 4]

Diante dessa demanda, o projeto Ações estruturantes e de inovação para o fortalecimento das cadeias produtivas da aquicultura no Brasil (BRS Aqua) apresenta o Componente Manejo e Gestão Ambiental da Aquicultura (PC MAN). Este pretende estabelecer conexões entre todos os componentes da cadeia produtiva (empresários, fiscalização, pesquisa e administração pública). Entre os resultados esperados está a sistematização de uma base de dados geoespaciais para dar suporte ao estabelecimento de uma Rede de Pesquisa e Monitoramento da Aquicultura brasileira. [5,6]

Os dados geoespaciais serão utilizados para a caracterização de reservatórios, integrações temáticas geoespaciais, elaboração de um WebGis e demais produtos direcionados para o desenvolvimento e fortalecimento da Rede. A coleta, organização, preservação e documentação desses dados envolve a tarefa de registro dos seus metadados [7], para os quais existe uma lista mínima e obrigatória a ser preenchida [8].

Os metadados geográficos tem as funções de identificar o produtor do dado, bem como de quem é a responsabilidade técnica. Permite, também garantir o compartilhamento e a transferência do dado tendo tanto a informação sobre a qualidade do dado, quanto dos requisitos para sua distribuição [8].

Neste trabalho são abordadas as estratégias para análise dos metadados de dados geoespaciais concluídos disponíveis na internet e para a descrição dos seus atributos com base no ambiente virtual (internet) da instituição proprietária dos dados. O objetivo é realizar uma análise desses metadados e verificar se é possível obter toda a informação obrigatória, além de descrever os atributos de cada dado

## 2. MATERIAIS E MÉTODOS

Os dados foram coletados para levantar o máximo de características disponíveis sobre os reservatórios com

atividades aquícolas. Em função disso, foram buscadas informações sobre suas características físicas e sobre a qualidade da água, além de dados socioeconômica e de uso e cobertura da terra da área de influência dos reservatórios.

Para isso, realizou-se a consulta a dados em instituições de administração pública e de pesquisa e de projetos multi-institucionais que fossem responsáveis em gerar a informação geoespacial demandada por este trabalho. Além disso, foi verificado se as instituições também disponibilizam esses dados no visualizador de mapas do Portal Brasileiro de Dados Geoespaciais (SIG Brasil) [9].

O preenchimento dos metadados foi realizado, quando possível, a partir da ficha de metadados disponibilizada pela instituição. Tanto na sua página quanto no catálogo de metadados do SIG Brasil [9]. No entanto, a descrição dos atributos foi realizada após o download dos dados. Para isso, verificou-se quais os atributos de cada dado e buscou-se na página da instituição as informações necessárias para os descrever.

Ressalta-se que a descrição dos atributos de dados vetoriais é toda informação que descreve o conteúdo dos campos da tabela de atributos. Enquanto que a dos arquivos matriciais é toda informação que descreve o significado do valor do pixel [10].

Finalmente, foi verificado o quanto a informação sobre os dados disponibilizados no *web site* por cada instituição permitiu preencher os campos da lista de metadados proposta e descrever os atributos desses dados.

### 2.1. Definição da lista de metadados

A lista metadados proposta foi baseada no mínimo de metadados obrigatórios no Perfil de Metadados Geoespaciais do Brasil (Perfil MGB) [8] e na Norma Brasileira de Referências [11]. Para a sua elaboração foi considerado que deveria haver a informação necessária para a correta utilização e citação dos dados. Outra necessidade foi o registro do mínimo de informações sobre a natureza e sobre a qualidade do dado disponibilizado. Além de, se possível, informar a responsabilidade pela autoria do dado dentro da instituição.

Dessa forma, foram elencados 15 campos a partir da lista sumarizada de MGB [12]. Dos quais 10 são obrigatórios (Ob) [8], 1 é condicional (Cn) - nesse caso, obrigatórios para os geoespaciais [8]. 3 opcionais (Op) [8], sendo 2 deles para referenciar os dados [11]. Finalmente, foi adicionado (Ad) 1 campo, que não consta na lista sumarizada, para informar a acesso online aos metadados do dado. Essa lista foi dividida em duas categorias de campos: a de informações sobre os dados e a de informações sobre o acesso a esses dados. Assim, essa lista ficou configurada como descrito a seguir:

Os campos sobre o dado, por sua obrigatoriedade, foram: Título (Ob); Data (Ob); Responsável (Ob); Idioma (Ob); Categoria Temática (Ob); Resumo (Ob); Formato de distribuição (Ob); Sistema de Referência (Ob); Extensão Geográfica (Cn); Resolução Espacial (Op); Acesso online (Op); Tipo de representação espacial (Ob) e Acesso online do metadado (Ad). Os campos sobre o acesso aos dados foram:

Responsável pelos metadados (Ob); Data dos Metadados (Ob).

O campo “Responsável” foi subdividido em: a) Autor, que foi a instituição responsável pelo recurso; b) Produção, que foi o responsável pela produção do recurso dentro dessa instituição e c) Origem, no caso de a origem ser um projeto institucional, como é comum na Embrapa. Adicionalmente, essa divisão foi pensada para facilitar referenciar o dado.

## 3. RESULTADOS E DISCUSSÃO

### 3.1. A lista de metadados

A lista de metadados sugerida foi utilizada para 40 arquivos. Desses, 39 foram no formato vetorial (36 shapefiles, 1 kml – que agrupa quatro temas) e 2 no formato matricial. Foi observado que o seu uso poderia ser aplicado em dados como relatórios, tabelas e documentos diversos, desde que os campos relativos exclusivamente aos dados geoespaciais fossem desconsiderados.

Os campos “Responsável pelos metadados” e “Data dos metadados” são obrigatórios para qualquer dado disponibilizado. Fornecem informação sobre o autor do metadado e do tempo que foi publicado. Contudo foi proposto nesse estudo utilizar primeiro para registrar o profissional da equipe que foi responsável pela coleta do dado e pelo cadastro dos seus metadados e o segundo para registrar a data em que esse dado foi acessado. Isso permitiu um registro sintético do dado coletado e da data de acesso para casos de referenciar o dado. Quanto à informação sobre a data do metadado, essa fica subentendida com anterior à da registrada nessa etapa.

Os campos “Resolução espacial” e “Acesso online” foram inseridos para atender requisitos de elaboração da referência do dado [11]. Por outro lado, é necessário conhecer a resolução espacial dado para garantir o seu uso correto [10].

O campo “Acesso online ao metadado” foi considerado para facilitar consultas posteriores. Além de permitir identificar se existe uma ficha de metadados disponível para consulta.

Finalmente, o campo “Status”, que é obrigatório, não foi considerado. Haja vista que o trabalho é de coleta de dados concluídos, ou seja, com status já definido. Contudo, se a coleta envolver dados que ainda estejam em elaboração, seria necessário informar o status do dado. Que poderia ser realizado em um campo de observações ou esse campo poderia ser adicionado à lista proposta.

Em síntese, a lista apresentada corresponde ao mínimo de informações do dado para o seu registro e busca em uma base de dados estruturada. Que pode ser utilizada tanto para cadastrar os dados utilizados no projeto, quanto para permitir consultas posteriores à base de dados cadastrada. Contudo recomendamos que haja um campo de observações para uso de quem coleta o dado, quando necessário.

### 3.2. Análise dos dados

**Tabela 1 – Dados coletados nas páginas web dos provedores.**

Instituição	Dado
1	Regiões hidrográficas
2	Sub bacias hidrográficas
3	Unidade de planejamento hídrico
4	Bacias Ottocodificadas nível 2
5	Comitê de bacias estaduais
6	Comitê de bacias interestaduais
7	Comitê de bacias únicos
8	Entidades com funções de agência de água
9	Ano de conclusão de planos estaduais de recursos hídricos
10	Enquadramento dos rios federais
11	Planos estaduais de recursos hídricos
12	Planos interestaduais de recursos hídricos
13	Fiscalização de barragens
14	Qualidade da água - Média da série disponível entre 2001 e 2014
15	Demanda Bioquímica de Oxigênio - média da série disponível entre 2001 e 2014
16	Fosforo total - média da série disponível entre 2001 e 2014
17	Oxigênio Dissolvido média da série disponível entre 2001 e 2014
18	Turbidez - média da série disponível entre 2001 e 2014
19	Volume armazenado dos reservatórios no início do ano de 2015
20	Precipitação média anual – série histórica de 1977 a 2006
21	Volume armazenado nos reservatórios no fim do ano de 2015
22	Eventos extremos de inundação nos municípios no ano de 2015
23	Eventos extremos de seca nos municípios no ano de 2015
24	Disponibilidade hídrica superficial
25	Vazão outorgada para a indústria
26	Vazão outorgada irrigação
27	Vazão outorgada outros usos
28	Vazão outorgada para abastecimento público
29	Florestas Públicas
30	Unidades de Conservação Ambiental
31	Zonas de Conservação de peixes em Santa Catarina
32	Hidrovias do Brasil
33	Elementos de transporte aquaviário
34	Terras indígenas em 2017
35	Terras indígenas em estudo em 2017
36	Dados geomorfométricos do Brasil
37	Limites estaduais
38	Limites municipais
39	Cobertura e Uso de Solo do Brasil de 1985 a 2017

Dos 39 dados coletados, 28 foram da Agência Nacional de Águas (ANA) [12], 4 do Ministério do Meio Ambiente (MMA) [13], 4 da Agência Nacional de Transporte Aquaviário (ANTAQ) [14], 2 da Fundação Nacional do Índio (FUNAI) [15], 1 do Instituto Nacional de Pesquisa Espacial (INPE) [16], 2 do Instituto Brasileiro de Geografia e Estatística (IBGE) [17] e 1 do projeto de Mapeamento Anual da Cobertura e Uso do Solo do Brasil (Projeto Mapbiomas) [18] (Tabela 1).

Os metadados sugeridos foram identificados na maioria dos dados geoespaciais disponibilizados pelas instituições pesquisadas. Também foi possível descrever os atributos para a maioria desses dados. Contudo, foi possível tanto preencher todos os campos de metadados quanto descrever todos os atributos para menos da metade dos dados (Tabela 2). Adicionalmente, a soma da quantidade de atributos desses dados foi de 499, dos quais, foi possível descrever 84% (422 atributos).

**Tabela 2 – Número de dados com metadados informados e atributos descritos.**

Total de dados	Preencheu Metadados	Preencheu Atributos	Preencheu Ambos
40	32	28	20
(100%)	(80%)	(70%)	(50%)

As informações sobre os campos “Sistema de Referência”, “Data”, “Resumo” - que são obrigatórios - e “Resolução Espacial” foram identificadas em pouco mais de 80% desses dados (Tabela 3). Em relação ao “Datum”, apesar da indicação do uso do SIRGAS 2000 para o Brasil [19], 27 dados foram disponibilizados em WGS 84, 5 em SIRGAS 2000, 4 em SAD 69 e 4 não tiveram essa informação disponível online.

**Tabela 3 – Metadados que foram informados para parte dos dados coletados.**

Metadado	Obrigatório	Porcentagem de preenchimento
Resolução espacial	Não	90,0
Sistema de referência	Sim	90,0
Data	Sim	90,0
Resumo do dado	Sim	90,0
Acesso online aos Metadados	Não	92,5

Em nível de instituição, 6 informam os campos “Sistema de Referência”, “Resolução Espacial” e “Data”, 4 fornecem o “Acesso Online aos Metadados” e 2 fornecem seus dados em SIRGAS 2000.

Assim, foi possível coletar a maioria das informações sobre os dados geoespaciais no ambiente web onde eles estão disponibilizados. Contudo, na maioria dos casos, foi necessário um esforço extra para a identificação dos metadados e para a descrição dos atributos. Pois esses, apesar de estarem disponíveis, encontram-se em várias seções das

páginas web ou em documentos disponibilizados pelo provedor.

Ocorreram quatro formas dessa disponibilização: 1) em um cadastro de metadados acessado na página do dado; 2) dispersos no conteúdo do site e de publicações fornecidas; 3) no cadastro de metadados da INDE ou 4) tanto em cadastro no site quanto no da INDE.

#### 4. CONCLUSÕES

Neste trabalho foi verificado que é possível identificar a os metadados mínimos recomendados no ambiente web da maioria dos provedores de dados que foram acessados. Também foi possível descrever a maioria dos atributos listados. Sendo, essa descrição mais fácil para dados matriciais. Contudo, as informações ausentes sobre metadados e atributos poderiam ser conseguidas consultando a instituição provedora.

Com esse tipo de trabalho, temos mais segurança tanto para disponibilizar essa informação aos outros membros da equipe, quanto para gerar os produtos que utilizarão esses dados. Pois, esse tipo de registro dos metadados e dos atributos permite conhecer com precisão a natureza da informação consultada. Inclusive, evitar o uso de dados sobre os quais haja quaisquer dúvidas, com uma argumentação sólida e consistente.

Essa etapa do trabalho permite confirmar a adequabilidade do dado para os propósitos do projeto AQUA-BNDES e evitar erros na sua aplicação. Além de viabilizar a citação correta das autorias. Finalmente, o registro estruturado dos metadados permitirá implementar uma base de dados com a possibilidade de consulta à informação armazenada a partir dessa base, sem a necessidade de acessar diretamente o local de armazenamento de dados do projeto.

#### 5. AGRADECIMENTOS

Os autores estendem sinceros agradecimentos ao Projeto BRS/AQUA, projeto componente Manejo e Gestão Ambiental da Aquicultura, com o apoio financeiro do BNDES

#### 6. REFERÊNCIAS

- [1] FAO. The State of World Fisheries and Aquaculture 2018 - Meeting the sustainable development goals. FAO. Rome, Italy: FAO. 19540EN: 227 p. 2018
- [2] OSTRENSKY, A.; BORGHETTI, J.R. & SOTO, D. Aquicultura no Brasil: o desafio é crescer. Brasília: FAO, 2008. 276p.
- [3] BOYD, C. E. ZAJICEK, P. W.; HARGREAVES, J. A., JENSEN, G. L. Development, implementation and verification of better management practices for aquaculture. In: TUCKER, C. S.; HARGREAVES, J. A. (Ed.). Environmental best management practices for aquaculture. Oxford: Wiley-Blackwell, p. 129-149, 2008.
- [4] BOYD, C. E.; QUEIROZ, J. F. de; McNevin, A. Perspectives on the responsible aquaculture movement. World Aquaculture, Baton Rouge, v. 44, n. 4, p. 14-21, 2013.
- [5] Empresa Brasileira de Pesquisa Agropecuária (Embrapa).. Projeto Manejo e gestão ambiental da aquicultura. Embrapa Meio Ambiente. 2017. Disponível em: < <https://www.embrapa.br/meio-ambiente/busca-de-projetos/-/projeto/214226/manejo-e-gestao-ambiental-da-aquicultura> >. Acesso em 01/10/2018
- [6] BRASIL. Portaria Nº 153, de 14 de agosto de 2018. Institui a Rede Nacional de Pesquisa e Monitoramento Ambiental da Aquicultura em Águas da União. Brasília, DF: Diário Oficial da União: 17/08/2018, 1 p. 2018.
- [7] BRASIL. Decreto N 6.666, de 27 de novembro de 2008. Institui, no âmbito do Poder Executivo federal, a Infra-Estrutura Nacional de Dados Espaciais - INDE, e dá outras providências. Diário Oficial da União - Seção 1 - 28/11/2008, Página 57
- [8] Comissão Nacional de Cartografia (CONCAR). Perfil de Metadados Geoespaciais do Brasil – Perfil MGB. Brasília: Ministério do Planejamento, 2009. 194p.
- [9] Infraestrutura Nacional de Dados Espaciais (INDE).-Disponível em: <<http://www.inde.gov.br/geo-servicos/visualizador-de-mapas>>. Acesso em: 11 out. de 2018.
- [10] BURROUGH, P. A.; MCDONNELL, R. A. Principles of geographical information systems. New York: Oxford University Press, 1998. 333 p.
- [11] Associação Brasileira de Normas Técnicas (ABNT). NBR-6023: Informação e documentação - Referências - Elaboração. Rio de Janeiro, 2002. 24p.
- [12] Agência Nacional de Águas (ANA). Dados Abertos da Agência Nacional de Águas. Dados Abertos para a Gestão de Recursos Hídricos. Disponível em: <<http://dadosabertos.ana.gov.br/>>; Acesso em: 20 Ago. 2018.
- [13] Ministério do Meio Ambiente (MMA). Data Download; Disponível em: <<http://www.mma.gov.br/>>; Acesso em: 20 de ago. 2018.
- [14] Agência Nacional de Transporte Aquaviário (ANTAQ). Informações Geográficas. Disponível em <<http://portal.antaq.gov.br/index.php/informacoes-geograficas/>>. Acesso em 11 out. 2018
- [15] Fundação Nacional do Índio (FUNAI). Geoprocessamento. Shape; Disponível em: <<http://www.funai.gov.br/index.php/shape>>; Acesso em: 20 ago. de 2018.
- [16] Instituto Nacional de Pesquisa Espacial (INPE). TOPODATA: banco de dados geomorfológicos do Brasil. São José dos Campos: DSR, INPE. 2011. Resolução:30m. Disponível em: <<http://www.dsr.inpe.br/topodata/>>. Acesso em: 15 ago. 2018.
- [17] Instituto Brasileiro de Geografia e Estatística (IBGE). Download. Geociências. Disponível em: <[https://downloads.ibge.gov.br/downloads\\_geociencias.htm](https://downloads.ibge.gov.br/downloads_geociencias.htm)> Acesso em 11 out. 2018.
- [18] Projeto MapBiomias. Coleção 3 da Série Anual de Mapas de Cobertura e Uso de Solo do Brasil. Disponível em <[http://mapbiomas.org/pages/database/mapbiomas\\_collection\\_download](http://mapbiomas.org/pages/database/mapbiomas_collection_download)>. Acesso em 7 ago. 2018.
- [19] Fundação Instituto Brasileiro de Geografia e Estatística, “Resolução da Presidência 01/2015”, 27/09/2018.