

TIPOLOGIAS DE DADOS ESPACIAIS NA EMBRAPA: UMA PROPOSTA CONCEITUAL PARA A SUA IDE - GEOINFO

DANIELA MACIEL PINTO¹
DEBORA PIGNATARI DRUCKER²
MARGARETH SIMÕES⁴
JÚLIO CÉSAR DALLA MORA ESQUERDO³
RICARDO DE OLIVEIRA DART⁵
JOÃO MARCELO DOS SANTOS MARQUES⁶
FRANCISCA RASCHE⁷
JULIANA MEIRELES FORTALEZA⁸
DAVI DE OLIVEIRA CUSTÓDIO⁹

1. EMBRAPA TERRITORIAL, CAMPINAS
E-MAIL: DANIELA.MACIEL@EMBRAPA.BR

2. EMBRAPA INFORMÁTICA, CAMPINAS, SP.
E-MAIL: DEBORA.DRUCKER@EMBRAPA.BR

3. EMBRAPA SOLOS, RIO DE JANEIRO, RJ.
HTTP://LATTES.CNPQ.BR/0134693605452082
E-MAIL: MARGARETH.SIMOES@EMBRAPA.BR

4. EMBRAPA INFORMÁTICA, CAMPINAS, SP.
E-MAIL: JULIO.ESQUERDO@EMBRAPA.BR

5. EMBRAPA SOLOS - RIO DE JANEIRO, RJ.
E-MAIL: RICARDO.DART@EMBRAPA.BR

6. DEPARTAMENTO DE INFORMÁTICA APLICADA, UNIVERSIDADE FEDERAL DO ESTADO DO RIO DE JANEIRO (UNIRIO). ANALISTA DE PLANEJAMENTO DE INFRAESTRUTURA DE INFORMAÇÕES GEOGRÁFICAS E ESTATÍSTICAS NO IBGE
E-MAIL: JOAO.MARQUES@UNIRIOTEC.BR

7. FRANCISCA RASCHE
EMBRAPA FLORESTAS, COLOMBO, PR.
E-MAIL: FRANCISCA.RASCHE@EMBRAPA.BR

8. JULIANA MEIRELES FORTALEZA
EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA (EMBRAPA) - BRASÍLIA, DF – BRASIL.
E-MAIL: JULIANA.FORTALEZA@EMBRAPA.BR

9. EMBRAPA TERRITORIAL - CAMPINAS, SP
DAVI.CUSTÓDIO@EMBRAPA.BR

espaciais, visando sua disponibilização à sociedade brasileira. Dentre eles, a Empresa Brasileira de Pesquisa Agropecuária (Embrapa) que lançou, em 2018, o GeoInfo, sua Infraestrutura de Dados Espaciais, aderente às orientações estabelecidas no referido Decreto, nas premissas de acesso aberto e ao novo paradigma da comunicação científica, o *e-Science*.

Os dados inseridos no GeoInfo pressupõem requisitos de qualidade, os quais, de acordo com Sayão e Sales (2015), identificam o valor atribuído às propriedades da informação disponibilizada, possibilitando identificar adequadamente os atributos dos dados disponíveis e averiguar inconsistências. Explicita-se o amplo conceito de qualidade dos dados, em que várias abordagens devem ser aplicadas durante todo o ciclo de desenvolvimento de um projeto de pesquisa, enfatizando-se a representação dos dados a partir de seus metadados. Desta forma, o objetivo deste trabalho é apresentar uma exploração conceitual a respeito de tipologias dos dados espaciais disponibilizados pela Embrapa, considerando-se a documentação dos dados espaciais (metadados) como um dos aspectos de qualidade para garantir usos atuais e futuros aos dados divulgados pela instituição. Assim, buscou-se a categorização dos diversos conjuntos de dados espaciais disponíveis no GeoInfo, com foco em sua vinculação à INDE. Adicionalmente, objetiva-se orientar pesquisadores, gestores de dados espaciais e equipes de pesquisa na Embrapa e parceiros para o uso e publicação de dados no GeoInfo.

Baseia-se nos conceitos de “Dados de pesquisa” e “Dados para a pesquisa” (Concordat on Open Research, 2016) e na conceituação sobre fontes de informação da área de Ciência da Informação (Pinheiro, 2006), destacando-se as “Fonte primária”, “Fonte secundária” e “Fonte Terciária”. Tomando-se o conceito da qualidade, a partir dos metadados, trabalhou-se a integridade da descrição dos dados, enfatizando-se os atributos referentes à metodologia para a produção dos dados espaciais (campo “Declaração da qualidade dos dados”), seu produtor (campo “Autores”) e a disponibilização dos dados para a INDE (campo “INDE”). A partir disso, são propostas três categorizações para as tipologias de dados espaciais da Embrapa, apresentando-se o tratamento para cada uma delas, frente aos atributos citados. As categorias de dados são: 1. **Dado primário**: refere-se ao dado gerado pelas Unidades da Embrapa, e parceiros, fruto da coleta, levantamento de campo, organização e/ou espacialização de dados brutos, para desenvolvimento da pesquisa científica. Dados com componente espacial, processados ou não, que representam informação inédita (levantamento de solos, Inventário florestal, mapas de produtividade advindos de agricultura de precisão, dados de sensores proximais, dados de instrumentação, dados advindos de drones, entre outros). No campo “Declaração da Qualidade dos dados” deve-se conter informações sobre a produção do dado em relação à coleta, levantamentos de campo, organização e/ou espacialização, devendo-se preencher todos os procedimentos adotados para a geração dos dados. No campo “Autores” devem ser inseridos os responsáveis na Embrapa, e parceiros (caso existam), responsáveis pela geração dos dados. Campo “INDE”: deve ser disponibilizado para a INDE. 2. **Dado derivado com aporte intelectual**: refere-se ao dado gerado a partir de sua agregação a outros dados, mas cuja integração foi fruto do esforço intelectual da equipe envolvida na pesquisa. O aporte pode ocorrer pela aplicação de metodologia científica ou alguma técnica que possa adicionar valor ao dado (classificação, mineração, modelagem, geoestatística etc). Dados com componente espacial, processados, que representam informação inédita. O campo “Declaração da Qualidade dos dados” deve conter informações sobre a linhagem do dado (incluindo referências no formato ABNT 6023) e os procedimentos para agregação ao dado da fonte original. No campo “Autores” devem ser adicionados os responsáveis, na Embrapa, e parceiros (caso existam), responsáveis pela geração dos dados. Campo “INDE”: deve ser disponibilizado para a INDE; 3. **Dado espacializado de fontes externas**: refere-se ao dado espacializado, obtido a partir de fontes oficiais, referências bibliográficas, sites etc, que recebe uma componente espacial, mas que não são alterados quanto à sua natureza informacional. Dados que não representam informação inédita (Tabela com dados censitários; Tabelas da Produção Agrícola Municipal (PAM), que apresentam geocódigo e que podem ser processada em um SIG). Campo “Declaração da Qualidade dos dados”: deve conter informações sobre a linhagem do dado, sendo preenchido a partir do referenciamento da fonte utilizada (segundo a ABNT 6023). Mencionar o responsável pela espacialização dos dados. Campo “Autores”: incluir a instituição, órgão, departamento, setor etc, responsável pela geração dos dados espacializados. Campo “INDE” este tipo de dado não deve ser disponibilizado para a INDE.

Buscou-se, sobretudo, apoiar as discussões acerca da qualidade do dado, a partir dos metadados, respeitando-se a propriedade intelectual dos mesmos quanto a sua produção e representação. Constatou-se que somente os dados "Primários" e "Dados Derivados com aporte intelectual" têm a relevância necessária para sua disponibilização na INDE, possibilitando à sociedade brasileira ter ciência da tipologia dos dados da Embrapa, que podem ser consumidos para geração de novos geoserviços, análises de dados e outras necessidades.

REFERÊNCIAS

BRASIL. Decreto n. 6.666, de 27 de novembro de 2008. Institui, no âmbito do Poder Executivo federal, a Infraestrutura Nacional de Dados Espaciais - INDE, e dá outras providências. **Diário Oficial [da] República Federativa do Brasil**, Poder Executivo, Brasília, DF, 28 nov. 2008. p. 57.

CONCORDAT OPEN RESEARCH DATA. 2016. Disponível em: <<https://www.ukri.org/files/legacy/documents/concordatonopenresearchdata-pdf/>>. Acesso em: 15 jun. 2020.

PINHEIRO, L. V. R. Fontes ou recursos de informação: categorias e evolução conceitual. **Pesquisa Brasileira em Ciência da Informação e Biblioteconomia**, v. 1, n. 1, 2006. Disponível em: <<https://ridi.ibict.br/bitstream/123456789/10/1/pbciblena.pdf>>. Acesso em: 29 jul. 2020.

SAYÃO, L. F.; SALLES, L. F. **Guia de gestão de dados de pesquisa para bibliotecários e pesquisadores**. Rio de Janeiro: CNEN, 2025. Disponível em: <>. Acesso em: 22 jul. 2020.