

A qualidade dos dados é avaliada considerando duas dimensões: completude e consistência [Sattler2009; DAMAUKWGQD2013].

A completude, `dados_completude`, mede quão bem os dados de solo contidos na planilha cobre a totalidade de dados de solo gerados pelo trabalho ou documento de origem. Em outras palavras, a completude é uma medida da proporção dos dados de solo existentes que estão contidos na planilha. Assim, ela aborda casos em que nem todos os eventos, camadas, variáveis ou repetições estão registrados na planilha. Medir objetivamente essa cobertura é uma tarefa complexa, razão pela qual são usadas classes de completude conforme descrito a seguir:

- Incompletos: menos de 70% de cobertura dos dados originais
- Aproximadamente completos: entre 70 e 95% de cobertura dos dados originais
- Virtualmente completos: mais de 95% de cobertura dos dados originais
- Completos: 100% de cobertura dos dados originais

Quando da submissão de dados de solo ao SoilData, a atribuição da classe de completude é de inteira responsabilidade do provedor dos dados. À medida que os dados de solo faltantes forem sendo adicionados, o autor das edições tem a responsabilidade de garantir a validade da classe de completude, atualizando-a quando pertinente.

A consistência, `dados_consistencia`, refere-se ao quão bem os dados atendem às restrições e regras estruturais e documentais do SoilData. Três níveis de consistência são considerados: básica, estrutural, e documental (Tabela 5). Esses níveis são incrementais, o que significa que quando os dados atingem o nível de consistência documental, ele também atende aos níveis inferiores de consistência básica e estrutural.

O SoilData recebe apenas dados que atendam ao nível básico de consistência. Cabe aos provedores dos dados garantir que esse nível de consistência seja atingido. Assim como para `dados_completude`, `dados_consistencia` será atualizado

à medida que os dados forem sendo editados para alcançar os níveis superiores de consistência.

Tabela 1. Campos utilizados para registro da qualidade dos dados.

campo_id	campo_descricao	campo_exemplo
dados_completude	Proporção aproximada dos dados contidos nos documentos ou trabalhos de origem que estão cadastrados na presente planilha eletrônica. Escolha uma opção válida da lista suspensa.	incompletos (menos de 70% de cobertura)
dados_consistencia	Nível de atendimento às restrições e regras estruturais e documentais do FEBR. Este campo pode ser deixado em branco, pois o nível de consistência é atribuído aos dados pelos gestores do FEBR.	consistência estrutural

Tabela 5. Níveis de consistência de conjuntos de dados.

Nível	Definição
Básica	Atende às restrições e regras básicas: - Dados estão armazenados em planilha eletrônica - Tabelas <code>identificacao</code> e <code>versionamento</code> estão presentes
Estrutural	Atende às restrições e regras básicas e estruturais: - Tabelas <code>metadado</code> , <code>observacao</code> , e <code>camada</code> estão presentes - Tabelas possuem campos obrigatórios - Identificadores de observações e camadas são únicos
Documental	Atende às restrições e regras básicas, estruturais e documentais: - Codificação padrão de variáveis e unidades de medida - Métodos analíticos estão documentados

Figura 1. Visão da tabela IDENTIFICACAO com os dois campos de qualidade dos dados preenchidos (linhas 19 a 20).

FEBR Modelo de Dados v2

Arquivo Editar Ver Inserir Formatar Dados Ferramentas Extensões Ajuda

150% R\$ % 123 Inconsolata 11

	A	B
1	campo	valor
2	dados_id	
3	dados_titulo	
4	dados_autor	
5	organizacao_nome	
6	organizacao_local	
7	organizacao_url	
8	dados_descricao	
9	dados_idioma	
10	area_conhecimento	
11	palavras_chave	
12	dados_publicacao	
13	fonte_financeira	
14	dados_licenca	
15	extensao_temporal	
16	extensao_especial	
17	extensao_altimetrica	
18	extensao_pedometrica	
19	dados_completude	incompletos (menos de 70% de cobertura)
20	dados_consistencia	consistência estrutural

Adicionar mais 1000 linhas ao fim.

+ LEIAME identificacao campanha metadado evento camada versionamento dicionario