

ESTRUTURAÇÃO DE UM MODELO PARA AVALIAÇÃO DA QUALIDADE DE METADADOS AGROMETEOROLÓGICOS

Milena de Jesus Maria¹, Wesley Vinicius Fernandes², Maria Salete Marcon Gomes Vaz³

¹Universidade Estadual de Ponta Grossa, Ponta Grossa, Brasil
(milena.mjesus@gmail.com)

²Universidade Estadual de Ponta Grossa, Ponta Grossa, Brasil (240401400010@uepg.br)

³Universidade Estadual de Ponta Grossa, Ponta Grossa, Brasil (salette@uepg.br)

Resumo: Este trabalho tem como objetivo desenvolver um framework conceitual para a avaliação da qualidade de metadados em dados agrometeorológicos. A pesquisa foca em superar barreiras de interoperabilidade entre os padrões ISO 19115, Darwin Core e WCMP. A metodologia envolveu auditoria técnica e formulação de métricas de verificação qualitativa. Como resultados parciais, construiu-se uma Matriz de Critérios de Qualidade para apoiar a redução de ilhas de dados.

Palavras-chave: Metadados; Qualidade da Informação; Dados Agrometeorológicos; Engenharia de Ontologias; Interoperabilidade Semântica.

INTRODUÇÃO

A crescente digitalização da agricultura e o avanço das técnicas de agricultura de precisão aumentaram significativamente a quantidade de dados gerados por sensores, satélites e estações meteorológicas. Para apoiar decisões estratégicas no campo, é necessário integrar dados heterogêneos de diferentes fontes em plataformas digitais agroalimentares (Urdu et al., 2024).

No entanto, a diversidade de padrões adotados para a descrição destes dados, como o ISO 19115 para dados geoespaciais, o Darwin Core para registros de biodiversidade e o WCMP para variáveis climáticas, cria grandes barreiras de interoperabilidade técnica e semântica. Esse cenário resulta na formação das chamadas "ilhas de dados", gerando informações isoladas e de difícil reutilização em sistemas automatizados (AgGateway, 2023).

Para resolver este problema, os metadados são fundamentais na sistematização e no fornecimento de contexto aos recursos, garantindo que os dados sejam encontráveis, acessíveis, interoperáveis e reutilizáveis (Ulrich et al., 2022; Vogt et al., 2025). Contudo, a garantia de que esses dados estão aptos para integração não depende apenas da sua exatidão técnica. É preciso avaliar dimensões como a completude, a consistência representacional e a acessibilidade, conforme ressaltado por Wang e Strong (1996).

Dessa forma, este trabalho tem por objetivo desenvolver um framework conceitual para avaliar a qualidade dos metadados agrometeorológicos. O

processo foca na formulação de métricas de verificação que atestem a consistência e a estruturação da informação.

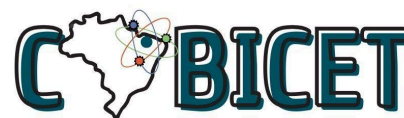
MATERIAL E MÉTODOS

O processo de desenvolvimento da pesquisa foi estruturado em quatro etapas práticas e sequenciais. O trabalho começou com a definição taxonômica, baseada no modelo de qualidade da informação de Wang e Strong (1996). Essa etapa selecionou dimensões de qualidade e as distribuiu nas categorias intrínseca, contextual, representacional e de acessibilidade.

Em seguida, a pesquisa focou na extração de requisitos por meio da auditoria técnica minuciosa da documentação oficial dos padrões ISO 19115, WCMP e Darwin Core. O objetivo dessa fase foi identificar os campos e propriedades críticos para a interoperabilidade agrometeorológica.

O processo continuou com a formulação dos critérios analíticos, transformando os requisitos técnicos extraídos em métricas de validação binárias ou escalares, aplicadas através de perguntas de verificação direta. Por fim, a consolidação dos dados ocorreu com o preenchimento de uma matriz de critérios de qualidade.

A execução dessa matriz foi feita de forma isolada para cada padrão auditado, visando garantir a profundidade técnica necessária na análise estrutural comparativa.



RESULTADOS E DISCUSSÃO

O processo de análise técnica dos padrões de metadados gerou a estruturação prática da avaliação de qualidade. A aplicação do modelo de Wang e Strong (1996) permitiu delimitar a Qualidade Intrínseca para medir a Acurácia e a Reputação da fonte. A Qualidade Contextual englobou a Completude e a Temporalidade do dado. Para a Qualidade Representacional, definiram-se a Consistência de formatos e a Interpretabilidade. Por fim, a Qualidade de Acessibilidade mediu a facilidade de recuperação e a segurança. A partir dessa base teórica, a auditoria revelou características específicas de cada padrão que impactam diretamente a integração das informações.

No escopo do padrão Darwin Core, focado em registros de biodiversidade, o objetivo foi avaliar a precisão da identificação biológica e a granularidade da ocorrência individual. As inconsistências encontradas mostraram que o uso isolado do nome científico da espécie gera ambiguidades (Joo et al., 2021; TDWG, 2025). Para resolver esse desafio, elaboraram-se métricas que exigem o preenchimento da hierarquia taxonômica completa, como reino e família, viabilizando filtros hierárquicos precisos. A análise também considerou a natureza do registro para evitar a mistura de dados de observação humana com espécimes de museu, validando o uso estrito de vocabulários controlados.

No caso dos dados meteorológicos regidos pelo WMO Core Metadata Profile (WCMP), a prioridade foi a capacidade de descoberta e o acesso direto aos dados brutos. Observou-se que a cobertura espacial representa áreas inteiras de atuação dos sensores, o que exigiu a elucidação precisa do objeto geométrico para validar as formas de polígonos (WMO, 2025). Além disso, elaborou-se uma métrica estrita de análise para verificar se o intervalo de tempo de medição possui datas de início e fim definidas, conforme o padrão internacional ISO 8601-1:2019, uma condição básica para a filtragem de séries de tempo histórico a longo prazo. A avaliação incluiu ainda a verificação de links e pontos de distribuição de dados.

Em relação ao padrão ISO 19115 para dados geoespaciais, a auditoria priorizou a conformidade estrutural e a consistência semântica. Um ponto problemático sério foi detectado no preenchimento inadequado ou ausência de campos obrigatórios definidos pela norma, como o título, a data de criação e a categoria do tema. A ausência dessas informações essenciais classifica o metadado como estruturalmente quebrado (ISO, 2014; Han e Han, 2022). A análise também abrangeu os vocabulários controlados, resultando na criação de métricas para garantir que os valores informados pertençam

estritamente às listas oficiais da ISO, evitando o uso indevido de texto livre em propriedades padronizadas.

A partir da consolidação desses requisitos extraídos e dos problemas mapeados, estabeleceu-se a matriz de critérios de qualidade. O detalhamento dos resultados demonstrou que as barreiras para a interoperabilidade agrometeorológica vão muito além de questões de formatação técnica. Identificou-se uma grande incompatibilidade informacional que gera as chamadas ilhas de dados (Meyer et al., 2021), com destaque para a diferença de granularidade entre os modelos. Enquanto o WCMP descreve as variáveis climáticas em um nível macro de pacotes de dados, o Darwin Core detalha informações biológicas em um nível micro de ocorrências individuais.

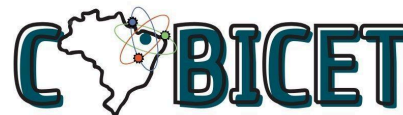
Estes resultados parciais consolidam a base estrutural do framework. Atualmente, o modelo encontra-se em transição para a sua fase empírica, na qual a matriz de critérios desenvolvida será aplicada e testada em repositórios reais de dados agrometeorológicos.

CONCLUSÃO

Os resultados parciais demonstram que as barreiras de interoperabilidade agrometeorológica vão além de questões de formatação técnica, envolvendo profundas incompatibilidades informacionais e diferenças de granularidade entre os padrões avaliados. A estruturação da matriz de critérios evidenciou que a harmonização conceitual e a padronização rigorosa — fundamentadas em dimensões de qualidade como completude e consistência representacional — são imperativas para reduzir as "ilhas de dados" e viabilizar a integração semântica em plataformas digitais. Nas próximas etapas da pesquisa, o framework conceitual será aplicado empiricamente em conjuntos reais de metadados provenientes de fontes abertas, como INMET, CPTEC/INPE e FAO. Essa fase prática objetiva validar o modelo proposto, atestar sua aplicabilidade em cenários reais e identificar oportunidades de aprimoramento para a gestão de dados na agricultura digital.

AGRADECIMENTOS

Agradeço ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela concessão da bolsa de Iniciação Científica, fomento essencial para a realização desta pesquisa, e à Universidade Estadual de Ponta Grossa (UEPG) pelo suporte institucional. Expresso também minha profunda gratidão à minha orientadora, Prof.^a Dr.^a Maria Salete Marcon Gomes Vaz, e ao meu coorientador, Wesley Vinicius Fernandes, pelo constante apoio, direcionamento e pelas valiosas contribuições ao longo de todo o desenvolvimento deste trabalho.



REFERÊNCIAS

- AGGATEWAY. Interoperability in Agriculture: 2023 Status Report. AgGateway, 2023.
- HAN, S.; HAN, J. Case study on an integrated interoperable metadata model for geoscience information resources. *Geoscience Data Journal*, v. 9, p. 355-370, 2022.
- ISO. ISO 19115-1: Geographic information – Metadata – Part 1: Fundamentals. ISO, Geneva, 2014.
- ISO. ISO 8601-1:2019: Date and time – Representations for information interchange – Part 1: Basic rules. ISO, Geneva, 2019.
- JOO, M. et al. Building core vocabulary of agriculture activity for agricultural data integration. *Lecture Notes in Computer Science*, v. 12612, p. 301-316, 2021.
- MEYER, R. et al. Aligning Standards Communities: Sustainable Darwin Core MIxS Interoperability. *Biodiversity Information Science and Standards*, v. 5, p. e73775, 2021.
- TDWG. Darwin Core RDF guide. Biodiversity Information Standards, 2025.
- ULRICH, H. et al. Understanding the Nature of Metadata: Systematic Review. *Journal of Medical Internet Research*, v. 24, p. e25440, 2022.
- URDU, D. et al. Aligning interoperability architectures for digital agri-food platforms. *Computers and Electronics in Agriculture*, v. 224, p. 109194, 2024.
- VOGT, L. et al. Suggestions for extending the FAIR Principles based on a linguistic perspective on semantic interoperability. *Scientific Data*, v. 12, p. 688, 2025.
- WANG, R. Y.; STRONG, D. M. Beyond accuracy: what data quality means to data consumers. *Journal of Management Information Systems*, v. 12, p. 5-33, 1996.
- WIECZOREK, J. et al. Darwin Core: An Evolving Community-Developed Biodiversity Data Standard. *PLoS ONE*, v. 7, p. e29715, 2012.
- WMO. WMO Core Metadata Profile (WCMP): Version 2. WMO, Geneva, 2025.
- WMO. WMO-No. 1060: Manual on the WMO Information System. WMO, Geneva, 2023.