

Qualidade de dados, score e classificação informacional

Fundamentos metodológicos para dossiês de imóveis informais

INORGIN — Instituto Nacional de Organização de Imóveis Informais

Estudo Técnico-Científico E-03/2026 | Edição 2026

Natureza da publicação. Estudo conceitual e metodológico. O modelo apresentado mede atributos da informação e não foi validado como instrumento preditivo de propriedade, valor, solvência, segurança jurídica ou probabilidade de regularização. Não constitui norma pública, certificação da ABNT, parecer ou avaliação imobiliária.

Resumo

O estudo apresenta fundamentos para avaliar a qualidade de dados em dossiês de imóveis informais e examina a separação metodológica entre Nível de Qualidade (NQ), score técnico e classificação por requisitos. A pesquisa é conceitual, baseada na literatura de qualidade da informação, assimetria informacional, gestão de riscos, legislação de proteção de dados e documentação do CDTI. Propõe-se que a unidade básica de avaliação seja a resposta vinculada à evidência, e não o documento tomado como selo autônomo. São examinadas as dimensões de completude, consistência, atualidade, proveniência, compreensibilidade, acessibilidade controlada e correspondência espacial. Demonstra-se que score e classificação cumprem funções diferentes: o score ordena o desempenho informacional segundo pesos; a classificação verifica conjuntos cumulativos de requisitos. O uso simultâneo pode melhorar a comunicação, mas também gerar falsa precisão, compensação indevida e viés se faltarem governança, versionamento e validação empírica. Recomenda-se um programa de testes de confiabilidade, validade, sensibilidade, equidade e calibração antes de atribuir aos indicadores qualquer interpretação além da qualidade do cadastro.

Palavras-chave: qualidade de dados; score técnico; classificação; Nível de Qualidade; governança informacional; imóvel informal.

1. Introdução

Dossiês imobiliários informais reúnem dados de naturezas distintas: declarações, certidões, contratos, fotografias, coordenadas, plantas e informações municipais. A simples presença de um arquivo não demonstra que o item esteja completo, atual, coerente ou relacionado ao objeto correto. Por isso, a avaliação precisa deslocar o foco da quantidade de documentos para a qualidade da informação produzida por eles.

O problema de pesquisa é definir como sintetizar qualidade sem transformar indicadores administrativos em aparência de certeza jurídica. A hipótese é que a separação entre avaliação do item, pontuação agregada e requisitos classificatórios reduz ambiguidades, desde que a metodologia seja transparente e o resultado permaneça rastreável às evidências.

O objetivo geral é propor fundamentos para uma arquitetura de mensuração aplicável a dossiês imobiliários. Os objetivos específicos são: definir dimensões; examinar o tratamento de ausência e não aplicabilidade; diferenciar score e classificação; identificar riscos metodológicos; e estabelecer agenda de validação para o CDTI.

2. Referencial teórico

2.1 Assimetria informacional

Akerlof (1970) demonstrou que a incerteza sobre a qualidade pode deteriorar mercados, porque o comprador não consegue distinguir bens melhores e reduz sua disposição a pagar. No contexto imobiliário informal, o problema não está apenas na qualidade intrínseca do direito, mas na impossibilidade de observar e comparar documentos, ocupação, localização e restrições.

Um dossiê padronizado atua como mecanismo de sinalização, porém o sinal somente é confiável se os critérios forem auditáveis. Um selo sem acesso à metodologia ou às fontes pode ampliar, em vez de reduzir, a assimetria.

2.2 Qualidade de dados além da exatidão

Wang e Strong (1996) organizaram a qualidade de dados em dimensões intrínsecas, contextuais, representacionais e de acessibilidade. A contribuição central é reconhecer que dado correto pode ser inadequado à tarefa, desatualizado, incompreensível ou inacessível.

Em dossiê imobiliário, a área transcrita fielmente de um contrato pode ser correta como transcrição e inadequada para representar a ocupação atual. A certidão registral pode ser autêntica, mas estar vencida para a finalidade da diligência. A fotografia pode ser nítida, porém sem data ou localização que permitam vinculá-la ao objeto.

2.3 Gestão de riscos

A ISO 31000:2018 estrutura a gestão de riscos como processo de identificação, análise, avaliação, tratamento, comunicação e monitoramento. Um score informacional não é, por si, matriz de risco. Risco envolve incerteza sobre objetivos, contexto, impacto e controles. Assim, a pontuação do cadastro deve ser um insumo, não substituto da análise jurídica, técnica ou transacional.

3. Método

O estudo utiliza construção conceitual e análise de artefato. Foram comparadas a literatura de qualidade de dados, as regras consolidadas do CDTI e requisitos de rastreabilidade, registro e proteção de dados. Não foram utilizados cadastros reais nem realizada análise estatística de desempenho.

Os critérios de avaliação do modelo são:

1. clareza do construto medido;
2. separação entre observação e inferência;
3. rastreabilidade do resultado;
4. tratamento explícito de ausência, divergência e não aplicabilidade;
5. resistência à compensação indevida;
6. versionamento e reprodutibilidade;
7. compatibilidade com proteção de dados e atribuições profissionais.

4. Unidade de avaliação: resposta vinculada à evidência

O documento não deve receber uma qualidade jurídica global. A avaliação incide sobre a resposta a uma pergunta. Uma matrícula pode sustentar adequadamente a identificação do titular registral, mas não responder sobre ocupação física ou regularidade da construção. Um contrato pode indicar preço e partes, porém descrever de forma insuficiente o objeto.

Cada item deve conter:

- pergunta e finalidade;
- tipo de resposta;
- evidência admitida;
- regra de completude;
- data de referência;
- responsável pela informação;
- estado de verificação;
- peso máximo, quando houver;
- relação com requisito classificatório;
- observação de divergência ou limitação.

Essa granularidade permite que a revisão altere apenas os itens afetados e preserve o histórico dos demais.

5. Dimensões propostas

Dimensão	Definição operacional	Exemplo de falha
Completude	presença dos elementos exigidos para a finalidade	contrato sem anexo ou assinatura necessária
Consistência	convergência lógica entre campos e fontes	áreas diferentes sem explicação
Atualidade	adequação temporal do dado	certidão antiga usada como situação vigente
Proveniência	identificação da fonte, emissor, data e responsável	arquivo recebido sem origem conhecida
Legibilidade e integridade	capacidade de leitura e ausência aparente de perda	página cortada, rasura ou baixa resolução
Correspondência	vínculo do dado com o objeto e as partes corretas	matrícula de outro imóvel ou foto não localizada
Compreensibilidade	representação clara para o usuário-alvo	sigla ou indicador sem definição
Acessibilidade	acesso proporcional à finalidade e permissão	exposição pública de CPF e assinatura
controlada		
Precisão espacial declarada	indicação do método e limite da localização	coordenada sem sistema de referência ou origem

As dimensões não são totalmente independentes. Documento legível pode ser desatualizado; informação completa pode ser inconsistente; dado preciso pode ter sido coletado sem necessidade. A avaliação deve conservar o diagnóstico dimensional antes de sintetizá-lo.

6. Nível de Qualidade (NQ)

As Normas INORGIN adotam quatro níveis: NQ100, NQ50, NQ25 e NQ0. Eles representam percentuais da pontuação possível do item. A escala é simples e operacional, mas precisa de âncoras descritivas específicas para cada pergunta.

Nível	Interpretação geral	Condição metodológica
NQ100	critério informacional integralmente atendido	evidência prevista, legível, coerente e suficiente para o item
NQ50	atendimento parcial relevante	parte substancial presente, com lacuna delimitada
NQ25	evidência limitada	indício ou condição mínima, insuficiente para atendimento parcial pleno
NQ0	não atendimento	dado ausente, incompatível, não informado ou alternativa sem pontuação

O NQ0 não deve misturar silenciosamente causas diferentes. A base precisa registrar motivo: “não

apresentado”, “não localizado”, “ilegível”, “divergente”, “não confirmado” ou “resposta negativa”. Para análise estatística, essas categorias possuem significados diferentes.

Item não aplicável também não equivale a NQ0. Se uma pergunta depende de construção existente e o terreno não possui edificação, a exclusão do denominador pode ser adequada. A regra precisa ser prévia, objetiva e versionada.

7. Score técnico

O score técnico agrega pontos por blocos e itens. Sua função legítima é resumir o grau de atendimento informacional segundo a matriz vigente. Em termos gerais:

$$[S = \sum_{b=1}^B \sum_{i=1}^{n_b} w_{\{bi\}} q_{\{bi\}}]$$

em que ($w_{\{bi\}}$) é o peso máximo do item e ($q_{\{bi\}}$) é o fator correspondente ao NQ. No CDTI, o teto total da Categoria A é 1000 pontos.

Três cuidados são essenciais. Primeiro, o peso expressa escolha normativa da metodologia, não importância jurídica universal. Segundo, a soma permite compensação: pontos em fotografias ou dados gerais podem elevar o total apesar de ausência documental crítica. Terceiro, diferenças pequenas de score podem não ter significado prático sem estudo de sensibilidade.

Por isso, o resultado deve apresentar total, pontuação por bloco, itens NQ0, versão da matriz e data. A exibição isolada do número reduz transparência.

8. Classificação por requisitos

A classificação Bronze, Prata e Ouro do CDTI é cumulativa e independente de faixas numéricas. Bronze exige todos os requisitos Bronze aplicáveis; Prata exige Bronze e Prata; Ouro exige os três conjuntos. Perguntas sem marca não alteram a classificação.

Esse desenho funciona como regra de barreira: um requisito indispensável não é compensado por pontos de outro bloco. A classificação responde “quais conjuntos mínimos foram atendidos?”, enquanto o score responde “quanto da matriz informacional foi pontuado?”.

Situação	Score	Classificação	Leitura correta
muitos itens acessórios e ausência de requisito Bronze	alto	sem Bronze	boa pontuação agregada, mas requisito mínimo não atendido
todos os requisitos Prata e poucos itens acessórios	médio	Prata	conjunto cumulativo atendido com score moderado
requisitos Ouro e documentos atuais	alto	Ouro	maior atendimento metodológico, sem garantia jurídica

A classificação não deve empregar denominações que sugiram domínio, risco baixo ou aprovação pública. Bronze, Prata e Ouro precisam aparecer acompanhados da expressão “classificação informacional CDTI”.

9. Validade e confiabilidade

9.1 Validade de conteúdo

Especialistas de diferentes áreas devem avaliar se os itens representam as dimensões relevantes. A revisão precisa incluir registral, jurídica, urbanística, técnica, corretagem, proteção de dados e experiência do usuário.

9.2 Confiabilidade entre avaliadores

Dois Certificadores diante do mesmo dossiê deveriam atribuir NQ semelhante. Isso pode ser testado com amostras duplicadas e coeficientes de concordância. Divergências indicam necessidade de melhorar âncoras, treinamento ou interface.

9.3 Validade de construto

Deve-se examinar se blocos supostamente distintos medem aspectos efetivamente diferentes. Correlações excessivas podem revelar redundância; baixa associação pode ser esperada entre documentação, fotografias e infraestrutura.

9.4 Validade preditiva

Somente pesquisa longitudinal pode verificar se score ou classificação predizem desfechos, como conclusão de regularização, redução de tempo de diligência ou menor incidência de divergências. Até que isso ocorra, é inadequado interpretar o score como probabilidade.

9.5 Sensibilidade

Pesos alternativos, tratamento de itens condicionais e mudança de um NQ devem ser simulados. Se pequenas alterações produzirem grande variação de resultado, a metodologia é instável e exige revisão.

10. Governança e versionamento

A plataforma deve armazenar, para cada cálculo: versão da matriz; pesos; requisitos; respostas; justificativas; documentos vinculados; usuário; data e hora. Mudança de metodologia não pode recalcular silenciosamente certificado histórico.

Recomenda-se governança com:

1. comitê metodológico multiprofissional;
2. processo documentado de proposta e aprovação de alteração;
3. consulta a usuários e análise de impacto;
4. publicação de nota técnica da versão;
5. auditoria de amostras e divergências;
6. canal de contestação e correção;
7. monitoramento de padrões anormais por Certificador ou região.

11. Proteção de dados, equidade e vieses

O dossiê contém dados pessoais e pode revelar situação econômica, familiar e territorial. A LGPD exige finalidade, adequação, necessidade, transparência, segurança e prevenção. A camada pública deve conter o mínimo necessário; anexos sensíveis precisam de controle de acesso, registro de consulta e retenção definida.

O sistema também deve investigar vieses. Regiões com menor capacidade administrativa podem apresentar menos documentos públicos, reduzindo score por fatores externos ao declarante. Certos tipos de ocupação possuem evidências diferentes das transações formais. A metodologia não deve

transformar desigualdade de acesso documental em julgamento moral ou exclusão automática.

12. Programa recomendado de validação

Propõe-se cinco fases:

1. **revisão de conteúdo:** painel de especialistas e teste cognitivo das perguntas;
2. **piloto controlado:** aplicação em conjunto diversificado de dossiês fictícios e reais autorizados;
3. **concordância:** dupla avaliação cega de amostra;
4. **análise estatística:** distribuição, teto, piso, redundância e sensibilidade dos pesos;
5. **acompanhamento longitudinal:** relação entre indicadores e desfechos operacionais claramente definidos.

Resultados devem ser publicados de forma agregada e, sempre que possível, anonimizados. Uma revisão independente acadêmica futura aumentaria a legitimidade, mas não deve ser anunciada antes de formalizada.

13. Considerações finais

Mensurar qualidade informacional é útil para tornar dossiês comparáveis, orientar pendências e apoiar auditoria. O desenho do CDTI avança ao separar NQ, score e classificação. Essa separação, contudo, só produz confiança se os usuários compreenderem o que cada indicador mede e, principalmente, o que não mede.

O princípio fundamental é a rastreabilidade: todo ponto e requisito deve retornar a pergunta, regra, evidência e versão. Sem essa trilha, o número substitui o diagnóstico; com ela, o indicador se torna síntese verificável de um cadastro técnico-declaratório.

Referências

- AKERLOF, George A. The market for “lemons”: quality uncertainty and the market mechanism. **The Quarterly Journal of Economics**, v. 84, n. 3, p. 488–500, 1970. DOI: 10.2307/1879431.
- ANPD — AUTORIDADE NACIONAL DE PROTEÇÃO DE DADOS. **Guia orientativo sobre segurança da informação para agentes de tratamento de pequeno porte**. Brasília, DF: ANPD, 2021.
- BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados Pessoais. Brasília, DF: Presidência da República, 2018.
- INORGIN. **Caderno Normativo: Normas Técnicas INORGIN, Série N-001/2026 a N-013/2026**. Versão consolidada 1.0. 2026.
- ISO — INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 31000:2018: Risk management — Guidelines**. Geneva: ISO, 2018.
- WANG, Richard Y.; STRONG, Diane M. Beyond accuracy: what data quality means to data consumers. **Journal of Management Information Systems**, v. 12, n. 4, p. 5–33, 1996. DOI: 10.1080/07421222.1996.11518099.

Identificação editorial: Estudo Técnico-Científico INORGIN E-03/2026.

Versão: 1.0 | **Data:** 12 de agosto de 2026.