Capítulo 2

O uso de Machine Learning para avaliação da qualidade de dados Linked Data

Ananda Fernanda de Jesus [1] *e José Eduardo Santarem Segundo* [2]

Introdução

O World Wide Web Consortium (W3C) é o responsável pelo desenvolvimento da Web e busca, desde o início dos anos 2000, estabelecer recomendações e elaborar instrumentos para incentivar e possibilitar a disponibilização, uso e reúso de dados nesse ambiente.

Um dos resultados desse processo é o conjunto de princípios denominado *Linked Data*, elaborado com a finalidade de permitir a conexão de dados na Web, com base no uso de identificadores únicos (*Uniform Resource Identifier* – URIs) e na adoção de uma estrutura conceitual (*Resource Description Framework* – RDF), facilitando o uso de agentes computacionais na busca por informações, bem como a reutilização de conjuntos de dados (Berners-Lee, 2006).

Embora essa estrutura confira aos dados certo nível de padronização e consistência sintática, os conjuntos de dados publicados como

[1] Doutoranda em Ciência da Informação pela Universidade Estadual Paulista – UNESP. E-mail: af.jesus@unesp.br. Lattes: http://lattes.cnpq.br/8553935355036462.

[2] Doutor em Ciência da Informação. Professor na Universidade de São Paulo – USP. E-mail: santarem@usp.br. ORCID: https://orcid.org/0000-0003-3360-7872. Lattes: http://lattes.cnpq.br/5562746387565465.

https://doi.org/10.36311/2025.978-65-5954-613-8.p35-48

Linked Data também estão sujeitos a questões relacionadas com a qualidade de dados.

A qualidade de dados pode ser definida com base em diferentes perspectivas, ou categorias de qualidade, sendo elas: contextual, intrínseca, representacional e acessibilidade (Nelson; Todd; Wixom, 2005; Nooghabi; Dastgerdi, 2016).

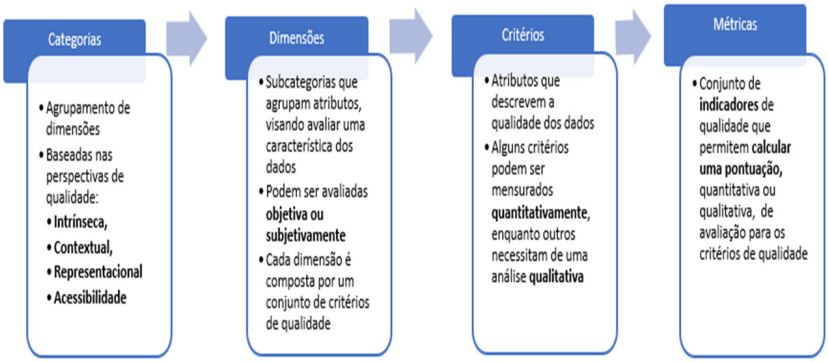
Em uma perspectiva contextual, a qualidade de um conjunto de dados é definida como "fitness for use". Nessa perspectiva, os problemas de qualidade de um conjunto de dados são identificados e discutidos levando em consideração o contexto e as atividades onde os dados serão aplicados (Juran, 1988; Wang; Strong, 1996; Zaveri *et al.*, 2012). Em uma perspectiva intrínseca, a qualidade de um conjunto de dados (*datasets*) está relacionada diretamente com as suas características sintáticas e semânticas, independentemente do contexto no qual serão aplicados. Já na perspectiva representacional, são avaliadas as questões de qualidade relacionada aos metadados elaborados para a sua representação. Em uma perspectiva de acessibilidade é levada em consideração a capacidade de acesso aos dados (Nelson; Todd; Wixon, 2005; Wang; Strong, 1996; Zaveri *et al.*, 2012).

A identificação de questões de qualidade e a seleção de conjuntos de dados para determinada aplicação ocorre por meio do processo de avaliação de qualidade. Esse processo pode ser compreendido por meio dos conceitos de: categorias, dimensões, critérios e métricas (Wang; Strong, 1996).

As categorias são o conceito mais abrangente, que refletem as principais perspectivas mencionadas anteriormente (contextual, intrínseca, representacional e acessibilidade). As dimensões são a expressão mais abrangente das características que serão observadas no processo de avaliação, cada dimensão é composta por um conjunto de critérios que descrevem a qualidade dos dados com base em um atributo específico. Para poder mensurar quão adequado um conjunto está em relação a determinado critério, precisam ser estabelecidas métricas, ou indicadores que permitem quantificar e qualificar um conjunto de dados em relação a determinada dimensão (Assaf; Senart; Troncy, 2016; Wang; Strong, 1996). A figura 1

apresenta os principais conceitos relacionados com o processo de avaliação de qualidade.

Figura 1 – Conceitos relacionados com a avaliação de qualidade



Fonte: Autores.

A avaliação de dados publicados como *Linked Data* demanda o uso de dimensões, critérios e métricas específicos, tanto pelo contexto no qual os dados são criados, como por sua estrutura pautada no uso do RDF, URIs e ligações entre diferentes fontes.

Ahmed (2017) aponta que um dos grandes desafios da qualidade de dados *Linked Data* é lidar com o volume e variedade dos dados, cuja qualidade varia de fonte para fonte. Cada fonte adota diferentes níveis de curadoria, a depender de suas necessidades.

Existem diferentes ferramentas para auxiliar no processo de avaliação de qualidade. Essas ferramentas podem ser automáticas, semiautomáticas ou manuais. As ferramentas muitas vezes são criadas em uma abordagem contextual, pensando em um cenário de aplicação específico, não atendendo a todas as dimensões e métricas do processo de avaliação de qualidade, fazendo com que seu uso em outros cenários seja inviável ou dependa de customização (Jesus; Santarem Segundo, 2022).

Em cenários com grandes quantidades de dados, o uso de ferramentas que dependam de uma abordagem semiautomática ou manual pode ser improdutivo ou mesmo inviável. Uma das possibilidades para processar

grandes quantidades de dados heterogêneos são as técnicas de *Machine Learning* (ML). Técnicas associadas à área da inteligência artificial e baseadas na

> […] construção de agentes computacionais capazes de aprender com a experiência, com base na aplicação de técnicas estatísticas, em especial, por meio de algoritmos, visando a identificação de padrões e a realização de predições (Jesus *et al.*, 2023, p. 168).

A aplicação depende do uso de algoritmos, agentes computacionais elaborados visando a identificação de padrões e que no geral

> […] tem uma base de treinamento, que permite que o algoritmo, ao entrar em execução, possa se utilizar dela, para gerar os resultados das análises. Assim, o *machine learning* tem sempre que consultar um conjunto de dados, e quanto melhor e com mais tempo de uso, o algoritmo tende a ser aprimorado e ter resultados mais efetivos (Coneglian, 2020, p. 127).

O processo de avaliação de qualidade é uma etapa importante da reutilização de dados publicados como *Linked Data*. Esse processo de avaliação pode ser custoso e depender da utilização de grandes recursos financeiros e humanos. Considerando o cenário apresentado, a presente pesquisa tem como objetivo verificar o potencial de aplicação das técnicas de ML para a avaliação de qualidade de dados *Linked Data*, partindo da hipótese de que essas técnicas podem contribuir para a eficiência do processo de avaliação, automatizando etapas repetitivas desse processo.

Essa pesquisa realiza um recorte no corpus de uma Revisão Sistemática da Literatura, buscando responder aos seguintes questionamentos: existem artefatos que utilizam técnicas de ML para avaliação de qualidade? Em caso positivo, em que contexto essas técnicas são aplicadas?

Procedimentos metodológicos

Essa pesquisa pode ser caracterizada por seus objetivos como exploratória e descritiva, com resultados quantitativos e qualitativos. Trata-se de um recorte da análise de dados de uma Revisão Sistemática da Literatura (RSL).

As RSLs são revisões de literatura, baseadas na documentação das decisões do pesquisador e orientadas por um protocolo de pesquisa, prezando pela sua reprodutibilidade e auditabilidade. Esse processo de documentação permite que o corpus teórico construído e os dados coletados possam ser analisados sob outras perspectivas, gerando novos resultados.

A condução da RSL teve como propósito identificar artigos que abordassem a qualidade de dados publicados como *Linked Data*. Entre os artigos aceitos foram identificados artefatos utilizados para avaliação de qualidade de dados *Linked Data*.

Para uma maior compreensão das etapas da presente pesquisa, os procedimentos metodológicos foram divididos em duas etapas principais: 1) Condução da RSL - que buscou estabelecer um corpus de artigos, discussões teóricas sobre os desafios e processos relacionados com a avaliação de qualidade de dados *Linked Data*, apontando as principais problemáticas, dimensões e métricas relacionados ao processo, bem como analisar os artefatos disponíveis para realização do processo de avaliação, e 2) Segunda rodada de seleção e de coleta de dados – onde buscou-se Identificar os artigos que discutem ou aplicam técnicas de ML na avaliação de qualidade de dados *Linked Data* e coletar informações a respeito desses estudos. O Quadro 1 apresenta uma versão resumida das informações utilizadas para estruturar a primeira etapa da pesquisa.

Na segunda etapa foi realizada uma busca pelo termo "*Machine Learning*", com a finalidade de identificar quais abordam diretamente o uso de ML na avaliação de qualidade de dados *Linked Data*. O quadro 2 apresenta o protocolo de pesquisa resumido dessa rodada.

Quadro 1 – protocolo de pesquisa resumido (1° rodada)

Protocolo de busca	
Pergunta de pesquisa (principal)	Como tem sido abordada a questão da qualidade de dados no contexto do *Linked Data*?
Objetivos	Identificar os principais enfoques temáticos através dos quais se discute qualidade de dados publicados como *Linked Data*.
Estratégia de busca	("Linked Data" OR "Linked Open Data") AND ("Data Quality")
Bases de dados consultada	1º rodada Web of Science 2º rodada ISTA; LISTA; 3° rodada BRAPCI (busca simples pelo termo "Qualidade de dados")
Período abrangido	Sem restrição temporal.
Idiomas	Português, inglês e espanhol.
Critérios de Inclusão	(I) Foco principal é voltado para discutir qualidade de dados publicados de acordo com os princípios do *Linked Data*
Critérios de exclusão	(E) Não está nos idiomas estabelecidos para a pesquisa; (E) Apenas menciona a temática de interesse; (E) Não aborda a temática de interesse; (E) Não foi possível obter acesso ao documento completo;
Formulário de extração	Desafios de qualidade
Data da coleta	Entre dezembro de 2021 e maio de 2022

Fonte: Autores.

Quadro 2 – protocolo de pesquisa (2° rodada)

Protocolo de pesquisa	
Pergunta de pesquisa (principal)	Existem artefatos que utilizam técnicas de ML para avaliação de qualidade? Em caso positivo, em que contexto essas técnicas são aplicadas?
Objetivos	Verificar o potencial de uso das técnicas de ML na avaliação de qualidade de dados *Linked Data*
Estratégia de busca	Busca simples nos PDFs dos artigos aceitos, pelo termo "*Machine Learning*", leitura dos resumos e palavras-chave dos artigos para checagem.
Critérios de Inclusão	(I) Discute um artefato de avaliação de qualidade de dados *Linked Data* que utiliza técnicas de ML (I) realiza uma discussão teórica sobre a possiblidade de uso de técnicas de ML no processo de avaliação de qualidade de dados *Linked Data*
Critérios de exclusão	(E) Não discute ou aplica técnicas de ML na avaliação de qualidade de dados *Linked Data* (E) Apenas menciona o uso de ML no processo de avaliação de qualidade

Formulário de extração	Artefato discutido; Contexto de aplicação; vantagens do uso ML para avaliação de qualidade de dados *Linked Data*; Desafios do uso de ML para avaliação de qualidade de dados *Linked Data*
Data da coleta	Julho de 2023

Fonte: Autores.

A próxima seção apresenta os resultados dos procedimentos apresentados.

Resultados e discussões

Na primeira etapa da pesquisa, de condução da RSL, foram recuperados 225 documentos. Desses, 30 foram identificados como documentos duplicados, 100 documentos foram rejeitados e 95 artigos foram aceitos para compor o corpus teórico da pesquisa.

Na segunda rodada foram analisados os 95 artigos, dos quais 73 (70,3%) puderam ser excluídos com base no critério "(E) Não discute ou aplica técnicas de ML na avaliação de qualidade de dados *Linked Data*". 15 artigos (15,70%) foram excluídos com base no critério "(E) Apenas menciona o uso de ML no processo de avaliação de qualidade". Parte desses artigos possuem uma seção de trabalhos correlatos ou realizaram uma revisão de literatura buscando identificar artefatos que realizassem a avaliação de qualidade de dados *Linked Data*.

Ao final, foram identificados 6 (6,31%) artigos que abordam diretamente a aplicação de ML em artefatos de avaliação de qualidade voltados para dados *Linked Data*. O quadro 3 apresenta os documentos aceitos.

Quadro 3 – documentos aceitos na segunda etapa da pesquisa

Autores	Título	Artefato	Abordagem
Hadhiatma (2018)	Improving data quality in the linked open data: a survey	Não se aplica (discussão teórica)	Discute as estratégias e os métodos para avaliação de qualidade de dados *Linked Data*. Entre outros métodos aborda o uso de técnicas de ML.

Autores	Título	Artefato	Abordagem
Liu *et al.* (2017)	Exploiting Source-Object Networks to Resolve Object Conflicts in *Linked Data*	ObResolution (object resolution) –	Os problemas de conflito de objeto são frequentes em contexto *Linked Data*, especialmente em dados convertidos de fontes semiestruturadas e não estruturadas. O artigo propõe uma ferramenta automática de resolução de conflitos de objetos,
Mihindukulasooriya e Rico (2018)	Type Prediction of RDF Knowledge Graphs Using Binary Classifiers with Structural Data	Data-driven type prediction approach	Apresenta ferramenta cuja finalidade é prever o tipo de informação, eliminando assim inconsistências na tipificação dos objetos, realiza uma abordagem baseada em pré-treinamento e ML.
Esteves *et al.* (2017)	Toward Veracity Assessment in RDF Knowledge Bases: An Exploratory Analysis	Deep Fact Validation (DeFacto)	Apresenta ferramenta que atua como validador de triplas RDF, criado para verificar a veracidade de fatos. A ferramenta tem seu funcionamento baseado em treinamento prévio, onde o algoritmo recebe um conjunto de dados para cada categoria de triplos, auxiliando na identificação da veracidade das informações.
Rico *et al.* (2018)	Predicting incorrect mappings: a data-driven approach applied to DBpedia	Não nomeado	A conversão de dados para *Linked Data* depende do estabelecimento de "template mappings", mapeamentos que estruturam como ocorrerá a conversão de dados para RDF, seguindo o model de triplas Recurso+Propriedade+Valor. Mais de um mapeamento pode ser utilizado em um mesmo projeto de conversão, levando a inconsistências. A ferramenta faz uso de ML para identificar esses potenciais inconsistências entre diferentes mapeamentos.

Autores	Título	Artefato	Abordagem
Elbattah e Ryan (2019)	Learning Sequence Patterns in Knowledge Graph Triples to Predict Inconsistencies	Não nomeado	Apresenta ferramenta que busca identificar inconsistências em dados *Linked Data*. Para isso, leva-se em consideração a estrutura Recurso+Propriedade+Valor do RDF, partindo da hipótese de que quando objetos possuem um mesmo valor, mas propriedades distintas, existe uma chance maior de inconsistências. Na ferramenta as técnicas de ML são aplicadas para prever essas inconsistências com base em padrões de erros.

Fonte: Autores.

Como é possível observar no quadro 3, apenas Hadhiatma (2018) realiza uma discussão teórica sobre o uso de ML para a avaliação de qualidade de dados *Linked Data*, embora esse não seja o foco do artigo. O uso de ML é discutido em meio a outros métodos. Após traçarem a discussão teórica, o autor apontam como trabalhos futuros o desenvolvimento de estudos para elaboração de um framework que permita o uso de ML no processo de avaliação.

Hadhiatma (2018) aponta que entre os desafios do uso de ML para avaliação de qualidade de dados *Linked Data* está a complexidade de realização do treinamento de algoritmos com dados no formato RDF, sendo a forma mais comum o treinamento com dados em colunas individuais, não relacionados. O autor aponta que esse processo pode ser complexo e custoso, dependendo do uso de técnicas mistas, e que existem ainda poucos estudos que realizem essa aplicação.

Ao analisar as aplicações identificadas no recorte da pesquisa, é possível notar que o uso de ML não é feito para a realização de um ciclo completo de avaliação de qualidade, com a verificação de dimensões e métricas abrangentes e fornecimento de um relatório de qualidade, mas sim em etapas pontuais desse processo.

Destaca-se especialmente o uso na identificação de inconsistências resultantes do processo de conversão de dados para *Linked Data*, como na

identificação da tipificação incorreta de objetos, uso incorreto de propriedades ou ainda na checagem de fatos falsos.

Nos artigos que abordam a elaboração de artefatos utilizando Ml, observou-se que boa parte não realiza um aprofundamento sobre a experiência com o uso da técnica empregada, não apresenta uma comparação que permita identificar a vantagem com resultados obtidos através de outras técnicas, e nem aponta os desafios enfrentados durante o processo. Apenas Elbattah e Ryan (2019, p. 438, tradução nossa) fornecem um maior detalhamento sobre esses desafios, indicando que

> Uma grande limitação para o desenvolvimento de uma abordagem baseada em *Machine Learning* foi a indisponibilidade de exemplos errados (ou seja, inconsistentes). Os triplos incluídos foram presumivelmente considerados como corretos. Mesmo que exemplos inconsistentes possam provavelmente existir, eles não foram explicitamente rotulados.

Ou seja, nesse modelo de uso, o desafio é que mesmo que existam dados presumivelmente incorretos, esses não estão rotulados como tal. Como o problema a ser sanado é justamente a identificação de inconsistências que se misturam a dados consistentes, a ausência de uma base que agrupe dados incorretos para ser usada no treinamento, torna inviável a preparação do algoritmo. Seria necessário um trabalho manual prévio de identificar quais são esses dados com inconsistências, agrupá-los em uma base de dados que seria usada no treinamento e então permitiria a identificação de inconsistências semelhantes.

Considerações finais

A presente pesquisa buscou responder aos seguintes questionamentos: existem artefatos que utilizam técnicas de ML para avaliação de qua-

lidade de dados *Linked Data*? Em caso positivo, em que contexto essas técnicas são aplicadas?

Em resposta à primeira pergunta de pesquisa, observa-se que sim, existem artefatos que se utilizam das técnicas e ML para avaliar a qualidade de dados *Linked Data*, mas que esses artefatos ainda são poucos, se comparados aos demais artefatos de avaliação de qualidade encontrados na literatura.

Apenas 6% dos artigos aceitos discutiram ou aplicaram o uso de técnicas de *Machine Learning* nesse processo. Outros 15 % dos artigos mencionam brevemente o uso de ML nesse processo, mas não realizam uma discussão aprofundada a respeito desse uso.

Em relação à segunda pergunta de pesquisa, observa-se que nos estudos aceitos, as técnicas de ML não são aplicadas em todo o processo de avaliação de qualidade e sim em etapas pontuais. Destaca-se seu uso na garantia da qualidade dos processos de conversão de dados não estruturados e semiestruturados para *Linked Data*. Nesses casos, as técnicas de *Machine Learning* foram aplicadas para prever inconsistências, por meio da identificação de padrões de erros.

O uso de ML em um processo de avaliação de qualidade completo não aparece entre os estudos identificados, abordando quais os potenciais contribuições a cada etapa. Com base na ausência desse tipo de análise e discussão, um dos maiores desafios é a ausência de estudos que permitam compreender quais as dificuldades e vantagens de se aplicar técnicas de ML no processo estudado.

As técnicas de ML ainda são pouco exploradas, pelo menos de maneira direta na literatura, entretanto, observou-se que em alguns casos, mesmo não mencionando diretamente o uso de *Machine Leraning*, os artefatos podem fazer uso de algoritmos de ML, ainda que esse uso não seja discutido de maneira direta nos artigos. Outros termos podem ser consultados para tentar identificar o uso indireto de ML nos artefatos, como o termo Algoritmos, associados a conceitos como treinamento e pré-treinamento. Também se torna necessário verificar a documentação desses

artefatos, onde pode existir um detalhamento maior de seu funcionamento em que o uso de ML seja indicado.

Como etapas futuras da pesquisa, propõe-se a análise da documentação oficial desses artefatos, buscando compreender se esses fazem uso de algoritmos de ML em seu funcionamento.

Dados os desafios identificados na literatura, pretende-se realizar ainda um aprofundamento no processo de treinamento de algoritmos de ML usando dados em formato RDF, que permita identificar se esse processo é de fato um impedimento ao uso de ML para a realização de um processo completo de avaliação de qualidade de dados Linked Data.

Agradecimentos

Agradecemos à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pelo financiamento recebido para o desenvolvimento dessa pesquisa.

Referências

AHMED, H. H. Data Quality Assessment in the Integration Process of Linked Open Data (LOD). **2017 Ieee/Acs 14Th International Conference On Computer Systems And Applications (Aiccsa)**, [*s.l.*], v. 14, n. 1, p. 1-6, out. 2017. Disponível em: http://dx.doi.org/10.1109/aiccsa.2017.178. Acesso em: 13 jul. 2023.

ASSAF, A.; SENART, A.; TRONCY, R. Towards An Objective Assessment Framework for Linked Data Quality. **International Journal On Semantic Web And Information Systems**, Hershey, v. 12, n. 3, p. 111-133, jul. 2016. Disponível em: http://dx.doi.org/10.4018/ijswis.2016070104. Acesso em: 27 maio 2022.

BERNERS-LEE, T. **Linked data**, 2006. Disponível em: https://www.w3.org/DesignIssues/LinkedData.html. Acesso em: 14 mar. 2021.

CONEGLIAN, C. S. **Recuperação da Informação com abordagem semântica utilizando Linguagem Natural:** a Inteligência Artificial na Ciência da Informação. 2020. 194 f. Tese (Doutorado) — Curso de Programa de Pós-Graduação em Ciência da Informação, Universidade Estadual Paulista, Marília, 2020. Disponível em: https://repositorio.unesp.br/bitstream/handle/11449/193051/coneglian_cs_dr_mar.pdf?sequence=3&isAllowed=y. Acesso em: 08 set. 2022.

ELBATTAH, M.; RYAN, C. Learning Triple Sequence Patterns in Knowledge Graphs to Predict Inconsistencies. **Proceedings Of The 11Th International Joint Conference On Knowledge Discovery, Knowledge Engineering And Knowledge Management**, [*s.l.*], p. 435-441, 2019. Disponível em: http://dx.doi.org/10.5220/0008494604350441. Acesso em: 13 jul. 2023.

HADHIATMA, A. Improving data quality in the linked open data: a survey. **Journal Of Physics**: Conference Series, Bristol, v. 1, n. 978, p. 1-7, Mar. 2018. Disponível em: http://dx.doi.org/10.1088/1742-6596/978/1/012026. Acesso em: 13 jul. 2023.

JESUS, A. F.; SANTAREM SEGUNDO, J. E. A questão da qualidade em dados publicados como linked data: um mapeamento sistemático da literatura. *In:* ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO, 22., 2022, Porto Alegre. **Anais** […]. p. 1-16. Disponível em: https://enancib.ancib.org/index.php/enancib/xxiienancib/paper/view/811. Acesso em: 30 mar. 2023.

JESUS, A. F.; TRIQUES, M. L.; SANTARÉM SEGUNDO, J. E.; ALBUQUERQUE, A. C. Processamento de linguagem natural e machine learning na categorização de artigos científicos: um estudo em torno do patrimônio cultural. **Revista Ibero-Americana de Ciência da Informação**, Brasília, v. 1, p. 164-184, 2023.

JURAN, J. M. **Quality Control Handbook**. New York: 1988.

LIU, W. *et al*. Exploiting Source-Object Networks to Resolve Object Conflicts in Linked Data. **The Semantic Web**, Amsterdam, v. 1, n. 10249, p. 53-67, 2017. Disponível em: http://dx.doi.org/10.1007/978-3-319-58068-5_4. Acesso em: 13 jul. 2023.

MIHINDUKULASOORIYA, N.; RICO, M. Type Prediction of RDF Knowledge Graphs Using Binary Classifiers with Structural Data. **Current Trends In Web Engineering**, Bari, n. 1, p. 279-287, June 2018. Disponível em: http://dx.doi.org/10.1007/978-3-030-03056-8_27. Acesso em: 13 jul. 2023.

NELSON, R.R.; TODD, P.A.; WIXOM, B.H. Antecedents of information and system quality: an empirical examination within the context of data warehousing. **Journal of Management Information Systems**, Londres, v. 21, n. 4, p. 199–235, 2005.

NOOGHABI, M. Z.; DASTGERDI, A. F. Proposed metrics for data accessibility in the context of linked open data. **Program**, Bingley, v. 50, n. 2, p. 184-194, 4 abr. 2016. Disponível em: http://dx.doi.org/10.1108/prog-01-2015-0007. Acesso em: 26 maio 2022.

RICO, M. *et al*. Predicting incorrect mappings. **Proceedings Of The 33Rd Annual Acm Symposium On Applied Computing**, v. 33, n. 1, p. 323-330, 9 abr. 2018. Disponível em: http://dx.doi.org/10.1145/3167132.3167164. Acesso em: 13 jul. 2023.

WANG, R. Y.; STRONG, D. M. Beyond accuracy: what data quality means to data consumers. **J. Manage. Inf. Syst**., Londres, v. 12, n. 4, p. 5–33, Jan. 1996.

ZAVERI, A. *et al*. Quality Assessment Methodologies for Linked Open Data. **Semantic Web journal**, Thousand Oaks, v. 1 n. 5, p. 1-33, 2012. Disponível em: http://www.semantic-web-journal.net/system/files/swj414.pdf. Acesso em: 16 ago. 2022.