

Design de Sistemas de Reparação de Alinhamento de Ontologias para Suporte à Especialistas de Domínio

(English Title: Design of Ontology Alignment Repair Systems to Support Domain Experts)

Miriam Oliveira dos Santos^{1,2}, Tadeu Moreira de Classe¹, Carlos Eduardo de Mello¹

¹Programa de Pós-Graduação em Informática (PPGI)
Universidade Federal do Estado do Rio de Janeiro (UNIRIO)
Rio de Janeiro – RJ – Brazil

²Fundação Oswaldo Cruz (Fiocruz)
Rio de Janeiro – RJ – Brazil

{miriam.santos,tadeu.classe,mello}@uniriotec.br

Abstract. *The ontology alignment aims to interoperability among systems connecting by ontologies in the same domain. In general, these ontology alignments are not effective, requiring domain experts to fix them and, there is no adequate software support to help them. Therefore, this paper presents the MRED, a guide of requirements to help design systems for ontology alignment repair (RAO). The demonstration follows the MRED to implement a system, which was evaluated in a quasi-experimental study. As result we observed good indicators in coverage, accuracy, and f-measure, and, thus, there are indications that the MRED supports the design of RAO systems that help domain expecialists in the RAO task.*

Resumo. *O alinhamento de ontologias busca a interoperabilidade entre sistemas através de ontologias de um mesmo domínio. Na maioria dos casos, os alinhamentos não são efetivos, necessitando de especialista de domínio para corrigi-los, os quais, nem sempre possuem softwares adequadas a isso. Neste contexto, este trabalho apresenta o MRED, um guia de funcionalidades para auxiliar no design de sistemas de reparação de alinhamentos de ontologias (RAO). O MRED foi demonstrado através da implementação de um sistema e avaliado em um estudo quasi-experimental. Como resultado, obteve-se bons indicativos de cobertura, acurácia e medida-f, e, portanto, há indícios que o MRED fornece requisitos ao design de sistemas de RAO que euxiliam expecialistas de domínio na tarefa de RAO.*

1. Introdução

Diferentes sistemas usam ontologias para o gerenciamento, interoperabilidade e compartilhamento de conhecimento [Nguyen et al. 2019]. Geralmente, elas descrevem um domínio de interesse ou alguma parte de conhecimento relacionado com determinada área (ou domínio). Elas especificam significados de termos que englobam o domínio, estabelecendo um relacionamento entre eles, tendo como características principais sua estruturação e organização hierárquica e lógica entre seus conceitos [Euzenat and Rousset 2020]. No geral, elas são aplicadas na integração, busca e análise de dados de diferentes fontes de informação. Na maioria das vezes existem diferentes ontologias para um domínio em particular, sendo cada uma delas modelada de formas distintas [Harrow et al. 2019].



Buscando melhorar a interoperabilidade e diminuir a distância entre os conceitos de diferentes ontologias dentro de um domínio em particular, o processo de alinhamento de ontologias é realizado [Gargouri and Jaziri 2010]. Tal processo consiste em destacar relacionamentos e/ou correspondências entre as entidades pertencentes as ontologias, direcionados a diminuir suas heterogeneidades semânticas [da Silva et al. 2016]. Como resultado, é gerado um mapeamento de ontologias, compreendendo a um conjunto de correspondências entre os conceitos de duas ou mais delas [Euzenat 2014].

Na literatura há indicações de que métodos de alinhamento totalmente automáticos não são totalmente efetivos [Li et al. 2019]. Mesmo usando estratégias automáticas, como comparação textual, linguística, alinhamento estrutural, análise semântica, por exemplo, os conceitos entre as ontologias podem não ser relacionados de maneira correta [Cheatham and Hitzler 2013, Duan et al. 2010].

Assim, é necessária a intervenção manual no processo de alinhamento de ontologias pelo especialista do domínio. A tarefa do *expert* é reparar o alinhamento criado por meios automáticos, confirmando e removendo correspondências existentes, ou, incluindo novas associações entre os conceitos. Essa tarefa tem o nome de **Reparação de Alinhamento de Ontologias (RAO)**, a qual objetiva garantir uma melhor precisão e cobertura entre os mapeamentos gerados [Meilicke 2011, Faria et al. 2013]. Existem pesquisas que investigam estratégias semiautomatizadas para a RAO, mas elas não fornecem o suporte adequado ao especialista de domínio, basicamente, limitando a exibir os mapeamentos gerados por meio de interfaces complexas [Faria et al. 2013, Massmann et al. 2011]. Neste contexto, o principal desafio da pesquisa está em: *como desenvolver sistemas que ajudem os especialistas de domínio no processo de RAO, contribuindo com a cobertura e precisão entre os alinhamentos.*

Esta pesquisa foi conduzida usando a metodologia *Design Science Research Methodology (DSRM)* [Peppers et al. 2007], baseada na *Design Science Research (DSR)*. Tal metodologia aborda a criação de um artefato tecnológico baseado em conjecturas e requisitos de design para a resolução de um problema, avançando conhecimentos tanto no campo técnico, quanto em científicos. Este trabalho apresenta o artefato principal da pesquisa, um guia com requisitos para o design de sistemas que forneçam suporte à reparação de alinhamento de ontologias para especialistas de domínio, chamado **MRED** e uma demonstração de design a partir do sistema *RAOSystem*. A avaliação do método foi realizada através de um *quasi-experimento*, no qual foram comparadas métricas de cobertura e precisão, obtidas a partir do uso do sistema em relação a alinhamentos de referências.

Desta forma, este artigo traz as seções seguintes organizadas em: seção 2 apresenta a pesquisa segundo a *DSRM*. A Seção 3 apresenta um referencial sobre alinhamento de ontologias e suas métricas para avaliação. A Seção 4, discute alguns trabalhos relacionados e como eles contribuíram com a propostas da pesquisa. O artefato da pesquisa, o *MRED* é apresentado e demonstrado na seção 5. O planejamento da avaliação da pesquisa é detalhado na seção 6. Na seção 7 são apresentados e discutidos os resultados da avaliação. E, a seção 8 apresenta as conclusões e *insights* futuros.

2. Design da Pesquisa

A DSR é voltada à ciência do artefato tendo como missão a realização de pesquisas baseadas na criação e desenvolvimento de conhecimentos válidos, tanto no campo científico, quanto no tecnológico. O conhecimento produzido pode ser usado na academia e por

profissionais para a criação de soluções a problemas reais, baseando-se em conjecturas teóricas ou técnicas [Vom Brocke et al. 2020]. Especificamente voltados para sistemas de informação, Peffers et al. criaram o método *DSRM* para a condução de pesquisas baseadas em DSR, nessa área. Devido a isso este método foi escolhido para conduzir as etapas nessa pesquisa [Peffers et al. 2007].

Baseando-se nesta metodologia, a pesquisa teve como contexto o processo de reparação de alinhamento de ontologias, objetivando apresentar uma solução satisfatória para o problema de *auxiliar os especialistas de domínio no processo de reparação de alinhamento de ontologias, desenvolvendo sistemas de RAO capazes de melhorarem os índices de cobertura e precisão*. Como enfoque de solução, foi proposto como artefato o **MRED**, sendo ele demonstrado pela implementação do sistema **RAOSystem**. As bases teóricas se concentraram em conceitos e definições da *web semântica, alinhamento e reparação de alinhamento* de ontologias. As bases técnicas foram concebidas a partir da identificação de características e funcionalidades de sistemas de RAO usados profissionalmente e na literatura existente sobre isso. O artefato busca analisar a conjectura teórica de que os especialistas de domínio, ao usar sistemas baseados nos requisitos do MRED, melhoram os índices de cobertura e precisão no processo de RAO. Além da verificação sobre a conjectura técnica de que MRED fornece suporte ao design de sistemas que contribuem na tarefa de RAO. O artefato e as conjecturas foram avaliadas através de um estudo *quasi-experimental*.

A DSR pode ser concebida por ciclos interativos de investigação, que geram conhecimentos tecnológicos e teóricos obtidos do design e da avaliação do artefato, originando, também, novos *insights*. Os *insights* fornecem observações para novos ciclos de investigação. Neste artigo, é apresentado o primeiro ciclo de investigação, compreendendo a versão preliminar do MRED, sua demonstração e avaliação.

3. Alinhamento e Reparação de Ontologias

Para promover a interoperabilidade entre os sistemas que as utilizam uma ou mais ontologias de domínio é necessário realizar um alinhamento entre os seus conceitos [da Silva et al. 2016]. Assim, alinhamentos de ontologias surgem como processos capazes de auxiliar no processo de mapeamento semântico entre duas ou mais ontologias, de forma a fazer a associação entre os conceitos do domínio [Gargouri and Jaziri 2010]. O mapeamento de ontologias é gerado como resultado desse processo, compreendendo um conjunto de correspondências que incluem a representação do relacionamento, equivalência, disjunção ou associação e, um valor que representa a força de similaridade entre os termos [Banouar and Raghay 2016].

Existem várias técnicas que podem ser aplicadas para a geração de alinhamento de ontologias, as quais se baseiam em comparação textual (*string*), linguística, estrutural e semântica [Banouar and Raghay 2016]. Ainda que haja diferentes técnicas automáticas para o alinhamento de ontologias, elas não são completamente efetivas [Li et al. 2019]. Devido à complexidade da tarefa de alinhamento, a intervenção humana do especialista de domínio é necessária para verificar e ajustar inconsistências no mapeamento, levando a resultados mais precisos [Li et al. 2015].

Levando isso em consideração o processo subsequente ao alinhamento é a **Reparação de Alinhamento de Ontologias (RAO)** [Pesquita et al. 2013]. Isto é, o refinamento do alinhamento, através da análise do especialista de domínio, o qual pode incluir, remover ou confirmar correspondências entre termos [Banouar and Raghay 2016].

A tarefa também permite a identificação de inconsistências lógicas, tornando o alinhamento final mais consistente [Pesquita et al. 2013]. A implementação de ações de RAO pode ser realizada usando estratégias automatizadas, semiautomatizadas ou manuais [Meilicke 2011].

Existem pesquisas que focam em algoritmos para automatizar a tarefa de RAO, desconsiderando o suporte cognitivo humano na análise e inclusão de novos mapeamentos [Falconer et al. 2007]. Em contrapartida, a atividade de reparação feita pelo especialista de domínio pode ser a chave para corrigir e confirmar alinhamentos de ontologias já gerados [Granitzer et al. 2010].

Nesta pesquisa considera-se a importância da intervenção do especialista de domínio no processo de RAO. Devido a isso, o seu escopo se concentra na criação de um conjunto de requisitos e funcionalidades que guie a criação de sistemas para auxiliar os especialistas nesta tarefa.

3.1. Métricas de Avaliação em Sistemas de Alinhamento de Ontologias

Existem métricas para avaliar o alinhamento gerado (G) (ou mapeamento) por um sistema de alinhamento de ontologias a partir da comparação com um alinhamento de referência (R). Um alinhamento de referência pode ser um mapeamento gerado por um sistema automatizado ou manualmente por um especialista de domínio, sendo consideradas, como próprio nome diz, a referência correta entre o alinhamento das ontologias [Euzenat et al. 2007, Pesquita et al. 2013].

As principais métricas de avaliação são: **precisão**, **cobertura** e **medida-f** (Figura 1A, tendo como base a classificação de uma matriz de confusão [Euzenat et al. 2007]. A matriz de confusão (Figura 1B) permite classificar os termos alinhados do mapeamento em: **Verdadeiro Positivo (VP)**: termos apontados no alinhamento gerado e de referência; **Verdadeiro Negativo (VN)**: termos não apontados no alinhamento gerado nem no de referência; **Falso Positivo (FP)**: termos apontados no alinhamento gerado mas não existe no de referência e; **Falso Negativo (FN)**: termos não apontado no alinhamento gerado mas existe no de referência. Assim, é possível averiguar os alinhamentos gerados a partir do cálculo das métricas descritas acima.

$$(A) \quad Precisão(G, R) = \frac{VP}{VP + FP} \quad Cobertura(G, R) = \frac{VP}{VP + FN}$$

$$MedidaF(G, R) = \frac{2(Precisão * Cobertura)}{Precisão + Cobertura}$$

		Alinhamento Gerado (G)	
		+	-
Alinhamento de Referência (R)	+	VP	FP
	-	FN	VN

Figura 1. A) Métricas de avaliação de alinhamentos. B) Matriz de confusão para análise de alinhamentos.

4. Trabalhos Relacionados

Existem iniciativas na literatura que, embora consideradas como referência no uso de estratégias automáticas ou semiautomáticas para a tarefa de RAO, não dão o suporte ne-

cessário para os especialistas de domínio realizarem suas contribuições. Essas ferramentas são limitadas a exibir apenas os mapeamentos que elas mesmas conseguiram elencar. Dentre as principais delas é possível destacar: *COMA (Combination of Schema Matching Approach)*, fornecendo uma plataforma para analisar mapeamentos, representando-os em árvores de conceitos [Massmann et al. 2011]. A *SMOV (Automated Semantic Matching of Ontologies With Semantic Validation)*, apresentando uma página web contendo os termos da ontologia e seus respectivos relacionamentos [Meilicke 2011]. E, a *AML (Agreement Maker Light)*, que segundo Faria et al. [Faria et al. 2013], apresenta grafos como forma de representação das ontologias.

Em relação a esses sistemas, uma característica sempre presente é o uso de visualização de informação, como os grafos, *trees*, relacionamento par-a-par, permitindo que os alinhamentos sejam observados e analisados pelos especialistas. Contudo, novos requisitos colaborativos, cognitivos e de interação, podem ser utilizados na construção destes sistemas, permitindo uma melhora no processo de RAO pelos especialistas do domínio, como apontado em vários trabalhos [Li et al. 2015, Granitzer et al. 2010, Falconer et al. 2007].

Neste trabalho, como é proposto um guia de requisitos para construção de sistemas para auxiliar os especialistas de domínio na tarefa de RAO, as funcionalidades e requisitos descritas em trabalhos de referência inspiraram os elementos do MRED. Diferentemente dos sistemas aqui mencionados, o MRED permite atividades de reparação manual pelos especialistas de domínio, não apenas a partir da visualização de informação, mas por meio de funcionalidades colaborativas, métricas de análise de trabalho, histórico de mudanças e outras. Esses requisitos do MRED são mais bem descritos na seção 5.

5. “MRED” - Guia de Requisitos e Funcionalidades para Sistemas de Reparação de Alinhamento de Ontologias

O MRED foi concebido, após a análise da literatura relacionada, no qual foram encontrados dezoito requisitos comuns em sistemas de reparação e alinhamento de ontologias. Assim, estes 18 requisitos foram considerados sugestões importantes para o design e desenvolvimento de sistemas para reparação de alinhamento de ontologia usados por especialistas de domínio. Ou seja, um sistema que apoie a tarefa de reparação de alinhamento de ontologias deve implementar esses requisitos. Estes requisitos se associam a cinco dimensões do MRED (Figura 2), descritos a seguir.

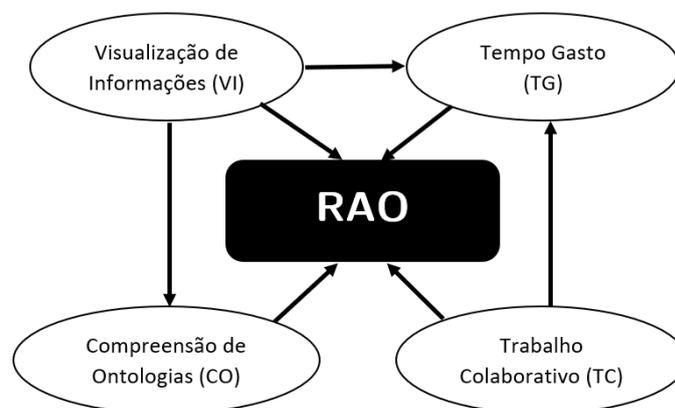


Figura 2. Dimensões do MRED

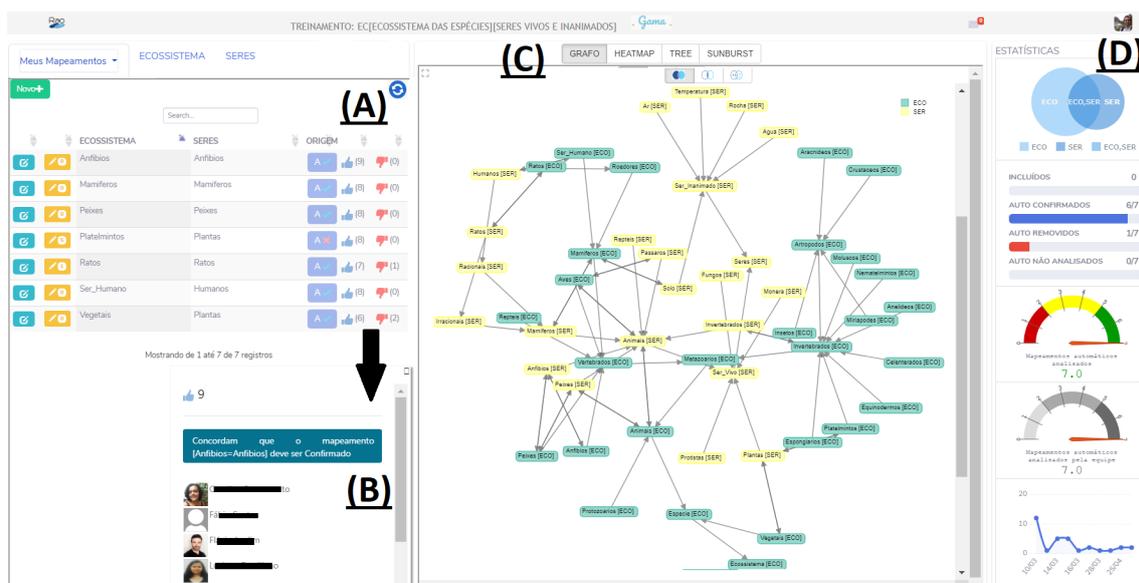


Figura 3. Interface principal do RAOSystem

- **Compreensão de Ontologias (CO):** requisitos que levam ao usuário entender os elementos e termos das ontologias observadas. Este é o primeiro passo para que o especialista de domínio consiga validar um alinhamento é ser familiarizado com a ontologia [Dragisic et al. 2016].
- **Trabalho Colaborativo (TC):** recursos direcionados a promover a colaboração e construção de conhecimento entre os especialistas de domínio. A interatividade de equipe permite ao time de especialistas analisarem, revisarem e discutirem o trabalho de reparação de alinhamento [Falconer et al. 2007];
- **Visualização de Informação (VI):** informações fornecidas por meio de recursos visuais ao especialista de domínio, contribuindo com o conhecimento das ontologias, seus relacionamentos e análises de trabalho. Esses recursos permitem a percepção simples e rápida do trabalho de RAO [Fu et al. 2013];
- **Tempo Gasto (TG):** tempo gasto está relacionado a recursos de promoção e continuidade do trabalho, incidindo sobre decisão dos especialistas e a qualidade da reparação [Falconer et al. 2007]
- **Reparação de Alinhamento de Ontologias (RAO):** requisitos específicos do sistema como persistência de informação, acesso de usuários, suporte a tomada de decisão etc., dando suporte a todos as demais dimensões.

5.1. Demonstração: O RAOSystem

O RAOSystem é um sistema que atende aos requisitos definidos pelo MRED (Figura 3A); visualização de dados para facilitar a análise, revisão, inclusão e exclusão de alinhamentos (Figura 3C); elementos colaborativos como troca de mensagens, comentários, aceitações e rejeições de termos e, a comunicação entre os especialistas de domínio, além de permitir-lhes trocar conhecimento e validar o trabalho uns dos outros (Figura 3B) e ; gráficos e estatísticas de trabalho que permitem os visualizar o progresso do trabalho e controlarem o tempo gasto (Figura 3D).

O sistema foi todo desenvolvido para Web. Seu núcleo construído com a linguagem PHP apoiado por HTML, CSS, JavaScript, sendo o front-end foi elaborado usando o framework Bootstrap. Para o banco de dados foi usado o MySQL 8.0.16. Todas essas tecnologias rodam em conjunto em um servidor Apache. O sistema foi desenvolvido

em uma arquitetura modular, onde todas os requisitos e funcionalidades propostos pelo MRED foram incorporados, como apresentados a seguir:

- **Módulo de Revisão e Reparo:** contém a interface principal, englobando a maior quantidade de funcionalidades do MRED, além de exibir a interface principal do programa e o acesso a todas as demais opções. Neste módulo os especialistas de domínio acessam os alinhamentos de ontologias disponíveis e que estão armazenados no módulo de armazenamento. Assim, os usuários podem visualizar os alinhamentos já mapeados (Figura 3A). A partir disso, eles podem executar ações como: remover, inserir, confirmar, comentar, discordar e concordar com os alinhamentos existentes (Figura 3B). Este módulo também fornece ao usuário acesso às visualizações de informações sobre ontologias (módulo de visualização), estatísticas relacionados ao processo de reparação (módulo estatístico) e, histórico de mensagens e comunicação entre os usuários (módulo de comunicação).
- **Módulo de Armazenamento:** é o módulo responsável pelo armazenamento dos alinhamentos de ontologias. Ainda, é o responsável pela consulta e busca de alinhamentos, pelo armazenamento das reparações, guardar o progresso das tarefas e comentários, além de dar suporte ao módulo estatístico.
- **Módulo de Visualização:** neste módulo as funcionalidades de visualização dos alinhamentos são exibidas, dando suporte à tomada de decisão pelos especialistas de domínio. As informações são exibidas em formatos de grafos (Figura 3D) e mapas de calor especificamente voltadas à percepção de relacionamentos entre os termos, enquanto as árvores e gráficos sunburst permitem a análise de conceitos através da comparação de estrutura ontológica. Cada visualização torna possível selecionar opções de união e intersecção entre os termos do alinhamento, tal como visualizar todos os conceitos, conceitos mapeados e, conceitos já visitados. Apesar disso, esse módulo permite a completa visualização das ontologias a partir do histórico de alterações dos usuários.
- **Módulo Estatístico:** o propósito deste módulo é deixar que os usuários observem o feedback do trabalho realizado (ou não), como também quanto da reparação de alinhamento ainda precisa ser feito, a revisão do mapeamento automático, a quantidade de elementos a ser incluída, removida e confirmada e a execução do trabalho realizado diariamente (Figura 3D).
- **Módulo de Comunicação:** como o próprio nome diz, é o módulo responsável pela comunicação entre os participantes. Foi desenvolvido para a colaboração, discussão e troca de conhecimento entre os especialistas de domínio sobre o alinhamento. Todo alinhamento no sistema apresenta seu respectivo sistema de comunica via chat.

Assim, o MRED apresenta uma série de requisitos e funcionalidades que, ao serem usadas, podem auxiliar os especialistas de domínio na tarefa de RAO.

6. Avaliação do Artefato da Pesquisa

Nesta pesquisa o método MRED foi avaliado por meio de um estudo *quasi-experimental*. Segundo Campbell e Stanley [Campbell and Stanley 2015], os *quasi-experimentos* pertencem a uma classe de estudos de natureza empírica, menos controlados que os experimentos tradicionais, porém, sem a necessidade da realização de uma seleção aleatória de grupos de participantes. Assim como os experimentos tradicionais, estudos *quasi-experimentais* seguem etapas metodológicas de estudo [Shadish et al. 2002], como: 1) definição; 2) planejamento; 3) execução; 4) análise e interpretação e; 5) conclusões.

6.1. Definição do Estudo

A **definição** apresenta os objetivos do estudo sendo descritos pela abordagem *GQM* (*Goal-Question-Metric*) [Basili 1992], como: **analisar** o MRED; **com o propósito de** avaliação; **no que diz respeito a** melhoria dos indicadores de cobertura, precisão e *medida-f*; sob a **perspectiva de** especialistas de domínio; **no contexto da** tarefa de reparação de alinhamento de ontologias.

Baseado nesta definição de estudo e, como este é um estudo *quasi-experimental*, é possível formular hipóteses nulas (H_0) e alternativas (H_1) para os indicadores de cobertura, precisão e *medida-f*:

- H_0 : O sistema desenvolvido com base do MRED **não melhora a cobertura, precisão e medida-f** na tarefa de RAO.
- H_1 : O sistema desenvolvido com base do MRED **melhora a cobertura, precisão e medida-f** na tarefa de RAO.

6.2. Planejamento e Execução do Estudo

Os **participantes** da pesquisa foram selecionados por conveniência, incluindo estudantes, professores e pesquisadores de programas de pós-graduação da área de computação. Por serem integrantes da academia, este público se alinha bem à ontologias de domínio usadas neste estudo: (*Conference*¹ e *Ekaw*² (*International Conference on Knowledge Engineering and Knowledge Management*)), ontologias do domínio de conferências, simpósios e eventos acadêmicos.

O estudo foi **projetado** baseando-se na comparação entre o alinhamento de referência³ entre essas duas ontologias e as reparações de alinhamento realizados pelos participantes (especialistas de domínio) utilizando o sistema *RAOSystem* [Santos et al. 2020]. Cada participante executou suas tarefas de maneira individual.

Além disso, anualmente acontece um desafio entre sistemas de alinhamento de ontologias proposto pela *Ontology Alignment Evaluation Initiative* (OAEI). No ano de 2019, os sistemas *SANOM*, *AML* e *LOGMAP*, se destacaram no alinhamento das ontologias *Conference* e *Ekaw*. Assim, o resultado obtido destes sistemas no desafio, serão usados como critério de comparação do alinhamento gerado pelos participantes neste estudo.

Portanto, o estudo teve as seguintes etapas: i) **instrução**: disponibilizando aos participantes vídeos tutoriais com conceitos teóricos e ajuda sobre o sistema; ii) **treinamento**: para se habituarem ao sistema foi realizada uma etapa de treinamento de reparação de alinhamento de ontologias simples e; iii) **execução**: reparação de alinhamento de ontologias de conferência.

Como **instrumentalização** no *RAOSystem*, foram apresentadas 25 correspondências para reparação, sendo que 16 delas estavam corretas de modo que os participantes deveriam apenas confirma-las, 9 delas estavam incorretas, devendo ser removidas e, 8 correspondências não foram exibidas, devendo serem incluídas. Ao final do trabalho, os resultados foram comparados ao alinhamento de referência para o cálculo dos indicadores de cobertura, precisão e *medida-f* e a sua comparação aos indicadores obtidos dos sistemas do desafio da OAEI.

¹<http://www.scholarlydata.org/ontology/doc/>

²<https://ekaw.vse.cz/>

³Alinhamento de referência: https://bit.ly/MRED_DataEval

A **execução** do estudo contou com 10 participantes voluntários, de forma online. A coleta dos dados aconteceu de forma individual, sendo adquirida diretamente do sistema de reparação. Todos os dados foram analisados e tratados através do *R Statistics 3.2.2* e suas bibliotecas, sendo sumarizados no *Microsoft Excel*.

6.3. Ameaças de Validade ao Estudo

A principal ameaça associada às **conclusões** do estudo está associada ao poder estatísticos dos métodos usados, devido à existência de diferentes modelos estatísticos e as várias maneiras de empregá-los. Para diminuir a ameaça, testes estatísticos de normalidade e inferência foram usados de acordo com o comportamento observado nos dados a partir de suas análises. Ainda como ameaça de conclusão, o MRED foi avaliado em um ambiente controlado, o que significa que não necessariamente os participantes são especialistas de domínio, implicado em uma ameaça de generalização. No domínio de ontologia de conferências, eles foram considerados especialistas uma vez que estão envolvidos na área acadêmica e possuem conhecimento sobre o jargão empregado em simpósios, conferências etc. Contudo, é necessário testar um sistema criado a partir do MRED em um contexto real, como o de bioinformática, por exemplo, motivacional deste trabalho.

As demais ameaças podem ser descritas como, **validade interna**: i) Desgaste do participante: Possibilidade de pausar e continuar a tarefa a qualquer momento; ii) Falta de treinamento: Gravação e disponibilização de vídeos e realização de seção de treinamento e iii) Troca de informações entre participantes: O sistema possui funcionalidades colaborativas, portanto a troca de informações é incentivada. **Ameaças de construção**: i) Participantes sem conhecimento do domínio de aplicação: Participantes que não tivesse conhecimento do domínio das ontologias ou não fizessem parte do ambiente acadêmico foram eliminados; ii) Conhecimento de idioma: Participantes sem conhecimento de inglês foram eliminados; iii) Sem conhecimento sobre ontologias: Aplicação de treinamento básico e vídeos de treinamento e; iv) Expectativa do pesquisador: Toda a explicação e treinamento dos participantes foi feita através de vídeo, não havendo o contato com o pesquisador. E as **Ameaças externas**: i) Restrição de tempo: Uso de ontologias menores que se adequem a um tempo razoável de experimento, sendo este limitado em até 2 horas e; ii) Generalização: Trabalho com uma ontologia de domínio para que os participantes possam ser considerados especialistas daquele domínio.

7. Resultados e Discussões

7.1. Apresentação dos Resultados

Após a execução do estudo *quasi-experimental*, todas as respostas dos participantes foram coletadas, analisadas e sumarizadas para o cálculo dos indicadores de precisão, cobertura e *medida-f*. A Tabela 1 exhibe os dados relativos à reparação de alinhamento de todos os participantes. Dos 16 termos corretos, os participantes obtiveram uma média de 14,2 em relação aos verdadeiros positivos (VP), ou seja, eles conseguiram confirmar 89% dos termos do alinhamento de referência. Destaque aos participantes P1, P4 e P10, que confirmaram todos eles. Em média foram apontados 5,7 termos que não existiam no alinhamento de referência (falsos positivos – FP) e, em relação aos falsos negativos (FN), 3,4 (média) termos existentes no mapeamento de referência não foram apontados na reparação.

Analisando os participantes em relação as métricas de precisão, cobertura e *medida-f*, de maneira geral, o participante 4 (P4) obteve os melhores índices na realização da tarefa de RAO em comparação ao alinhamento de referência, obtendo uma precisão de 73% dos termos, uma cobertura de 67% dos termos e uma *medida-f* em 70%.

Tabela 1. Resultados dos Participantes

	VP	VN	FP	FN	Precisão	Cobertura	Medida-F
P1	16,00	8,00	8,00	2,00	0,67	0,67	0,67
P2	13,00	3,00	11,00	6,00	0,81	0,54	0,65
P3	14,00	6,00	10,00	3,00	0,70	0,58	0,64
P4	16,00	6,00	8,00	3,00	0,73	0,67	0,70
P5	15,00	5,00	9,00	4,00	0,75	0,63	0,68
P6	13,00	6,00	11,00	3,00	0,68	0,54	0,60
P7	15,00	7,00	9,00	2,00	0,68	0,63	0,65
P8	10,00	5,00	14,00	4,00	0,67	0,42	0,51
P9	14,00	4,00	10,00	5,00	0,78	0,58	0,67
P10	16,00	7,00	8,00	2,00	0,70	0,67	0,68
Méd.	14,20	5,70	9,80	3,40	0,72	0,59	0,65
DP.	1,78	1,42	1,78	1,28	0,05	0,07	0,05

A partir do cálculo das médias dos indicadores obtidos pelos participantes, é possível compará-los com os 3 sistemas do desafio da OAEI (Tabela 2). A Figura 4 ilustra a comparação entre esses sistemas de informação, no qual é possível perceber que, em alguns momentos, o *RAOSystem* foi superior a algum dos sistemas do desafio. Por exemplos, em quase todos os critérios, o sistema desenvolvido com bases no MRED, teve melhores valores do que o sistema *LOGMAP*. Em relação a precisão, o *RAOSystem* obteve valores melhores ou iguais aos sistemas *LOGMAP* e *SANOM*. E, em relação ao critério de falso negativo (FN) o *RAOSystem* obteve melhores resultados que todos os outros comparados, deixando de apontar apenas 3 termos que existem no alinhamento de referência.

Tabela 2. Comparativo dos sistemas do desafio da OAEI com o baseado no método MRED

	SANOM	AML	LOGMAP	RAOSystem
VP	18,00	18,00	12,00	14,50
FP	7,00	5,00	5,00	9,50
FN	6,00	6,00	12,00	3,00
Precisão	0,72	0,78	0,71	0,72
Cobertura	0,75	0,75	0,50	0,59
Medida-F	0,73	0,77	0,59	0,65

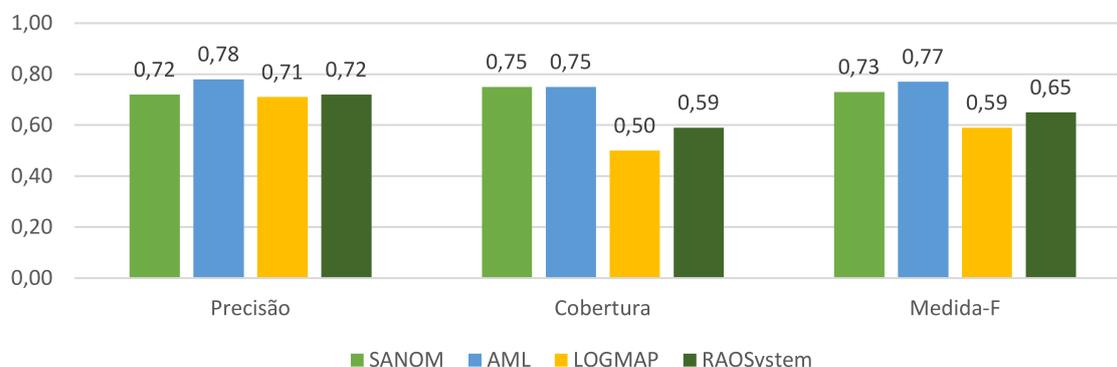


Figura 4. Comparativo de indicadores dos sistemas

Se comparados esses sistemas com apenas um dos participantes, por exemplo, o

participante 4 que obteve os melhores resultados, é possível perceber que os valores se igualam ou superam alguns dos sistemas do desafio: precisão = 0,73 (melhor que *SANOM* e *LOGMAP*), cobertura = 0,67 (melhor que *LOGMAP*) e *medida-f* = 0,70 (melhor que *LOGMAP*).

A partir dessas análises, percebe-se que o sistema desenvolvido com base no MRED, indicou valores um pouco melhores para cobertura, precisão e *medida-f* se comparado ao sistema *LOGMAP*. Este resultado, aparentemente valida a hipótese alternativa (H_1), especificamente comparando o *RAOSystem* com o *LOGMAP*. Para comprovar este resultado testes estatísticos de inferências foram usados.

Os resultados dos testes estatísticos podem ser observados na Tabela 3. Para as análises foi usado um intervalo de confiança de 95% ($\alpha=0,05$). Os dados dos participantes (Tabela 1) foram submetidos ao teste de *Shapiro-Wilk* (mais apropriado para a quantidade de participantes deste estudo), para averiguar o comportamento de normalidade dos dados, indicado qual seria o melhor teste de hipótese a seguir. Como resultado do teste de normalidade, os indicadores de confiança ($p\text{-value}=0,1199$) e precisão ($p\text{-value}=0,1151$) apresentaram comportamento **normal** ($\alpha>0,05$), enquanto a *medida-f* ($p\text{-value}=0,0147$) **não apresentou tal comportamento**. Devido a isso, para validar a hipótese, o *Teste T* foi aplicado aos indicadores de precisão e cobertura e o teste de *Wilcoxon* foi aplicado à *medida-f*. Além disso, o teste de normalidade ajudou a definir o tipo de teste de tamanho de efeito a ser usado, sendo o *Cohen'D* para precisão e cobertura (relativo ao Test T) e, o teste de *Vargha & Delaney's (A₁₂)* (relativo ao teste de Wilcoxon) para *medida-f*.

Tabela 3. Testes de normalidade, hipótese e tamanho de efeito

Indicador	Shapiro-Wilk		Hipótese		Tamanho Efeito		
	<i>p-value</i>	Nomal	Teste	<i>p-value</i>	Teste	<i>p-value</i>	Interpretação
Cobertura	0,1199	SIM	Teste T	5,15E-12	Cohen'D	6,2628	Médio
Precisao	0,1151	SIM	Teste T	2,03E-09	Cohen'D	2,0900	Pequeno
MedidaF	0,0147	NÃO	Wilcoxon	5,83E-03	A12	0,9000	Grande

Portanto, ao observar os dados da Tabela 3 e a Figura 4, é possível dizer com pelo menos 95% de confiança que os resultados indicam que o sistema desenvolvido a partir do método MRED, conseguiu ter uma melhor cobertura, precisão e *medida-f* se comparado aos resultados do sistema *LOGMAP*, usando o alinhamento de referências das ontologias *Conference* e *Ekaw*. Não sendo possível afirmar o mesmo para os sistemas *SANOM* e *AML*.

7.2. Limitações e Insights

Embora o alinhamento de ontologias e a reparação de alinhamento de ontologias sejam tarefas distintas, o seu resultado final é um mapeamento entre termos de uma ou mais ontologias. Assim, entende-se que comparar os resultados dos sistemas de alinhamentos automáticos do desafio anual da OAEI, com os resultados da tarefa de RAO, contribuíram para verificar e validar as métricas de cobertura, precisão e *medida-f* e, com isso, a relevância do MRED para a guiar o design de sistemas para a tarefa de RAO. Porém, o sistema gerado a partir do MRED, obteve resultados melhores que apenas um dos sistemas comparados. Desta forma não é possível generalizar que o MRED é capaz de auxiliar no desenvolvimento de sistemas que ajudem a melhorar a cobertura, precisão e *medida-f* para todos os casos.

Um dos maiores problemas enfrentados no uso do *RAOSystem* relatado pelos participantes é a dificuldade de encontrar novas correspondências e, também, destacar as incorretas. Isso impacta diretamente nas métricas. Portanto, é preciso investir novas recomendações e funcionalidades para o MRED, que estimule e facilite para os especialistas de domínio a identificação de inconsistências de novas correspondências.

Todas essas limitações, fornecem os *insights* necessários para a evolução do MRED, melhorando as suas recomendações e reflexões e, originando novos sistemas para auxiliar no trabalho de RAO.

8. Conclusões e Trabalhos Futuros

O objetivo principal deste artigo foi apresentar e avaliar o MRED, feito para guiar o design de sistemas para auxiliar especialistas de domínio na tarefa de RAO. O MRED foi demonstrado a partir da construção do *RAOSystem*, sistema esse usado para demonstrar a viabilidade de uso e na avaliação com participantes do estudo *quasi-experimental*. A conjectura teórica que se buscava satisfazer era que os *sistemas de RAO desenvolvidos com base no MRED eram capazes de melhorar os índices de cobertura, precisão e medida-f*, dando maior qualidade aos alinhamentos de ontologias. Essa conjectura pode ser **parcialmente satisfeita** a partir da análise dos dados colhidos no estudo, os quais indicaram que, ao comparar o SI baseado no MRED, com sistemas automáticos de alinhamento de ontologias, em pelo menos um deles (*LOGMAP*), os indicadores foram melhores.

A partir da metodologia *DSRM*, este trabalho apresentou o primeiro ciclo de investigação da pesquisa, com uma versão preliminar do MRED. A partir dos resultados deste estudo foi possível demonstrar a viabilidade do método e observar o uso do sistema de RAO pelos participantes, suas principais dificuldades e demandas. Desta forma, a partir dos *insights* deste ciclo, em um próximo, uma nova versão do MRED será explorada e, conseqüentemente, uma nova versão do SI para a tarefa RAO. Como os participantes do estudo apontaram, encontrar novos relacionamentos foi uma das principais dificuldades, assim, serão estudadas maneiras de estimular e facilitar este trabalho. Além disso, é necessário realizar novas avaliações, inclusive, usando ontologias maiores em contextos reais para que os resultados da pesquisa possam ser generalizados e aplicados na prática.

Referências

- Banouar, O. and Raghay, S. (2016). Interoperability of information systems through ontologies: State of art. *International Journal of Computer Science and Information Security*, 14(8):392.
- Basili, V. R. (1992). Software modeling and measurement: the goal/question/metric paradigm. Technical report.
- Campbell, D. T. and Stanley, J. C. (2015). *Experimental and quasi-experimental designs for research*. Ravenio Books.
- Cheatham, M. and Hitzler, P. (2013). String similarity metrics for ontology alignment. In *International semantic web conference*, pages 294–309. Springer.
- da Silva, J., Baião, F. A., and Revoredo, K. (2016). Alinhamento interativo de ontologias usando anti-padrões de alinhamento: Um primeiro experimento. In *Anais Principais do XII Simpósio Brasileiro de Sistemas de Informação*, pages 208–215. SBC.

- Dragisic, Z., Ivanova, V., Lambrix, P., Faria, D., Jiménez-Ruiz, E., and Pesquita, C. (2016). User validation in ontology alignment. In *International Semantic Web Conference*, pages 200–217. Springer.
- Duan, S., Fokoue, A., and Srinivas, K. (2010). One size does not fit all: Customizing ontology alignment using user feedback. In *International Semantic Web Conference*, pages 177–192. Springer.
- Euzenat, J. (2014). First experiments in cultural alignment repair (extended version). In *European Semantic Web Conference*, pages 115–130. Springer.
- Euzenat, J. and Rousset, M.-C. (2020). Semantic web. In *A Guided Tour of Artificial Intelligence Research*, pages 181–207. Springer.
- Euzenat, J., Shvaiko, P., et al. (2007). *Ontology matching*, volume 18. Springer.
- Falconer, S. M., Noy, N. F., and Storey, M.-A. D. (2007). Ontology mapping—a user survey. In *OM*. Citeseer.
- Faria, D., Pesquita, C., Santos, E., Palmonari, M., Cruz, I. F., and Couto, F. M. (2013). The agreement makerlight ontology matching system. In *OTM Confederated International Conferences "On the Move to Meaningful Internet Systems"*, pages 527–541. Springer.
- Fu, B., Noy, N. F., and Storey, M.-A. (2013). Indented tree or graph? a usability study of ontology visualization techniques in the context of class mapping evaluation. In *International Semantic Web Conference*, pages 117–134. Springer.
- Gargouri, F. and Jaziri, W. (2010). *Ontology theory, management, and design: Advanced tools and models*. Information Science Reference.
- Granitzer, M., Sabol, V., Onn, K. W., Lukose, D., and Tochtermann, K. (2010). Ontology alignment—a survey with focus on visually supported semi-automatic techniques. *Future Internet*, 2(3):238–258.
- Harrow, I., Balakrishnan, R., Jimenez-Ruiz, E., Jupp, S., Lomax, J., Reed, J., Romacker, M., Senger, C., Splendiani, A., Wilson, J., and Woollard, P. (2019). Ontology mapping for semantically enabled applications. *Drug discovery today*, 24:2068–2075.
- Li, H., Dragisic, Z., Faria, D., Ivanova, V., Jiménez-Ruiz, E., Lambrix, P., and Pesquita, C. (2019). User validation in ontology alignment: functional assessment and impact. *The Knowledge Engineering Review*, 34.
- Li, Y., Stroe, C., and Cruz, I. F. (2015). Interactive visualization of large ontology matching results. In *VOILA@ ISWC*, page 37.
- Massmann, S., Raunich, S., Aumüller, D., Arnold, P., and Rahm, E. (2011). Evolution of the coma match system. In *Proceedings of the 6th International Conference on Ontology Matching-Volume 814*, pages 49–60. CEUR-WS. org.
- Meilicke, C. (2011). *Alignment incoherence in ontology matching*. PhD thesis, Universität Mannheim.
- Nguyen, A., Gardner, L., and Sheridan, D. (2019). Towards ontology-based design science research for knowledge accumulation and evolution. In *Proceedings of the 52nd Hawaii International Conference on System Sciences*.
- Peffer, K., Tuunanen, T., Rothenberger, M. A., and Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of management information systems*, 24(3):45–77.

- Pesquita, C., Faria, D., Santos, E., and Couto, F. M. (2013). To repair or not to repair: reconciling correctness and coherence in ontology reference alignments. In *Proc. 8th ISWC ontology matching workshop (OM), Sydney (AU)*, page this volume.
- Santos, M. O., Mello, C. E. R., and Classe, T. M. (2020). A useful tool to support the ontology alignment repair. In *Brazilian Conference on Intelligent Systems*, pages 201–215. Springer.
- Shadish, W. R., Cook, T. D., and Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin,.
- Vom Brocke, J., Winter, R., Hevner, A., and Maedche, A. (2020). Special issue editorial—accumulation and evolution of design knowledge in design science research: A journey through time and space. *Journal of the Association for Information Systems*, 21(3):9.