

## Otimização de Método de Alinhamento de Ontologias

Roberto Monteiro Dias<sup>1</sup>, Fernanda Araújo Baião<sup>1</sup>, Kate Revoredo<sup>1</sup>

<sup>1</sup>Programa de Pós-Graduação em Informática (PPGI) – Universidade Federal do Rio Estado do Rio de Janeiro (UNIRIO)  
Av. Pasteur 458, Rio de Janeiro, Brasil

roberto.dias@uniriotec.br, fernanda.baiao@uniriotec.br, katerevoredo@uniriotec.br

**Previsão de Defesa:** dezembro de 2017

**Abstract.** *With the advancement of the Semantic Web, libraries were created that store collections of diverse ontologies. These libraries allow users to identify, reuse, and publish ontologies that properly describe their domain. With the increasing number of ontologies being hosted in these libraries by users, they often need to calculate the alignment between ontologies. As new ontologies are being submitted or updated, alignment must be performed correctly and quickly. Therefore, the process of aligning ontologies becomes a challenge for these libraries. The ALIN approach proposes an interactive strategy for the alignment of ontologies of these libraries. It seeks to increase the expertise of the domain specialist through the use of anti-alignment standards. However, the order in which the matches are shown to the expert influences the use of anti-patterns. Then, if the correct alignments are presented to the specialist earlier, more inconsistent alignments will be removed with the use of anti-patterns. This can lead to a large decrease in the number of interactions with this expert and, at the same time, increase the effect of the propagation of their responses, which may contribute to the improvement of the quality of the alignments generated.*

**Resumo.** *Com o avanço da Web Semântica, foram criadas bibliotecas que armazenam coleções de diversas ontologias. Estas bibliotecas permitem que os usuários possam identificar, reusar e publicar ontologias que descrevam o seu domínio de forma apropriada. Com o aumento da quantidade de ontologias sendo hospedadas nessas bibliotecas pelos usuários, estas precisam, muitas vezes, calcular o alinhamento entre as ontologias. À medida que novas ontologias vão sendo submetidas ou são atualizadas, o alinhamento deve ser realizado de forma correta e rápida. Portanto, o processo de alinhamento de ontologias torna-se um desafio para essas bibliotecas. A abordagem ALIN propõe uma estratégia interativa para o alinhamento de ontologias destas bibliotecas. Ela procura aumentar a participação do especialista do domínio através da utilização de anti-padrões de alinhamento. Porém, a ordem em que as correspondências são mostradas ao especialista influencia na utilização dos anti-padrões. Então, se forem apresentados inicialmente os alinhamentos corretos ao especialista, mais alinhamentos inconsistentes serão retirados com o uso de anti-padrões. O que pode gerar uma grande diminuição no número de interações com este especialista e ao mesmo tempo, aumentar o efeito da propagação de suas respostas, o que pode contribuir para melhoria da qualidade dos alinhamentos gerados.*

### 1. Introdução

Com o avanço da Web Semântica, foram criadas bibliotecas que armazenam coleções de diversas ontologias. Estas bibliotecas permitem que os usuários possam identificar, reusar e publicar ontologias que descrevam o seu domínio de forma apropriada [1]. Com o aumento do número de ontologias sendo hospedadas nessas bibliotecas pelos usuários, estas precisam, muitas vezes, calcular o alinhamento entre as ontologias. À medida que novas ontologias vão

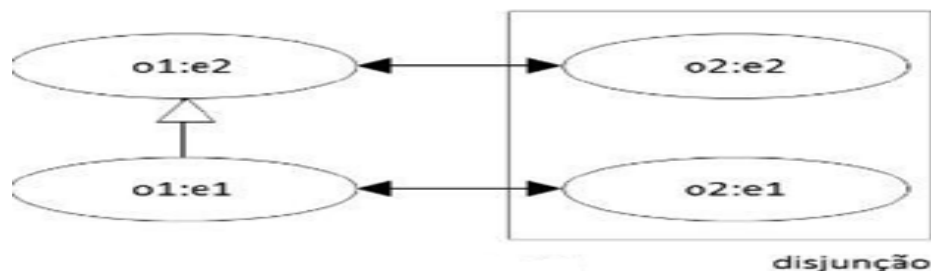
sendo submetidas ou são atualizadas, o alinhamento deve ser realizado de forma apropriada e rápida. Assim, o processo de alinhamento de ontologias torna-se um desafio para essas bibliotecas [2]. Em cenários como estes, em que diferentes ontologias que descrevem o mesmo domínio precisam interoperar, podem-se encontrar diversos problemas, como, por exemplo, a incompatibilidade semântica gerando resultados imprevisíveis, ambíguos ou até mesmo incompletos, o que vai contra ao propósito da Web Semântica [3]. Então, como resolver os problemas de alinhamento de ontologias nessas bibliotecas de forma mais correta e rápida possível?

## 2. Processo de Alinhamento de Ontologias

Pode-se resolver este problema através do processo de alinhamento de ontologias, mas ainda existem muitos erros relacionados aos alinhamentos gerados [4]. Pode-se destacar algumas limitações, como por exemplo, as abordagens automáticas têm problemas de detecção dos significados dos termos de forma precisa semanticamente e a abordagem manual é extremamente trabalhosa e bastante demorada. Uma abordagem que é bastante comentada na literatura e tem tido bons resultados devido a sua eficiência, é a abordagem interativa. Ela considera a participação do especialista do domínio com o propósito de gerar um alinhamento de excelente qualidade. Porém, mesmo nesta abordagem, também pode-se encontrar erros nos alinhamentos gerados devido ao mal-uso do especialista do domínio. Uma técnica que tem sido bastante utilizada para tornar eficiente a participação do especialista é o uso de anti-padrões que é uma combinação de correspondências que geram uma insatisfabilidade de um conceito [5].

## 3. Anti-padrões de Alinhamentos de Ontologias

Um Anti-padrão de alinhamento reúne um conjunto de correspondências que, em conjunto com os relacionamentos das próprias ontologias (hierarquia de conceitos, por exemplo), necessariamente gera uma inconsistência lógica no alinhamento [4].



**Figura 1.**Exemplo de anti-padrão das ontologias O1 e O2 ( Disjunção e Generalização) [4].

O alinhamento da Figura 1, por exemplo, é inconsistente logicamente, pois cada instância da classe o1:e1 é também instância da classe o2:e1 (pela equivalência, indicada pela seta de duas pontas), e cada instância de o1:e1 também é instância de o1:e2 (pela herança).

No exemplo da figura 2 abaixo, é apresentado dois pares de entidades e1 e e2 de duas ontologias do domínio de Conferências (OAEI) que se desejam alinhar:

Id	e1	e2	Inconsistência Lógica
4	Author	Regular_author	31
23	Paper	Paper	31
31	Paper_Abstract	Abstract	

**Figura 2:** Pares de entidades de duas ontologias do domínio de Conferências (OAEI).

Ao passar pela interação com o especialista, os pares de entidades de **e1** e **e2** de **id = 31** são considerados correspondentes, retirando automaticamente, através do uso de anti-padrões, os pares de correspondência que são inconsistentes logicamente no alinhamento.

#### 4. Apresentação do Problema

A abordagem ALIN é uma estratégia interativa para o processo de alinhamento de ontologias que integra o uso de classificadores e clusterizadores com anti-padrões a cada interação com o especialista [4]. A utilização dos anti-padrões ocorre na ocasião da interação com o especialista. Se o especialista responder “sim”, indicando que um par de entidades é uma correspondência, todos os pares incompatíveis com ele, de acordo com os anti-padrões, serão classificados como “não”, indicando que estas correspondências não pertencem ao alinhamento. Isso amplifica as respostas dadas pelo especialista na interação, podendo-se saber a resposta para várias correspondências com somente uma pergunta ao especialista [4]. Porém, a ordem em que as correspondências são mostradas ao especialista influencia na utilização dos anti-padrões. Então, se forem apresentados na primeira interação os alinhamentos corretos ao especialista, mais alinhamentos inconsistentes serão retirados com o uso de anti-padrões. Então, como podemos otimizar a interação do especialista com o uso de anti-padrões na abordagem ALIN?

#### 5. Proposta de Solução

Como a ordem em que as correspondências são mostradas ao especialista influem na eficiência da utilização dos anti-padrões, entende-se que, se as correspondências corretas forem mostradas nas primeiras interações com o especialista, mais alinhamentos inconsistentes serão retirados com a utilização de anti-padrões.



**Figura 2. Fluxograma de Interação com o Especialista do Domínio.**

Então, a proposta de solução é um método que estenda o ALIN para que sejam apresentados na interação com o especialista somente aquelas correspondências que influenciem diretamente na eficiência da utilização da técnica de anti-padrões de tal forma que tenhamos um alinhamento com menor número de interações com especialista. O que pode gerar um alinhamento em menor tempo de execução por aumentar o efeito da participação do especialista somente sobre os pares de entidades que sejam realmente relevantes.

#### 6. Projeto de Avaliação

Será realizado um experimento utilizando todos os datasets disponibilizados pela OAEI (conference, anatomy e largebio). A abordagem vai calcular o número de interações com o especialista e os resultados em termos de precisão, cobertura e medida-F para cada par de ontologias.

## 7. Referencias

- [1] M. d'Aquin and N. F. Noy, "Where to publish and find ontologies? A survey of ontology libraries," *Web Semant. Sci. Serv. Agents World Wide Web*, vol. 11, pp. 96–111, 2012.
- [2] U. Thayasivam and P. Doshi, "Speeding up batch alignment of large ontologies using MapReduce," *Proc. - 2013 IEEE 7th Int. Conf. Semant. Comput. ICSC 2013*, pp. 110–113, 2013.
- [3] N. Noy, "Ontology Mapping and Alignment," *3rd Summer Sch. Ontol Eng.*, p. 48, 2005.
- [4] J. Da Silva, F. A. Baião, and K. Revoredo, "Alinhamento Interativo de Ontologias usando Anti- Padrões de Alinhamento: Um Primeiro Experimento Alternative Title: Interactive Ontology Alignment using Alignment Antipatterns: A First Experiment," 2016.
- [5] C. Meilicke, "Alignment Incoherence in Ontology Matching," *Univ. Mannheim*, p. 190, 2011.