

Uma metodologia gamificada para construção de ontologias, apoiada por um Chatbot.

Adriano Ferraz
Sistemas de Informação
UEG - Universidade Estadual de Goiás
Santa Helena-GO, Brasil
adriano.ferraz@ueg.br

Fred Freitas
Centro de Informática - Cin
Universidade Federal de Pernambuco - UFPE
Recife-PE, Brasil
fred@cin.ufpe.br

Resumo—Ontologias são o alicerce da Web Semântica, criadas na IA simbólica, são uma maneira formal de representar conhecimento que permite reuso, raciocínio e inferências lógicas. Diversos domínios de conhecimento têm se favorecido do uso de ontologias. Porém, criar ontologias ainda não é uma tarefa fácil e metodologias tradicionais ainda são as mais utilizadas. Em tarefas que exigem maior conhecimento técnico a gamificação demonstrou ser um método facilitador. Este trabalho propõe o uso de uma metodologia gamificada apoiada por um Chatbot para construção de ontologias.

Keywords-Ontologia; Chatbot; Gamificação; Representação de conhecimento;

I. INTRODUÇÃO

Ontologias têm sido amplamente utilizadas e se demonstraram importantes nos mais diversos domínios [1] [2]. Ontologias permitem interoperabilidade entre softwares, reuso e raciocínio sobre o conhecimento representado. No entanto, criar ontologias ainda não é uma tarefa trivial, que exige conhecimento de uma linguagem formal de construção de ontologias.

Ontologia é um dos principais componentes da pilha de tecnologias que compõem a Web Semântica[3]. Por meio de triplas que representam classes, instâncias e relações é possível construir um grafo de conhecimento sobre um determinado domínio.

Devido a grande quantidade de dados disponíveis na web e a necessidade de interoperabilidade entre softwares, é necessário estruturar conhecimento para que ele possa ser disponibilizado em uma linguagem formal. Por isso, criar ferramentas ou metodologias que permitam que especialistas de um domínio representem conhecimento é fundamental [4].

Construir boas ontologias vai além da construção de triplas [5]. Boas ontologias são construídas com o uso de axiomas lógicos que proveem quantificadores existenciais ou universais, disjunções, conjunções, complemento, entre outros. O uso destes operadores amplia a expressividade da ontologia e permite checagem de consistência e descoberta de novos fatos [6].

Embora modelar conhecimento e torna-lo explícito através de ontologias seja vantajoso, criar ontologias ainda é uma tarefa onerosa que depende não só de conhecimento do domínio a ser representado, mas de ferramentas e técnicas de modelagem que permitam uma representação coerente e organizada. O reuso de outras ontologias também não é uma tarefa trivial, pois exige conhecimento da ontologia a ser reutilizada [7].

Propomos neste artigo uma metodologia apoiada por um Chatbot para que um especialista de domínio construa uma ontologia a partir de um diálogo em linguagem natural. Interagindo com um Chatbot em um processo gamificado, conseguimos converter sentenças em linguagem natural para axiomas lógicos em linguagem OWL.

Arandu¹, o nome do Chatbot desenvolvido, encoraja o usuário a informar definições que serão transformadas em axiomas lógicos na ontologia em desenvolvimento. Arandu sugere a inserção de quantificadores existenciais ou universais para as relações informadas. Utilizando uma ontologia superior, ou ontologia de topo, Arandu sugere classes que possam ser reutilizadas, isto amplia a cobertura da ontologia em desenvolvimento. Arandu também faz checagem de consistência e informa ao usuário caso haja alguma definição contraditória.

Contudo, utilizando a abordagem epistemológica *Design Science Research - DSR* [8], validamos o artefato de software desenvolvido. Arandu, de forma gamificada guia o usuário no processo de construção de uma ontologia utilizando linguagem natural e fazendo reuso de uma ontologia superior.

Este artigo segue a seguinte estrutura: na seção II apresentamos alguns fundamentos teóricos, na seção III temos os principais trabalhos relacionados, na seção IV apresentamos a metodologia, na seção V mostramos um exemplo de construção de uma ontologia utilizando a abordagem proposta, na seção VI apresentamos os resultados, e por fim na seção VII apresentamos a conclusão e trabalhos futuros.

¹Arandu é uma palavra do dicionário *Guarani* que significa aquele que sabe, que tem conhecimento.

II. FUNDAMENTAÇÃO TEÓRICA

A. Ontologia

A palavra Ontologia é uma palavra derivada do dicionário grego *ontos*, ser, e *logos*, palavra. Os primeiros estudos sobre Ontologias foram iniciados por Aristóteles, em um contexto filosófico onde ele atribuía esse termo a um ramo da Metafísica para classificar as coisas, ou seja, como uma ciência para descrever “o ser”, isto é, o estudo dos atributos que pertencem às coisas por causa de sua própria natureza [9].

Em Ciência da Computação, uma Ontologia define um conjunto de primitivas representacionais específicas para modelar um domínio do conhecimento. Estas primitivas representacionais são classes, instâncias e relações que formam um grafo, onde os nós representam classes ou instâncias e os vértices representam as relações. Ontologias têm se tornado objeto de estudo em diversos domínios, com o objetivo de representar conhecimento. A criação de Ontologias permite que sejam construídas bases de dados e de conhecimento compartilháveis e reutilizáveis [10].

Segundo Guarino [5] uma Ontologia pode ser classificada em três níveis:

- *Ontologia de Topo* - As ontologias de Topo ou superiores descrevem conceitos gerais como espaço, tempo, matéria, objeto, evento, ação, etc. Elas são independentes de um determinado problema ou domínio.
- *Ontologia de domínio* - São Ontologias que expressam um vocabulário relacionado a um domínio genérico (como medicina ou automóveis) ou uma tarefa ou atividade genérica (como diagnosticar ou vender), especializando os termos de uma Ontologia de alto nível. Segundo [11] as Ontologias de domínio cobrem dúzias de áreas específicas, incluindo governo mundial, finanças e economia e vírus biológicos entre outros.
- *Ontologias de Aplicação* - São Ontologias apropriadas para uma aplicação em questão, que fazem sentido no contexto da aplicação. São pouco reusáveis por terem um objetivo bem específico.

Uma ontologia de topo ou uma ontologia superior, descreve conceitos gerais que são independentes de domínio.

A ontologia *Suggested Upper Merged Ontology* - SUMO é uma ontologia de topo. Com cerca de 1000 termos e 4000 axiomas usando declarações SUO-KIF [11].

SUO-KIF é uma variante da linguagem KIF (*Knowledge Interchange Format*), uma linguagem declarativa capaz de expressar declarações arbitrárias em lógica de primeira ordem com suporte a raciocinadores.

Algumas ontologias são de domínio, mas por serem tão abrangentes em alguns domínios, podem ser utilizadas como ontologias superiores. Um exemplo é a ONTOBIO, uma ontologia de biodiversidade desenvolvida pelo Instituto Nacional de Pesquisa Amazônica (INPA) [12].

B. Processamento de Linguagem Natural - PLN

Processamento de Linguagem Natural - PLN é uma área da Ciência da Computação, dentro do campo da Inteligência Artificial, que estuda os problemas da geração e compreensão automática de línguas humanas naturais [13].

Alguns desafios de PLN envolvem a compreensão e geração de textos em linguagem natural conectando o computador e o ser humano, uma tarefa considerada computacionalmente difícil.

O PLN permite que desenvolvedores criem, por exemplo, sistemas de diálogo, resumos automáticos, sistemas de tradução de textos, reconhecimento de entidades, extração de relacionamentos, análise de sentimentos, reconhecimento de fala e segmentação de tópicos, entre outros.

O trabalho de PLN é um processo de análise, interpretação e geração da linguagem e pode ser dividido em vários estágios visualizados na Figura 1 [14].

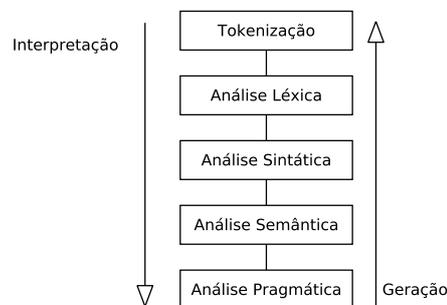


Figura 1. Estágios do Processamento de Linguagem Natural

O processo de **Tokenização** é o processo de identificação das palavras em uma sentença, também chamadas de *tokens*. Nem todas as línguas fornecem as palavras em forma de texto separadas por espaços, por exemplo a língua Chinesa. Neste caso, o processo de tokenização é um processo fundamental para a análise da sentença, exemplo na figura 2.

A **Análise Léxica** em PLN refere-se aos mecanismos para análise textual em nível de palavra, um processo conhecido da análise léxica é o processo de *lemmatização*, que consiste em reduzir a palavra ao seu radical utilizando recursos de análise linguística, exemplo a palavra *pedalando* possui *lemma* ou radical, *pedal*.

A **Análise Sintática** é a fase de PLN baseada em uma gramática. Nesse processo os *tokens* são etiquetados com suas respectivas funções na sentença, por exemplo sintagma verbal, sintagma nominal, substantivo, verbo, adjetivo, adverbio, etc. Esta análise gera uma árvore hierárquica permitindo uma futura análise semântica, exemplo na Figura 2.

Após o processo de análises léxicas e sintáticas, é necessário que faça a **Análise Semântica**. Neste processo são adicionadas definições a árvore hierárquica que permitem

uma melhor compreensão da sentença. Estas definições são por exemplo a indicação de um agente e de um paciente em uma determinada sentença, exemplo na Figura 2 a relação *nsubj*.

A **Análise Pragmática** vai além da análise da estrutura de uma frase, analisando em toda a sentença, ou conjunto de frases, a compreensão do contexto que não esteja definido na frase, como por exemplo referências pronomiais.

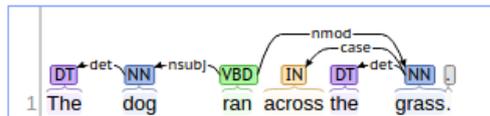


Figura 2. Exemplo de árvore Sintática.

C. Gamificação

O uso de elementos de jogos digitais em uma tarefa real pode torna-la mais lúdica, atrativa e inclusiva [15]. O uso destes elementos em tarefas cotidianas é conhecido como gamificação.

Gamificação, segundo Deterding [16], pode ser definida como o uso de elementos de design de jogos em contextos que não são jogos.

A gamificação pode ser aplicada em ambientes digitais ou não, e seus componentes e elementos vão depender do contexto que estão inseridos [17].

Utilizando elementos da gamificação como interação e *feedback*, Arandu promove uma experiência de construção da ontologia de uma forma que o usuário não se sinta isolado.

D. Design Science Research - DSR

Design Science Research é um paradigma de ciência que vem se popularizando na área de Sistemas de Informação [18]. Foi proposta por Herbert Simon, autor do livro "As Ciências do Artificial" (*The Sciences of the Artificial*) [8].

A abordagem DSR provê rigor científico e relevância adequada para o desenvolvimento de artefatos tecnológicos [8].

Em DSR normalmente realizasse três avaliações. A primeira se o artefato satisfaz aos requisitos. A segunda se as conjecturas teóricas são válidas, e a terceira avaliação se o artefato é satisfatório [19].

III. TRABALHOS RELACIONADOS

Na literatura é possível encontrar várias metodologias para construção de ontologias, muitas inspiradas na engenharia de software com a participação de seres humanos e muitas vezes apoiadas por software. Algumas metodologias bastante conhecidas são, METHONTOLOGY [20], OTK [21], Methodology 101 [22], RapidOWL [23], etc.

Métodos caracterizados como Aprendizado de Ontologias, que criam ontologias de forma automática ou semi-automática, possuem bastante relação com este trabalho [24].

Porém, este trabalho propõe o uso de uma interface de conversação para construção de uma ontologia, convertendo texto para DL, com interação com um usuário. Diferente de [25], por conseguir reutilizar conhecimento de outra ontologia.

Usuários de software têm demonstrado interesse por utilizar interfaces de conversação [26], e utilizar estas interfaces para desenvolvimento de software podem ser promissoras.

IV. METODOLOGIA

A. Design do artefato

Arandu, por meio de um diálogo em Linguagem Natural Controlada - LNC, guia o usuário no processo de construção de uma ontologia, recomendando classes de uma ontologia superior, fazendo inferências lógicas por meio de um raciocinador e verificando inconsistências lógicas.

Para apoiar o desenvolvimento do Chatbot Arandu elaboramos um mapa apresentado na figura 3, baseado no mapa de elementos esperados uma pesquisa DSR proposta em [18].

Este mapeamento contribui para direcionar o desenvolvimento do artefato tecnológico que guiará o usuário no processo de construção da ontologia, através de conjecturas teóricas fundamentadas por um quadro teórico estabelecido por uma análise do estado da arte.

Diante deste contexto serão realizadas duas avaliações. A primeira avalia se os requisitos do artefato foram atendidos, e a segunda se este artefato contribui para a solução do problema observada nas conjecturas teóricas.

B. Arandu Chatbot

Nesta seção é apresentada a arquitetura utilizada para a construção do Chatbot Arandu. Esta arquitetura foi projetada com um acoplamento mínimo, que permite que módulos sejam substituídos ou adicionados.

Os canais de interação entre o usuário e o Chatbot Arandu podem ser os mais diversos, desde que permitam integração Web. Exemplo, páginas Web, mobile apps, etc.

Devido ao avanço dos dispositivos eletrônicos é possível que muitos aparelhos que sejam utilizados para interagir com o Chatbot possuam nativamente recursos de interação por voz, vídeo ou teclados específicos para portadores de alguma necessidade especial, que convertam a entrada do usuário para texto. Estes recursos ampliam a experiência do usuário com o Arandu.

Para que seja possível criar uma ontologia a partir de um sistema de diálogo são realizadas algumas tarefas de NLP (*Natural Language Processor*) e utilizadas APIs (*Application Programming Interface*) de manipulação de ontologias. A Figura 4 representa a arquitetura do Chatbot Arandu.

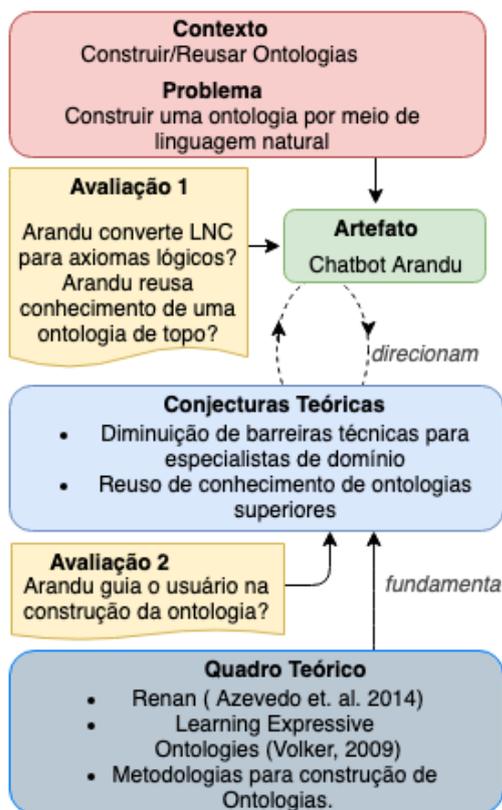


Figura 3. Mapeamento dos elementos DSR do Chatbot Arandu, baseado em [18]

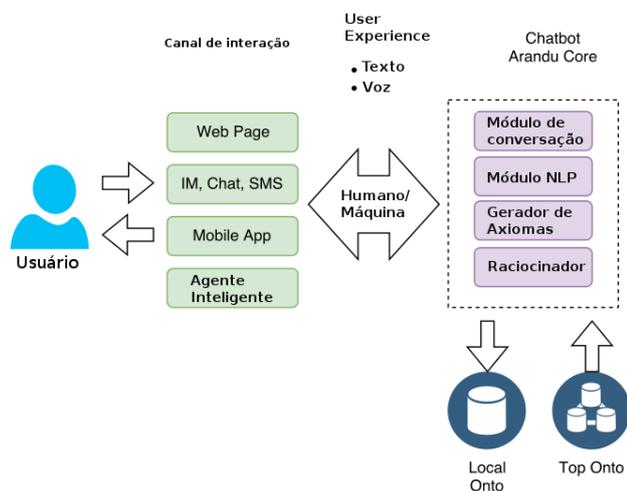


Figura 4. Anatomia do Chatbot Arandu.

C. Módulo de Conversação

Arandu, por meio do módulo de conversação, propõe uma metodologia interativa apoiada por software. Arandu combina elementos de outras metodologias já existentes com

uma proposta ágil e com foco na facilidade do usuário estruturar conhecimento.

A metodologia Arandu se enquadra no manifesto ágil pelo fato de se basear em valores, princípios e práticas.

Valores estão implicitamente contidos na proposta de construção através de linguagem natural, que garante simplicidade. Com foco no usuário, o Chatbot interage com o usuário e o encoraja a informar definições, exigindo o mínimo de conhecimento de uma linguagem formal para construção de ontologias.

Princípios estão em seu processo axiomático de construção. O processo é incremental e convergente. Através da indução de definição de classes já extraídas a construção é observável e coerente, com menor risco de falta de definições que compõem um axioma inicial. O feedback é rápido, tanto na avaliação quanto na inferência de novos axiomas por meio do raciocinador.

Práticas podem ser notadas na possibilidade de *short releases* da ontologia, que evolui a cada interação garantindo consistência desde a primeira iteração. O reuso de termos de uma ontologia de topo com suporte de anotações permite que não seja necessário usuários *experts* para conseguir reutilizar conhecimento de outra ontologia.

Uma representação visual desta metodologia pode vista na figura 5.

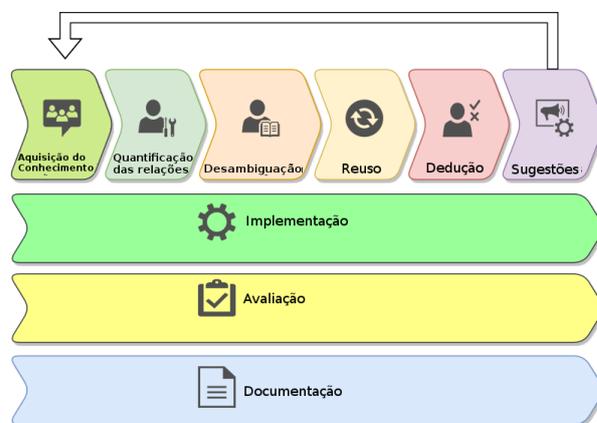


Figura 5. Metodologia de construção de Ontologias Arandu.

Esta metodologia foi construída para ser utilizada por um usuário humano e um Chatbot através de um diálogo em linguagem natural controlada, essa característica permite que algumas fases possam ser quebradas durante o diálogo. Porém, o módulo de conversação descrito anteriormente possui estratégias para tratar qualquer quebra de diálogo e manter o usuário neste fluxo.

Algumas fases podem não ocorrer, por exemplo, caso não haja conhecimento para ser reutilizado da ontologia de topo.

D. Módulo NLP

No módulo NLP são realizadas as tarefas relacionadas a NLP para extração de classes, relações e axiomas das sentenças informadas pelo usuário durante a interação com o Chatbot Arandu. Para a extração dessas relações são realizadas atividades de análise sintática e semântica das classes.

O processo de análise sintática é uma atividade de PLN que tem como objetivo analisar as palavras que compõem o discurso, seguindo regras gramaticais.

Para esta atividade o Arandu utiliza a API do *Stanford NLP Group*. [27]

Este módulo tem como principal função identificar os *tokens* da frase, identificar os sintagmas nominais que pertence cada *token* e o seu papel na árvore de dependência.

Após terem sido identificados os sintagmas a que pertencem os *tokens* que compõem o diálogo, bem como identificado a árvore de dependência que fornece a estrutura semântica do diálogo, são removidas as palavras conhecidas como *stop words*, estas são palavras como por exemplo, *a, an, the, this, etc.*

A próxima atividade é a análise semântica das classes, que consiste na identificação dos *tokens* candidatos a classes na ontologia. Os sintagmas são identificados pelas tags **NN**, **NNP**, **NNS** que representam os substantivos, **JJ** representa os adjetivos, **VB**, **VBZ**, **VBP**, **VBN**, **VPG** representam os verbos e a tag **NP** identifica o sujeito da frase.

Para que um *token* seja candidato a uma classe ele deve se encaixar em uma das seguintes regras:

- Se o sintagma a que pertence o *token* for um **NN**, **NNP** ou **NNS**.
- Ser composição de um sintagma **NP**

Dada uma sentença como exemplo: S1= "The orange is the fruit of the citrus species.", os *tokens* "orange" e "fruit" atendem a essas regras e então se tornam candidatos a classes na ontologia. Porém analisando o restante da sentença, identificamos os seguintes *tokens* "citrus" e "species". Ambos atendem as regras definidas, porém pertencem ao mesmo sintagma **NP** e analisando o contexto da sentença é notável que o objetivo é que estes dois *tokens* sejam concatenados e formem um *token* único "citrusSpecies" e esse sim seja um candidato a classe na ontologia. Isso foi tratado com a adição da seguinte regra.

- Se um *token* com sintagma **NN**, **NNP** ou **NNS** for precedido de um *token* com sintagma **JJ**, **NN**, **NNP** ou **NNS**, pertencente ao mesmo sintagma **NP** eles devem ser concatenados.

Após realizada a etapa de identificação das classes que irão compor a ontologia, é necessário identificar as relações que possuem estas classes. As relações são os verbos, e possuem sintagmas **VB**, **VBZ**, **VBP**, **VBN**, **VPG**.

Poderão ocorrer dois tipos de relações. Relação de **subsunção** e a relação de **equivalência**.

A relação de subsunção é a relação em que uma classe se torna subclasse de outra classe. Em linguagem natural essa relação é dada pelo verbo *be*.

V. CONSTRUÇÃO DA ONTOLOGIA

O primeiro passo para a construção é definir a IRI e inserir uma definição para a ontologia. Um exemplo de diálogo para esse passo pode ser visto na tabela I.

Tabela I
EXEMPLO DE CONSTRUÇÃO DE IRI.

Arandu :	Whats the name of your ontology?
User :	Cerrado Ontology
Arandu :	I can create your IRI thats "www.arandu.com.br/ontologies/example/cerrado_ontology" ? (yes/no)
User :	yes
Arandu :	Ok! get me a description about your ontology:
User :	Its a domain ontology to describe the ecosystem cerrado.

A. Aquisição do Conhecimento

O processo de aquisição do conhecimento é realizado através da extração de classes e relações das sentenças informadas pelo usuário em linguagem natural, que serão convertidas para axiomas em DL. Na tabela II temos um exemplo de um axioma de hierarquia de classes.

Tabela II
EXEMPLO DE CONSTRUÇÃO DE SUBCLASSES.

User:	Cerrado is an ecosystem and terrestrial.
-------	--

Na tabela III temos um exemplo de extração de uma relação do tipo *Object Property*.

Tabela III
EXEMPLO DE CONSTRUÇÃO DE RELAÇÕES.

User:	Cerrado has plants and animals.
-------	---------------------------------

Arandu também permite a construção de axiomas de disjunção. Na tabela IV temos um possível axioma que pode ser informado pelo usuário.

Tabela IV
EXEMPLO DE DISJUNÇÃO.

User:	Cerrado has no Sea.
-------	---------------------

B. Quantificação das relações

Quando forem encontradas relações no axioma que possam ser quantificadas, o Chatbot questionará o usuário qual o quantificador que ele deseja inserir. O usuário poderá optar pelo quantificador existencial *some* ou pelo quantificador universal *only*. Um exemplo deste diálogo pode ser visto na tabela V.

Tabela V
EXEMPLO DE QUANTIFICAÇÃO DE RELAÇÕES.

Arandu:	<i>"Cerrado have only plants and animals" or other things? (some/only)</i>
User:	<i>some</i>

C. Desambiguação

O Chatbot Arandu fará uma busca na ontologia de topo por classes que foram extraídas da sentença informada pelo usuário, e caso algumas dessas classes extraídas da sentença sejam encontradas na ontologia de topo, o Chatbot Arandu mostrará a definição desta classe obtida e o usuário confirmará com sim ou não se esta classe é compatível com o contexto da ontologia que está sendo criada. Essa desambiguação é essencial para que se possa reutilizar conhecimento da ontologia de topo. A tabela VI exemplifica esse diálogo.

Tabela VI
EXEMPLO DE DESAMBIGUAÇÃO DE CLASSES.

Arandu:	<i>Plant is: An Organism having cellulose cell walls, growing by synthesis of Substances, generally distinguished by the presence of chlorophyll, and lacking the power of locomotion.</i>
User:	<i>yes</i>

D. Reuso

Se for encontrado alguma classe na ontologia de topo que possa ser reutilizada o Chatbot Arandu irá sugerir superclasses, ou outros axiomas, para que o usuário adicione na ontologia local que está sendo criada. Poderão ser sugeridas pelo Arandu uma lista de superclasses ou axiomas encontrados na ontologia de topo.

A ontologia SUMO, por exemplo, pode fornecer as superclasses *Physical*, *Agent*, *OrganicThing*, *Object*, *CorpuscularObject*, *OrganicObject*, *Organism* para a classe *Plant*. Parte do código² OWL fornecido por SUMO, pode ser visto na figura 6.

E. Dedução

Por meio de um raciocinador em DL, o Chatbot poderá inferir algum conhecimento que esteja implícito e sugerir ao usuário para que ele adicione este axioma na ontologia.

F. Sugestão

Para fomentar o usuário a informar mais definições o Chatbot Arandu antes de pedir uma nova definição, ele perguntará ao usuário definições sobre as classes já informadas que não foram encontradas na ontologia de topo.

²Disponível em: <http://sigma.ontologyportal.org:8080/sigma/OWL.jsp?kb=SUMO&term=Plant>

```

<rdf:RDF xmlns="http://www.ontologyportal.org/SUMO.owl#" xmlns:wnd="http://www.ontologyportal.org/2002/07/owl#" xmlns:xsd="http://www.w3.org/2001/XMLSchema#" xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:owl="http://www.w3.org/2002/07/owl#" xml:base="http://www.ontologyportal.org/SUMO.owl" >
  <owl:Ontology rdf:about="http://www.ontologyportal.org/SUMO.owl" >
    <rdf:comment xml:lang="en">
      A provisional and necessarily lossy translation to OWL. Please see www.ontologyportal.org for more information. This
      software is released under the GNU Public License www.gnu.org.
    </rdf:comment>
    <rdf:comment xml:lang="en">Produced on date: Fri Jan 11 16:49:59 UTC 2011
    </owl:Ontology>
    <owl:Class rdf:about="#Plant">
      <rdf:isDefinedBy rdf:resource="http://www.ontologyportal.org/SUMO.owl"/>
      <rdf:subClassOf rdf:resource="#SelfConnectedObject"/>
      <rdf:subClassOf rdf:resource="http://www.w3.org/2002/07/owl#Thing"/>
      <rdf:subClassOf rdf:resource="#Physical"/>
      <rdf:subClassOf rdf:resource="#Agent"/>
      <rdf:subClassOf rdf:resource="#OrganicThing"/>
      <rdf:subClassOf rdf:resource="#Object"/>
      <rdf:subClassOf rdf:resource="#CorpuscularObject"/>
      <rdf:subClassOf rdf:resource="#OrganicObject"/>
      <rdf:subClassOf rdf:resource="#Organism"/>
    </owl:Class>
    <rdf:comment xml:lang="en">
      An Organism having cellulose cell walls, growing by synthesis of Substances,
      lacking the power of locomotion.
    </rdf:comment>
  </rdf:RDF>

```

Figura 6. Parte do código OWL da classe *Plant* em SUMO.

G. Implementação

O processo de implementação é a codificação destes axiomas em uma linguagem formal. O Chatbot Arandu converte os axiomas para linguagem OWL 2 garantindo expressividade *ALC*.

H. Avaliação

A avaliação é realizada através da checagem de consistência dos axiomas pelo raciocinador. Caso alguma inconsistência seja encontrada na ontologia, o Chatbot Arandu informará ao usuário o axioma que provocou esta inconsistência.

I. Documentação

A documentação é realizada durante todo o processo de construção da ontologia, através da análise dos logs da interação, gerando um documento com as definições informadas, classes e relações extraídas, axiomas gerados, e axiomas reutilizados.

VI. RESULTADOS

Para testar o Arandu propomos uma ontologia de domínio para representar um ecossistema conhecido como cerrado³. Escolhemos este domínio por ser de simples representação e por permitir alinhamento com as ontologias SUMO e Ontobio.

A tabela VII apresenta seis sentenças criadas para simular um diálogo com o Arandu.

Foram executados testes sem o reuso de uma ontologia de topo, e reutilizando classes de SUMO e de Ontobio separadamente. Na tabela VIII temos as métricas obtidas para a ontologia de domínio proposta.

De acordo com as avaliações propostas no mapeamento DSR proposto na figura 3, observamos os seguintes resultados:

³<https://en.wikipedia.org/wiki/Cerrado>

Tabela VII
SENTENÇAS CRIADAS PARA GERAR A ONTOLOGIA DE DOMÍNIO
CERRADO ONTOLOGY.

S1	Cerrado is an ecosystem and terrestrial.
S2	Cerrado has plants and animals.
S3	Cerrado has no Sea.
S4	Pequizeiro is a tree that has Pequi.
S5	Pequi is a fruit.
S6	Jaguatirica is an animal that lives in the cerrado.

Tabela VIII
MÉTRICAS DA ONTOLOGIA DE DOMÍNIO CERRADO ONTOLOGY.

Métricas	sem ontologia de topo	SUMO	Ontobio
<i>Axiom</i>	23	81	27
<i>Logical axiom count</i>	10	50	12
<i>Class count</i>	11	29	13
<i>SubClassOf axiom count</i>	10	50	12
<i>Object Property count</i>	2	2	2
<i>DL expressivity</i>	<i>ACC</i>	<i>ACC</i>	<i>ACC</i>

Avaliação 1: Arandu converte LNC para axiomas lógicos? Arandu reusa conhecimento de uma ontologia de topo?: analisando a tabela VIII temos as métricas obtidas para a conversão das sentenças propostas na tabela VII para linguagem OWL. Arandu satisfaz aos requisitos de software propostos convertendo LNC para axiomas lógicos e demonstrou capacidade de reuso de termos de duas ontologias de topo. Foram realizados três testes, um sem o uso de uma ontologia de topo, um utilizando a ontologia SUMO e outro teste utilizando a ontologia Ontobio.

Avaliação 2: Arandu guia o usuário na construção da ontologia?: a partir dos diálogos obtidos nas tabelas I, II, III, IV, V, Arandu guia o usuário no processo de construção de uma ontologia de acordo com a metodologia proposta na figura 5.

VII. CONCLUSÃO

Os resultados alcançados demonstram a capacidade de converter um diálogo em linguagem natural para axiomas lógicos. O reuso de termos é um fator importante no processo de construção de uma ontologia, porém exige que o engenheiro de ontologias conheça a ontologia a ser reutilizada.

Arandu demonstrou ser possível reutilizar conhecimento de outra ontologia de maneira fácil, este reuso além de ampliar o modelo esperado, mapeia ontologias de domínio para ontologias superiores, isso permite interoperabilidade entre sistemas que utilizem estas ontologias.

A metodologia proposta apoiada por um Chatbot e com elementos de gamificação, permite checagem de consistência e sugestão de axiomas propostas pelo Arandu durante o processo de construção, permite que especialistas de domínio com menor conhecimento técnico construam ontologias.

Em trabalhos futuros, realizaremos testes com usuários para ampliar o módulo de conversação e adicionar novas

features que permitam maior expressividade OWL, estes testes permitirão a terceira avaliação proposta pelo método de pesquisa *Design Science Research* - DSR, que verifica se o artefato é satisfatório.

REFERÊNCIAS

- [1] R. Lima, B. Espinasse, and F. Freitas, “Ontoilper: an ontology- and inductive logic programming-based system to extract entities and relations from text,” *Knowl. Inf. Syst.*, vol. 56, no. 1, pp. 223–255, 2018.
- [2] V. Maran, G. M. Machado, A. Machado, I. Augustin, J. a. C. D. Lima, and J. P. M. d. Oliveira, “Database ontology-supported query for ubiquitous environments,” in *Proceedings of the 23rd Brazilian Symposium on Multimedia and the Web*, ser. WebMedia '17. New York, NY, USA: ACM, 2017, pp. 185–188. [Online]. Available: <http://doi.acm.org/10.1145/3126858.3131575>
- [3] T. Berners-Lee, J. Hendler, and O. Lassila, “The semantic web,” *Scientific American*, vol. 284, no. 5, pp. 34–43, May 2001. [Online]. Available: <http://www.sciam.com/article.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21>
- [4] C. M. de Oliveira Rodrigues, F. L. G. de Freitas, E. F. S. Barreiros, R. R. de Azevedo, and A. T. de Almeida Filho, “Legal ontologies over time: A systematic mapping study,” *Expert Systems with Applications*, vol. 130, pp. 12 – 30, 2019. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0957417419302398>
- [5] N. Guarino, *Formal Ontology in Information Systems: Proceedings of the 1st International Conference June 6-8, 1998, Trento, Italy*, 1st ed. Amsterdam, The Netherlands, The Netherlands: IOS Press, 1998.
- [6] F. Freitas and J. Otten, “A connection calculus for the description logic ALC,” in *Canadian Conference on AI*, ser. Lecture Notes in Computer Science, vol. 9673. Springer, 2016, pp. 243–256.
- [7] E. Blomqvist, K. Hammar, and V. Presutti, “Engineering ontologies with patterns - the extreme design methodology,” in *Ontology Engineering with Ontology Design Patterns* .:, ser. Studies on the Semantic Web, 2016, no. 25, pp. 23–50.
- [8] H. A. Simon, *The Sciences of the Artificial (3rd Ed.)*. Cambridge, MA, USA: MIT Press, 1996.
- [9] N. Guarino, D. Oberle, and S. Staab, “What is an ontology?” in *Handbook on Ontologies*, 2nd ed., S. Staab and R. Studer, Eds. Springer, 2009.
- [10] T. R. Gruber, “Toward principles for the design of ontologies used for knowledge sharing,” *Int. J. Hum.-Comput. Stud.*, vol. 43, no. 5-6, pp. 907–928, Dec. 1995. [Online]. Available: <http://dx.doi.org/10.1006/ijhc.1995.1081>
- [11] A. Pease, *Ontology: A Practical Guide*. Angwin, CA: Articulate Software Press, 2011.
- [12] A. Albuquerque, J. Campos dos Santos, and A. N. de Castro, “Ontobio: A biodiversity domain ontology for amazonian biological collected objects,” 01 2015, pp. 3770–3779.

- [13] E. D. Liddy, “Natural Language Processing,” in *Encyclopedia of Library and Information Science*, 2nd ed. NY: Marcel Decker, Inc., 2001.
- [14] R. Vieira and Vera, “Linguística computacional: princípios e aplicações,” in *As Tecnologias da informação e a questão social*. Ana Teresa Martins and Dibio Leandro Borges, 2001.
- [15] A. Brazil and L. Baruque, “Gamificação aplicada na graduação em jogos digitais,” in *SBIE. Simpósio Brasileiro de Informática na Educação*, 10 2015, p. 677.
- [16] S. Deterding, D. Dixon, R. Khaled, and L. Nacke, “From game design elements to gamefulness: Defining “gamification”,” in *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, ser. MindTrek '11. New York, NY, USA: ACM, 2011, pp. 9–15. [Online]. Available: <http://doi.acm.org/10.1145/2181037.2181040>
- [17] M. Bacelar, D. L. Costa, D. Abrantes, and F. Albert, “Revisão bibliográfica dos aspectos e métodos componentes da gamificação na educação,” in *SBC – Proceedings of SBGames 2018*, 11 2018, p. 1232.
- [18] M. PIMENTEL, D. FILIPPO, and F. M. SANTORO, *Metodologia de Pesquisa em Informática na Educação: Concepção da Pesquisa*. Porto Alegre: SBC, 2019, ch. Design Science Research: fazendo pesquisas científicas rigorosas atreladas ao desenvolvimento de artefatos computacionais projetados para a educação.
- [19] D. P. A. J. J. A. V. DRESCH, Aline; LACERDA, *Design Science Research: método de pesquisa para avanço da ciência e tecnologia*. Porto Alegre: Bookman, 2015.
- [20] M. Fernandez-Lopez, A. Gomez-Perez, and N. Juristo, “Methontology: from ontological art towards ontological engineering,” in *Proceedings of the AAAI97 Spring Symposium*, Stanford, USA, March 1997, pp. 33–40.
- [21] H. Akkermans, J. Broekstra, J. Davies, Ying, Ding, A. Duke, R. Engels, D. Fensel, I. Horrocks, V. Iosif, A. Kampman, A. Kiryakov, M. Klein, T. Lau, D. Ognyanov, U. Reimer, K. I. Simov, Rudi, D. Studer, J. van der Meer, and F. van Harmelen, “On-to-knowledge : Semantic web enabled knowledge management,” 2002.
- [22] N. F. Noy and D. L. McGuinness, “Ontology development 101: A guide to creating your first ontology,” Tech. Rep., 2001.
- [23] S. Auer and H. Herre, *RapidOWL — An Agile Knowledge Engineering Methodology*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 424–430. [Online]. Available: https://doi.org/10.1007/978-3-540-70881-0_36
- [24] J. Volker, *Learning Expressive Ontologies: Volume 2 Studies on the Semantic Web*. Amsterdam, The Netherlands, The Netherlands: IOS Press, 2009.
- [25] R. R. D. Azevedo, F. Freitas, R. G. C. Rocha, J. A. A. D. Menezes, C. M. D. O. Rodrigues, and G. D. F. P. e. Silva, “An approach for learning and construction of expressive ontology from text in natural language,” in *Proceedings of the 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT) - Volume 01*, ser. WI-IAT '14. Washington, DC, USA: IEEE Computer Society, 2014, pp. 149–156. [Online]. Available: <http://dx.doi.org/10.1109/WI-IAT.2014.28>
- [26] A. Følstad and P. B. Brandtzæg, “Chatbots and the new world of hci,” *interactions*, vol. 24, no. 4, pp. 38–42, Jun. 2017. [Online]. Available: <http://doi.acm.org/10.1145/3085558>
- [27] D. Klein and C. D. Manning, “Accurate unlexicalized parsing,” in *Proceedings of the 41st Annual Meeting on Association for Computational Linguistics - Volume 1*, ser. ACL '03. Stroudsburg, PA, USA: Association for Computational Linguistics, 2003, pp. 423–430. [Online]. Available: <http://dx.doi.org/10.3115/1075096.1075150>